Spring 3-13-2023

AN EXPLORATORY STUDY OF ELEMENTARY SCHOOL STUDENTS’ READING PERFORMANCE SCORES BEFORE AND AFTER COVID-19

Timothy Goodman

Follow this and additional works at: https://firescholars.seu.edu/coe

Part of the Educational Assessment, Evaluation, and Research Commons, Elementary Education Commons, and the Language and Literacy Education Commons

Recommended Citation

This Dissertation is brought to you for free and open access by FireScholars. It has been accepted for inclusion in Doctor of Education (Ed.D) by an authorized administrator of FireScholars. For more information, please contact firescholars@seu.edu.
AN EXPLORATORY STUDY OF ELEMENTARY SCHOOL STUDENTS’ READING
PERFORMANCE SCORES BEFORE AND AFTER COVID-19

By
TIMOTHY E. GOODMAN

A doctoral dissertation submitted to the
College of Education
in partial fulfillment of the requirements
for the degree Doctor of Education
in Curriculum and Instruction

Southeastern University
March 2023
AN EXPLORATORY STUDY OF ELEMENTARY SCHOOL STUDENTS’ READING PERFORMANCE SCORES BEFORE AND AFTER COVID-19

by

TIMOTHY E. GOODMAN

Dissertation Approved:

__________________________
Patty LeBlanc, PhD, Dissertation Chair

__________________________
Thomas Gollery, EdD, Committee Member

__________________________
Kari Allen, EdD, Committee Member

__________________________
Lisa Coscia, EdD, Dean of the College of Education
DEDICATION

Four years ago, I was inspired to begin this journey after much prayerful consideration. I would first like to thank my Lord and Savior, Jesus Christ, for blessing me with such a supportive and understanding family. I have nothing but love and gratitude for my wife, daughter, and son, who have been with me through the extra strain and stress during my time in Southeastern University’s doctoral program. I love you more than words can tell. I am also very thankful for the support and encouragement from my church and my colleagues from Ocala Springs Elementary.
I would like to acknowledge and express my gratitude to the faculty and support staff of Southeastern University. I am very grateful that in my search for the right university to pursue my doctorate, I decided to commit to Southeastern’s doctoral program. I can truly say that it was not only worth the challenge but also that the quality of instructors, instruction, and coursework has prepared me to face the next steps in my career. During the first intensive class to begin the program, Dr. Yates, Dr. Deck, and Dr. Ingle were open and honest about the challenge and certain that God had chosen those of us in Cohort K to undertake this journey together. I would especially like to acknowledge and thank Dr. LeBlanc for guiding me through the dissertation process, praying for my family and me along the way, and helping to keep me sane throughout. I would also like to acknowledge and express my gratitude to my dissertation committee, comprising Dr. LeBlanc, Dr. Gollery, and Dr. Allen; my editor, Dr. Hoskins; and also Benjamin Whitehouse and Joelene Vining from my school district for their support.
Abstract

The purpose of this study was to explore the reading performance scores of elementary school students in one school district before and after school closures due to COVID-19. This nonexperimental, exploratory study of quantitative, archival data was designed to explore a single cohort (N = 2,006) of third- through fifth-grade students’ i-Ready reading diagnostic scale scores from five different assessment periods: before school closures in spring 2020 and four subsequent assessment periods after face-to-face instruction resumed in fall 2020. The research sample excluded exceptional student education students and English language learners who had been retained. Mean scale scores of the research cohort were compared to the school district’s 2018-2019 i-Ready norms and to the 2018-2019 i-Ready national norms for each of the five assessment periods. The results of the comparisons revealed that the research cohort’s mean reading scale scores were significantly different from the district and national norm groups’ mean reading scale scores. The research sample’s mean reading scale scores were slightly higher than those of the target district’s norm group; however, the mean scores and the effect sizes were small. The research cohort’s reading scores were significantly lower than the scores of the 2018-2019 national norm group. Although the students in the research cohort demonstrated small increments of reading progress over time, the rate of progress was not commensurate with the 2018-2019 national norm group’s rate of progress.

Keywords: COVID-19, pandemic, reading performance, reading achievement, elementary education, reading development, i-Ready
# TABLE OF CONTENTS

Dedication ....................................................................................................................... iii

Acknowledgments .......................................................................................................... iv

Abstract .......................................................................................................................... v

Table of Contents .......................................................................................................... vi

List of Tables .................................................................................................................... viii

List of Figures .................................................................................................................. ix

I. INTRODUCTION ............................................................................................................. 1

  Problem Statement ........................................................................................................ 1
  Background of the Study ................................................................................................. 2
  Theoretical Framework ................................................................................................... 3
  Purpose Statement .......................................................................................................... 8
  Research Question .......................................................................................................... 8
  Research Design ............................................................................................................. 8
  Overview of Methodology .............................................................................................. 8
    Data Collection ........................................................................................................... 8
    Population and Sample ............................................................................................... 9
    Instrumentation ........................................................................................................... 9
    Data Analyses .............................................................................................................. 10
  Ethical Considerations .................................................................................................. 11
  Limitations ..................................................................................................................... 11
  Definition of Key Terms ............................................................................................... 12
  Significance of the Study ............................................................................................... 12

II. REVIEW OF LITERATURE .......................................................................................... 14

  Pre-COVID Learning Loss ............................................................................................. 14
  COVID-19 Studies ......................................................................................................... 23
  Summary ....................................................................................................................... 47
III. METHODOLOGY ........................................................................................................ 49
   Data Collection ........................................................................................................... 49
   Sample ......................................................................................................................... 50
   Instrumentation .......................................................................................................... 50
   i-Ready Validity and Reliability ............................................................................... 52
   Data Analyses ............................................................................................................. 53

IV. RESULTS .................................................................................................................. 55
   Methods of Collecting Data ....................................................................................... 55
   Sample ......................................................................................................................... 56
   Preliminary Analyses and Results ........................................................................... 56
   Descriptive Analyses and Results ........................................................................... 57
   Inferential Analyses and Results ............................................................................. 57
   Summary ..................................................................................................................... 59

V. DISCUSSION .............................................................................................................. 61
   Review of Methodology ........................................................................................... 61
   Research Population and Sample ............................................................................. 62
   Instrumentation .......................................................................................................... 62
   Analyses ..................................................................................................................... 63
   Summary of Results ................................................................................................... 64
   Preliminary Results ................................................................................................... 64
   Inferential Results ....................................................................................................... 64
   Implications ................................................................................................................ 65
   District Norm Comparisons ...................................................................................... 70
   National Norm Comparisons .................................................................................... 72
   Mitigating Factors ...................................................................................................... 73
   Recommendations for Future Research ................................................................. 76
   Limitations .................................................................................................................. 78
   Significance of the Study ........................................................................................... 79
   References ................................................................................................................... 80
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1: Chall’s Stages of Reading Development</td>
<td>4</td>
</tr>
<tr>
<td>Table 2: Mean Difference Scores in Reading Between Spring and Fall Terms by Grade Level Using Two Metrics</td>
<td>17</td>
</tr>
<tr>
<td>Table 3: Mean MSA Scores During School Years with Frequent Unscheduled School Closures Compared to Mean MSA Scores and Adequate Yearly Progress</td>
<td>19</td>
</tr>
<tr>
<td>Table 4: Descriptive Analyses of Eleven Studies of Student Achievement Published Between March 1, 2020, and April 30, 2021</td>
<td>38</td>
</tr>
<tr>
<td>Table 5: Descriptive Statistics Summary Table</td>
<td>57</td>
</tr>
<tr>
<td>Table 6: Mean Scale Score Comparisons of the Study Cohort and Mean i-Ready District Scale Scores by Assessment Period</td>
<td>58</td>
</tr>
<tr>
<td>Table 7: Results of One-Sample t-Test Comparisons of Cohort Mean Scale Scores to i-Ready National Mean Scale Score Norms by Assessment Period</td>
<td>59</td>
</tr>
<tr>
<td>Table 8: i-Ready Scale Score Differences Between Means of Research Cohort and 2018-2019 National Norms</td>
<td>72</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1: Estimated Percentages of Students Participating in Three Delivery Modes During School Closures Due to COVID-19 in Spring 2020</td>
<td>25</td>
</tr>
<tr>
<td>Figure 2: Growth Curve in Oral Reading Fluency for First-Grade Students in Fall 2018 Through Fall 2021</td>
<td>29</td>
</tr>
<tr>
<td>Figure 3: Comparisons of Pre-COVID Achievement to Post-COVID Achievement in English Language Arts (ELA) and Mathematics by Grade Level on Three Instruments</td>
<td>35</td>
</tr>
<tr>
<td>Figure 4: Measures of Academic Progress Assessment Mean Percentile Rank Changes of Students in Spring 2019 and Students in Spring 2021</td>
<td>42</td>
</tr>
<tr>
<td>Figure 5: i-Ready Reading Scale Score Growth of 2nd–4th Grade Students by Pre- or Postpandemic Time Period and Instructional Location</td>
<td>46</td>
</tr>
<tr>
<td>Figure 6: Mean Scale Score Comparison by Assessment Period: Study Sample vs. i-Ready 2018-2019 Academic Year National Norms</td>
<td>60</td>
</tr>
<tr>
<td>Figure 7: i-Ready Diagnostic Reading Assessment Results From Fall 2020 and Fall 2021</td>
<td>66</td>
</tr>
<tr>
<td>Figure 8: i-Ready Diagnostic Reading Assessment Results from Spring 2021 and Spring 2022</td>
<td>67</td>
</tr>
<tr>
<td>Figure 9: Percentage of Students Below Grade Level Phonics Skills on i-Ready Diagnostic Assessments</td>
<td>70</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

The onset of the COVID-19 pandemic in the spring of 2020 disrupted educational institutions as they abruptly transitioned from traditional face-to-face (FTF) instruction to online delivery of instruction. Although online instruction existed as an option for students at all levels of education before the global emergency, virtual learning became compulsory in many school districts across the United States during school closures in March 2020 (Sato, 2020). According to recent research from The Department of Education’s Office of Civil Rights (2021), most educators felt unprepared to provide high-quality instruction in online formats, especially when the transition demanded rapid deployment. Additionally, many families lacked the technical resources to support their children in online environments.

Problem Statement

The COVID-19 pandemic highlighted the collective lack of experience and preparation of most educators to plan, implement, and evaluate the transition from F2F instruction to online delivery of instruction (Domingue et al., 2021). However, the influence of the pandemic on students’ academic performance is still under review and may have a far-reaching influence on crucial foundational skills in reading, especially for primary-grade students (Engzell et al., 2021).

Definitive evidence related to student academic performance during the pandemic is variable and still under investigation (Domingue et al., 2021). Recent research conducted since the onset of the pandemic generally falls into one of four common thematic categories: the
expectation of learning loss, the emotional influences on educators and families, the shortfalls of online education for emergency purposes, and necessary actions to mitigate learning loss (Karakose & Demirkol, 2021). An exploratory study of reading performance during the pandemic may help researchers and educators understand elementary students’ reading growth over time after the transition from F2F to virtual instruction (Kuhfeld et al., 2020).

**Background of the Study**

Many researchers have investigated the responses of government and school districts to school closures (Carrigan et al., 2021; Sato, 2020; Soni, 2020; Storey & Slavin, 2020). Most published studies referred to the psychological strain endured by educators and students’ families (Hasan & Bao, 2020). However, few studies have provided consistent, concrete evidence of verifiable learning loss among students, especially with regard to reading achievement.

Analysts at most state and national departments of education in the United States have collected and compiled available data related to student achievement since school closures; however, the results of the data analyses are not easily interpreted. For example, in a press release by the Florida Department of Education (2021a), officials reported that the percentage of third-grade students who demonstrated acceptable levels of reading proficiency dropped from 58% in 2019 to 54% in 2021 on the state’s annual standardized reading assessment, which is administered each spring semester. However, this percentage is likely mediated by the fact that most students in Florida were not required to take the state assessment in spring 2020 after the school closures. The next required administration of the assessment was in spring 2021, after most students had returned to F2F classrooms for the 2020-2021 school year. Due to the variable delivery systems used during the pandemic and its aftermath, the mean reading scores of third-grade students in 2019 may not necessarily be comparable to the mean reading scores of third-
grade students in 2021. More explicit interpretation of the data and additional data analyses are needed to guide policymakers as school districts and states return to F2F instruction while implementing the lessons learned from virtual instruction.

Most COVID-19 research studies to date have focused on (a) the transition from FTF instruction to online instruction, (b) the need to improve online instructional methods, (c) the well-being of all stakeholders in education, and (d) the need to develop remedies or professional development to address insufficient or ineffective models of online instruction (Karakose & Demirkol, 2021). Because the pandemic is an ongoing international crisis, more research is needed to understand the relationships between extended school closures and student achievement in all subject areas, especially reading (Domingue et al., 2021; Gore et al., 2021). Many researchers have also suggested that the pandemic may influence low-income students and students from minority groups differentially when compared to non-minority peers (e.g., Kaffenberger, 2021; Mutch, 2021). Storey and Slavin (2020) suggested that solid, verifiable research evidence on the influence of pandemic-influenced long-term school closures and student performance is highly variable and ambiguous.

**Theoretical Framework**

Jeane Sternlicht Chall was a psychologist and literacy scholar at the Harvard Graduate School of Education from 1965 until her retirement in 1991 (Harvard Graduate School of Education, 2020). On the website created to celebrate 100 years of the Harvard Graduate School of Education, Chall was described as a literacy pioneer without whom it “would be difficult to discuss what we know about early literacy” (Harvard Graduate School of Education, 2020). Chall began her seminal research in the 1950s as an advocate for early childhood literacy and the use of phonics instruction (Semingson & Kerns, 2021). Beginning in the early 1960s, Chall
served on numerous federal and state advisory boards, such as the Project Literacy steering committee for the U.S. Office of Education, the Commission on Reading for the National Academy of Education, and the Elementary Education Study Group for the U.S. Department of Education.

Chall’s (1995) stages of reading development were a significant impetus to promote intensive study on the science of reading during her lifetime as a researcher and reading advocate. Her research on reading as a developmental process continues to influence scholars, educators, policymakers, and publishers (Semingson & Kerns, 2021). Chall’s research provided evidence of reading development that progresses through six basic stages over time (Chall & Jacobs, 2003). These stages are depicted in Table 1.

**Table 1**

*Chall’s Stages of Reading Development*

<table>
<thead>
<tr>
<th>Stage</th>
<th>Approximate Age/Grade</th>
<th>Characteristics and Masteries by End of Stage</th>
<th>How Acquired</th>
<th>Relationship of Reading to Listening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 0: Pre-reading “pseudo reading”</td>
<td>6 months–6 years preschool</td>
<td>Child “pretends” to read, retells story when looking at pages of a book previously read to him/her, names letters of the alphabet, recognizes some signs; prints own name; plays with books, pencils, and paper</td>
<td>Being read to by an adult (or older child) who responds to and warmly appreciates the child’s interest in books and reading, being provided with books, paper, pencils, blocks, and letters. Dialogic reading.</td>
<td>Most can understand the children’s picture books and stories read to them. They understand thousands of words they hear by age 6 but can read few, if any, of them.</td>
</tr>
<tr>
<td>Stage 1: Initial reading and decoding</td>
<td>6–7 years old 1st grade and beginning of 2nd</td>
<td>The child learns the relation between letters and sounds between printed and spoken words; the child is able to read simple text</td>
<td>Direct instruction in letter-sound relations (phonics) and practice in their use. Reading simple stories using words with phonic</td>
<td>The level of difficulty of language read by the child is much below the language understood when</td>
</tr>
<tr>
<td>Stage</td>
<td>Approximate Age/Grade</td>
<td>Characteristics and Masteries by End of Stage</td>
<td>How Acquired</td>
<td>Relationship of Reading to Listening</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------</td>
<td>-----------------------------------------------</td>
<td>--------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Stage 1:</td>
<td></td>
<td>containing high-frequency words and phonically regular words; uses skill and insight to “sound out” new one-syllable words.</td>
<td>elements taught and words of high frequency. Being read to on a level above what a child can read independently to develop more advanced language patterns, vocabulary, and concepts.</td>
<td>heard. At the end of Stage 1, most children can understand up to 4000 or more words when heard but can read about 600.</td>
</tr>
<tr>
<td>Stage 2: Confirmation and fluency</td>
<td>7–8 years old 2nd and 3rd grade</td>
<td>The child reads simple, familiar stories and selections with increasing fluency. This is done by consolidating the basic decoding elements, sight vocabulary, and meaning context in the reading of familiar stories and selections.</td>
<td>Direct instruction in advanced decoding skills; wide reading (instruction and independent levels) of familiar, interesting materials that help promote fluent reading. Being read to at levels above their own independent reading level to develop language, vocabulary, and concepts.</td>
<td>At the end of Stage 2, about 3000 words can be read and understood, and about 9000 are known when heard. Listening is still more effective than reading.</td>
</tr>
<tr>
<td>Stage 3: Reading for learning new ideas</td>
<td>9–13 years old 4th–8th grade intermediate 4th–6th junior high school 7th–9th</td>
<td>Reading is used to learn new ideas, gain new knowledge, experience new feelings; to learn new attitudes, generally from one viewpoint.</td>
<td>Reading and study of textbooks, reference works, trade books, newspapers, and magazines that contain new ideas and values, unfamiliar vocabulary, and syntax; systematic study of words and reacting to the text through discussion, answering questions, writing, etc. Reading of increasingly more complex text.</td>
<td>At the beginning of Stage 3, listening comprehension of the same material is still more effective than reading comprehension. By the end of Stage 3, reading and listening are about equal for those who read very well. Reading may be more efficient.</td>
</tr>
<tr>
<td>Stage</td>
<td>Approximate Age/Grade</td>
<td>Characteristics and Masteries by End of Stage</td>
<td>How Acquired</td>
<td>Relationship of Reading to Listening</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------</td>
<td>---------------------------------------------</td>
<td>--------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Stage 4: Multiple viewpoints</td>
<td>15–17 years old 10th– 12th grade</td>
<td>Reading widely from a broad range of complex materials, both expository and narrative, with a variety of viewpoints.</td>
<td>Wide reading and study of the physical, biological, social sciences, and the humanities, high quality and popular literature, newspapers, and magazines, systematic study of words and word parts.</td>
<td>Reading comprehension is better than listening comprehension of materials of difficult content and readability. For poor readers, listening comprehension may be equal to reading comprehension.</td>
</tr>
<tr>
<td>Stage 5: Construction and reconstruction</td>
<td>18+ years old college and beyond</td>
<td>Reading is used for one’s own needs and purposes (professional and personal); reading serves to integrate one’s knowledge with that of others, to synthesize it, and to create new knowledge. It is rapid and efficient.</td>
<td>Wide reading of ever more difficult materials, reading beyond one’s immediate needs, writing of papers, texts, essays, and other forms that call for integration of varied knowledge and points of view.</td>
<td>Reading is more efficient than listening.</td>
</tr>
</tbody>
</table>

**Note.** Adapted from *Stages of Reading Development* (2nd ed.), by J. S. Chall, 1995, Wadsworth Pub Co.

Although individual students progress through the stages of reading at different rates, Chall (1995) reported that children’s development of reading was invariant. In other words, each stage builds on the previous stage as children mature and develop discrete reading skills that continue to grow throughout the readers’ lifetimes. To become fluent readers, students must simultaneously master the discrete skills of phonemic awareness, phonics, decoding,
automaticity, fluency, vocabulary development, and comprehension. Each developmental stage provides teachers with clear learning targets for all students as they move through the stages.

Chall’s (1995) theories on the stages of reading provide curriculum designers and educators with an evidence-based framework to guide and evaluate students’ reading progress. Her research and theories continue to influence reading curricula, reading instruction, and assessments designed to measure reading development (Semingson & Kerns, 2021). According to Semingson and Kerns (2021), today’s curriculum designers often develop lessons and assessments focused on student progression through each stage and are better able to make appropriate evidence-based instructional decisions for each student. Although Chall’s (1995) stage theory of reading development is not without opposition, especially among supporters of whole-language approaches to reading (Semingson & Kerns, 2021), many publishers continue to use Chall’s (1995) reading development paradigm to develop and market diagnostic reading assessments and personalized instruction for learners. For example, i-Ready is a widely used reading program designed to diagnose student performance on the discrete subtasks of reading and provide personalized instruction for individual students via computer-based tutorials, practice, and assignments (Curriculum Associates, 2019).

Each growth phase in i-Ready aligns with Chall’s (1995) stages of reading development and Scarborough’s (2001) description of the discrete reading skills in her reading rope. The reading rope consists of lower and upper strands of reading skills; when all the strands are holistically combined, readers can read accurately, fluently, and with high levels of comprehension. The lower strands include phonological awareness, decoding skills, and sight word recognition of common words. The upper strands include background knowledge, vocabulary acquisition and comprehension, language structures, verbal reasoning, and literacy
knowledge. Chall’s stages of overall reading development and Scarborough’s indicators of skills within the stages help to clarify the overall complexity of learning to read by breaking the process down into developmental steps. The publishers of i-Ready designed its diagnostic tests to measure individual student’s mastery of developmental skills and stages and to provide instructors with specific data to help each reader succeed.

**Purpose Statement**

The purpose of this study was to explore the reading performance scores of elementary school students before and after school closures due to COVID-19 in one school district.

**Research Question**

What were the reading performance scores of third-grade students before and after the COVID-19 pandemic?

**Research Design**

This study was a nonexperimental, exploratory study of archival quantitative data. The research population included all students in third grade from one school district in the Southeast region of the United States.

**Overview of Methodology**

**Data Collection**

After approval by the Institutional Review Board at Southeastern University, the researcher submitted an official request to the target school district to obtain student-level and district-level reading performance data from the district’s Performance Matters (Unify, 2020) database, which includes i-Ready composite scale scores for each student on each quarterly administration of i-Ready’s diagnostic assessment under study. In addition, the researcher requested students’ demographic data to include exceptional student education (ESE)
designation, English language learner (ELL) designation, and students who had been previously retained.

**Population and Sample**

Approximately 41,000 students in the target school district were enrolled in kindergarten through Grade 12 during the 2020-2021 school year (Florida Department of Education, 2021b). The race and ethnic origins of the target district were identified as 48% White, 20% Black or African American, 25% Hispanic, 3% Asian or Pacific Islander, and 5% other. The percentage of households with a computer was 88%, and 81% of the households had broadband internet access. The target population in this study consisted of all third-grade students in the district. The sample was composed of 2,006 third-grade students.

**Instrumentation**

This study measured reading performance over time using individual students’ mean composite reading scale scores from computer-based i-Ready diagnostic tests of reading from five designated assessment periods. The diagnostic reading assessments are adaptive tests that measure both overall reading comprehension and student skill levels for each stage of reading development. The publisher of the i-Ready program, Curriculum Associates, created the diagnostic assessments to measure elementary students’ discrete skill development in phonological awareness, phonics, high-frequency words, vocabulary, reading comprehension for literature, and reading comprehension for informational text.

Each i-Ready diagnostic assessment provides an overall scale score that represents a comparison of student performance to grade-level norms and grade-level equivalents for each of the measured stages of reading (Curriculum Associates, 2019). For example, a second-grade student might earn a scale score of 432 in the first assessment period, which represents reading at
a first-grade level. Along with the scale score and accompanying grade-level equivalent, the same student might test out of phonological awareness and perform at first-grade level for phonics, second-grade level for high-frequency words, and first-grade level for vocabulary, literature comprehension, and comprehension of informational text.

Each student who participates in the i-Ready program in the target district is administered the diagnostic tests at the beginning, middle, and end of each academic year. After completing the first administration, the student receives a scale score that represents a goal of approximately one year’s growth as a summative measure (Curriculum Associates, 2019). Students also receive a stretch goal, which is designed to quantify larger steps for below-level readers to assist them in reaching grade level by the end of the school year. These target scores can assist teachers as they form guided reading groups and individualized reading instruction. The diagnostic tests are computer-based and adaptable to individual student progress. For example, as students progress through instruction on phonics, the online assessment of that skill decreases in frequency, and comprehension-based skill assessments increase in frequency. As the student’s reading ability improves throughout the school year, growth is reflected in subsequent assessment results.

**Data Analyses**

After receiving the archived datasets from the district, the researcher cleaned the data in an Excel document to identify missing data and any other problems that compromised the integrity of the datasets. In addition, the data were disaggregated by ESE and ELL designation; these datasets will be examined in future studies. The researcher used 5 one-sample $t$ tests to compare the third-grade cohort’s mean composite scale scores for five sets of i-Ready reading performance: January 2020 (prior to school lockdowns), August 2020 (after schools reopened),
January 2021 (winter/midyear assessment period), May 2021 (spring assessment period), and August 2021 (beginning of year assessment).

After disaggregating and cleaning the data in Excel, the researcher conducted the following analyses in conjunction with the dissertation chair and methodologist:

- tests for the normalcy of data;
- Cronbach alpha tests to determine the reliability of scores;
- descriptive statistics, including frequencies, means, and standard deviations;
- one sample $t$-test comparisons of the research cohort’s mean composite scale scores to district and national i-Ready norms.

**Ethical Considerations**

The researcher requested approval from Southeastern University’s Institutional Review Board and the target school district. Because the researcher obtained archival data from the district, no student or parent permissions were required to conduct the study. All data from the target school district were anonymized by district personnel to protect the students’ privacy. The only researchers on the committee who had access to the anonymized data were the committee chair, the methodologist, and the student researcher. The student researcher maintained the raw data on a password-protected laptop in a locked office with the required two-step authentication to access the dataset. The researcher employed appropriate ethical standards for the treatment of data from human subjects when reviewing, analyzing, and reporting the study results.

**Limitations**

All students from kindergarten through Grade 8 in the target school district took the i-Ready Diagnostic Reading Assessment. This research study focused only on diagnostic reading assessments for third- through fifth-grade students. This study addressed five specific i-Ready
assessment periods: January 2020, August 2020, January 2021, May 2021, and August 2021. Although many school districts use i-Ready as an assessment tool, this study focused on one school district. The i-Ready program includes math diagnostic assessments; however, this study examined only reading assessment data. The i-Ready diagnostic assessment measures were designed to measure reading performance at different periods during the school year; as such, the instruments are considered performance-based, formative assessments and should not be considered summative achievement assessments (Curriculum Associates, 2019).

**Definition of Key Terms**

*Reading performance* is the measurement of reading skill acquisition designed to provide ongoing feedback to guide instruction. The purpose of a formative reading assessment is to provide students and instructors with the data necessary to improve performance in particular stages or skills of overall reading development (Curriculum Associates, 2019).

*Reading achievement* is designed to measure the totality of acquired reading skills at the end of a specified period, such as end-of-year exams (Barkley & Major, 2016).

**Significance of the Study**

This research study adds to the body of research regarding school closures and virtual instruction during the COVID-19 pandemic and students’ reading performance. Although a growing body of research related to the pandemic and its relationship to learning has been published, minimal research evidence exists to address specific student outcomes in reading performance during the school closures due to the pandemic. According to Karakose and Demirkol (2021), most of the scientific studies on the pandemic and education focused on governments’ initial responses to the crisis, the transition from FTF instruction to online delivery of instruction, unequal access to required resources for online instruction, and the mental health
of students, families, and teachers. Further exploration of the pandemic’s influence on elementary students’ reading performance can assist educators and educational institutions in responding to students’ academic needs for reading instruction in possible disruptions in the future. Another common theme in the existing literature related to the COVID-19 pandemic was the assumption that the disruptions to the educational system would result in learning loss (Cummiskey & Stern, 2020; Dorn et al., 2020; Engzell et al., 2021; Godsey, 2020). This research study adds to the body of knowledge related to school interruptions, remote instruction, and assumptions of learning loss.
II. REVIEW OF LITERATURE

This study was designed to explore the reading performance scores of elementary school students before and after school closures due to COVID-19 in one school district. The first section of this literature review focuses on research on scheduled and unscheduled interruptions of instruction and their influence on student academic growth and progress before the COVID-19 pandemic; these studies examined the influence of summer vacations and weather events on academic achievement or progress. The second section of this literature review describes recent math and reading studies conducted since school closures and the transition to remote instruction in the spring of 2020.

Pre-COVID Learning Loss

Cooper et al. (1996) conducted an early study on summer learning loss. The authors conducted a meta-analytic review of 39 studies on the relationships between summer vacation and math and reading achievement scores from 1906 to 1995. The sample sizes, student grade levels, and target populations of the 39 studies varied widely. The researchers separated the 39 studies into two groups: pre-1975 and post-1975. Sample sizes for the pre-1975 studies ranged from seven to 370,000 students. Sample sizes for the post-1975 studies ranged from 59 to 39,000 students from the United States and Canada.

Cooper et al. (1996) included background information to contextualize the differences between school years, academic expectations, and assessments over time, as well as the
challenges of generalizing the results of such diverse studies and populations. In fact, the authors concluded that the differences between older and more modern school years and grade-level expectations introduced spurious results. For example, 20th-century students in agricultural regions of the United States often attended school for approximately 6 months, while students in urban areas of the nation attended school for 10 to 12 months of the same calendar year. According to the authors, the percentage of families that depended on agriculture for a living declined from 85% at the beginning of the century to approximately 3% in the final 2 decades; a 9-month school year eventually emerged as the norm for most regions in the United States.

Consequently, Cooper et al. (1996) divided the learning loss studies into two groups: pre-1975 publications and post-1975 studies. The researchers synthesized the results of 26 pre-1975 studies using a vote count or tally method to describe the findings. A vote count is a method of synthesizing research in which the researchers divide the studies into three categories: studies that showed significant positive results, studies that showed significant negative results, and studies with nonsignificant results.

Cooper et al. (1996) further categorized the results of the 26 pre-1975 studies into six academic subskills from four broad categories of math, language arts, history, and science for first- through eighth-grade students. Most of the studies included data for more than one subject area. Researchers compared 80 assessments of different subject areas to examine the test scores before and after the summer vacation. Out of 80 comparisons in which students demonstrated either a loss or gain, 48 (60%) of the comparisons revealed a loss in skills over the summer. However, the results differed depending on the subject area, and the differences resulted in a statistical trend \((p < .08)\). All study results from math computation assessments indicated a loss in skills for students in the first through ninth grades from 1906 to 1968. The comparisons of
reading assessments showed a gain in 10 of 17 (58%) studies. Out of the six academic subskills under study, only math computation and spelling assessments revealed a loss of skills over the summer. Little to no summer learning loss was observed in math concepts, reading comprehension, or content-area knowledge.

Cooper et al. (1996) subsequently conducted a separate meta-analysis of the 13 post-1975 (i.e., 1976-1995) studies to further examine the influence of summer break on students’ academic achievement in the first through eighth grades. The researchers reported that the average length of time between spring and fall assessments was 131 days, and the average length of the summer break was 101 days, making comparisons difficult. The authors found that math and reading skill losses over the summer approximated 1 month of learning or one-tenth of a standard deviation relative to the spring assessments for reading and math.

Cooper et al. (1996) used two different metrics to calculate and interpret the results of the influence of summer break on reading and math achievement. The first metric, the $d$-index, was the mean difference between the spring and fall achievement scores. The second metric, the difference in grade-level equivalence (DGLE), was used to express the difference between each grade level’s spring and fall grade-level equivalent scores to compare the change in scores to national norms. Similar to the pre-1975 results, students in the post-1975 group demonstrated skill loss in math but not in reading. In fact, most students made reading gains in the summer. One possible explanation for the differential result is that math skills are more likely to be practiced in a school setting, while reading and language skills are likely to be practiced throughout the summer. The mean differences between spring and fall reading achievement scores are presented in Table 2.
Table 2

Mean Difference Scores in Reading Between Spring and Fall Terms by Grade Level Using Two Metrics

<table>
<thead>
<tr>
<th>Grade</th>
<th>d-index</th>
<th>DGLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of samples</td>
<td>Sample size</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>1,967</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>2,189</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>2,169</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>2,778</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>4,056</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>1,212</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>642</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>235</td>
</tr>
</tbody>
</table>

Note. $d$-index = measure of effect size; DGLE = the difference in grade level equivalents.


* $p < .05$.

Cooper et al. (1996) reported that students from low-income homes demonstrated lower reading skill scores after the summer break; in contrast, students from middle-income families experienced little or no loss in reading skills over the summer. The researchers suggested that this disparity might be attributed to the availability of opportunities and resources for students to practice reading skills during the summer. No gender or race differences were observed.
In a more recent study, Marcotte and Hemelt (2008) conducted a quantitative study of third-, fifth-, and eighth-grade students to determine whether unscheduled school closures influenced student performance on standardized reading assessments. Unscheduled school closures in the state of Maryland included weather events such as excessive snow, ice-covered roads, and other weather events that caused dangerous driving conditions. The Maryland State Department of Education provided standardized test data in reading for all third-, fifth-, and eighth-grade students from 1994 to 2005. The Maryland State Assessment (MSA) replaced the Maryland State Performance Assessment Program (MSPAP) in 2003. The MSPAP and the MSA were administered on the same day in all Maryland school districts each year in the spring. Marcotte and Hemelt established the criteria for reading mastery as the percentage of students who met the Maryland State Department of Education guidelines on the MSPAP and MSA math and reading assessments. The researchers also examined data on the number of annual unscheduled school closing days from 1994 to 2005 and snow accumulation from the National Oceanic and Atmospheric Administration’s National Climatic Data Center to determine the number of school days lost due to inclement weather. The authors found that Maryland school districts averaged five unscheduled school closures per year from 1994 to 2005. The researchers conducted estimated models of the relationships between unscheduled closures in a year and student performance on the state math and reading assessments for each subject and grade level under study. Each model was estimated using “weighted least squares to adjust for heteroskedasticity in the error term”; the “possibility of serial correlation in the error term” was accounted for using Huber-White robust standard errors (Marcotte & Hemelt, 2008, p. 322).

The results of Marcotte and Hemelt’s (2008) analyses revealed that in years with frequent, unscheduled school closures (more than 5 days), students’ standardized reading and
math test scores were lower than the mean performance levels on both the MSPAP and the MSA math and reading assessments in third, fifth, and eighth grades. Unscheduled closures before February revealed a significant negative effect for third-grade reading and math assessments only. Unscheduled school closures between February and April were significantly and negatively correlated to student performance in all three grade levels (Grades 3, 5, and 8). During years with few unscheduled school closures, students’ scores on math and reading standardized tests were significantly higher than the mean for each of the three grade levels on the MSPAP and the MSA assessments (Marcotte & Hemelt, 2008, p. 322). In addition, the percentage of students who made adequate yearly progress was lower for third-grade students in both reading and math than for fifth- and eighth-grade students during academic years with frequent school closures.

Significant results of the study are depicted in Table 3.

**Table 3**

*Mean MSA Scores During School Years with Frequent Unscheduled School Closures Compared to Mean MSA Scores and Adequate Yearly Progress*

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>N</th>
<th>Overall t</th>
<th>Reading t</th>
<th>Math t</th>
<th>Reading %</th>
<th>Reading SE</th>
<th>Math %</th>
<th>Math SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Third</td>
<td>6,178</td>
<td>4.54</td>
<td>1.61</td>
<td>1.72</td>
<td>−1.024</td>
<td>0.400*</td>
<td>0.963</td>
<td>0.331**</td>
</tr>
<tr>
<td>Fifth</td>
<td>6,148</td>
<td>4.40</td>
<td>2.56</td>
<td>0.94</td>
<td>−0.325</td>
<td>0.140*</td>
<td>−0.612</td>
<td>0.279*</td>
</tr>
<tr>
<td>Eighth</td>
<td>1,838</td>
<td>3.35</td>
<td>−1.27</td>
<td>2.36</td>
<td>−0.028</td>
<td>0.131</td>
<td>−0.645</td>
<td>0.226**</td>
</tr>
</tbody>
</table>


*p < .05. **p < .01.*
The mean percentage of mastery by third-grade students on the reading assessment was 43.5%, 46.2% for fifth-grade students, and 37.6% for eighth-grade students. Between 2003 and 2004, during the highest recorded period of unscheduled school closings (between 10 and 12 closures), the third-grade mean percentage of reading mastery dropped by more than 6.36%. Conversely, during a period of fewer unscheduled school closings in 2005 and another period between 1997 and 2000, third-grade students performed between 2% and 4% above the mean of 43.5% on the reading assessment.

Marcotte and Hemelt (2008) concluded that frequent unscheduled school closings were significantly correlated to lower standardized test scores in reading for third-grade students only. In light of reading development theory, the influence of above-average school closure days on third-grade students is an important finding; the loss of more than 5 days of school in an academic year demands further investigation. The percentage of fifth- and eighth-grade students’ reading mastery scores were not significantly related, possibly because the reading fluency of older students was more developed and less influenced by school closures.

Sandberg Patton and Reschly (2013) conducted a quantitative study of students’ oral reading fluency (ORF) in grades two through five to determine the influence of summer break on reading growth. The study included a curriculum-based measurement of ORF using the Dynamic Indicators of Basic Early Literacy Skills Oral Reading Fluency (DIBELS) measures. The purpose of the study was to investigate any differences in ORF scores from spring to the following fall using the DIBELS and to augment the body of literature on changes in reading progress over the summer.

The initial sample size for Sandberg Patton and Reschly’s (2013) study consisted of 404 students from second through fifth grades in the Southeast region of the United States. After
excluding students who had been retained and students without either spring or fall ORF measurements, the final sample size was 317 students in the following grade levels: first grade ($n = 81$), second grade ($n = 86$), third grade ($n = 61$), and fourth grade ($n = 89$). All the students attended a Title 1 elementary school in a rural region. The researchers followed four cohorts from the end of the 2010-2011 school year through the fall of the 2011-2012 school year. The authors compared the mean pre- and posttest data for each of the four cohorts.

Sandberg Patton and Reschly (2013) examined changes in oral reading progress over summer break using the DIBELS (University of Oregon, 2022). The DIBELS was administered individually at least two times per year in most school districts to measure students’ oral reading skills, fluency, and reading outcomes in Grades 1 through 4. The teacher or assessor asked students to read a grade-level passage aloud for 1 minute while the teacher marked incorrect word recognition. The outcome of each assessment was the number of words read correctly by a student in 1 minute (WCPM). In Sandberg Patton and Reschly’s (2013) study, the spring and fall assessments were benchmark assessment periods in which the student read three different grade-level passages aloud; the median score of the three passages was recorded. In spring 2011, all first- through fifth-grade students took the DIBELS. During the fall of 2011, the same students took the assessment; however, the researcher only used data from students in first through fourth grade in 2010 for comparison purposes.

Sandberg Patton and Reschly (2013) compiled the DIBELS test results for Grades 1 through 4 for analysis. They first disaggregated the ORF data by gender, ethnicity, free and reduced lunch eligibility (FRL), ELL, and ESE status. The total sample ($n = 317$) was approximately 54% female, 61% White, 12% Black, 21% Hispanic, and 6% other. Other subgroups of the overall sample consisted of 71% FRL, 14% ELL, and 13% in an ESE program.
Sandberg Patton and Reschly (2013) screened the resulting data for adherence to normality, parametric assumptions, and outliers; 14 cases were eliminated as outliers after examining histograms and boxplots. The researchers then computed descriptive statistics and checked for correlations between variables. The researchers subsequently used repeated-measures analyses of variance to compare the spring 2011 ORF scores to the fall 2011 ORF scores for each student in the sample. However, since DIBELS ORF scores are not comparable across grade levels, each grade level was analyzed separately. The researchers used the demographic variables as between-subjects factors; the within-subjects factor was the period in which the assessments were administered. The results of Pearson r correlation analyses revealed that spring and fall ORF scores were significantly different for Grades 1 through 4 ($p < .01$). The overall correlation for students in second grade in spring 2010 was .930. The correlation for second-grade students was .901 among third-grade students, .916 for fourth-grade students, and .931 for fifth-grade students. These results indicated that the spring and fall ORF scores were highly correlated.

Sandberg Patton and Reschly (2013) also shared the results of the repeated measures analyses of variance. In third grade, the results were significant for the effect of time of test administration ($p < .001$) and the main effect of ESE status ($p = .013$). The researchers reported that third-grade students lost an average of nine WCPM over the summer months. Second-grade students lost approximately five WCPM. In the fourth and fifth grades, the researchers observed no learning loss. However, the main effect for ESE status was statistically significant among fourth-grade students, $F(1, 56) = 8.10, p = .006$, and fifth-grade students, $F(1, 56) = 9.37, p = .003$. The loss in WCPM in second and third grades, but not in fourth and fifth grades, may relate to the developmental stages of reading (Chall, 1995). Sandberg Patton and Reschly (2013)
suggested that second- and third-grade students require more help with reading practice when learning to read, while students in later stages of reading development can read independently. Students who can read independently during the summer may have more opportunities to maintain levels of reading skills over the summer break. However, younger students possibly experience greater assistance from caregivers during the summer, while independent readers may experience less assistance and accountability. The long-term nature of the COVID-19 pandemic, school closures, and rapidly deployed remote instruction differs considerably compared to a typical summer break. A study of the influence of a long-term break from F2F instruction on the development of reading skills may improve the ability of educators to address future disruptions in educational programming.

Based on the results of their study, Sandberg Patton and Reschly (2013) stated that the reading skills of certain student subgroups were disproportionately influenced by the summer break. For example, differential losses were observed in second grade among FRL and ESE students. Second-grade students with FRL status lost seven WCPM, while students without the FRL designation did not display a significant change in WCPM. Similar to students with the FRL designation, second-grade ESE students also showed a decline in WCPM, although the loss was not statistically significant compared to non-ESE students. These results indicated that differential effects were mediated by students’ ages, development, and academic status.

**COVID-19 Studies**

In the fall of 2020, Dorn et al. (2020) conducted a nonexperimental, archival study of reading and math assessment data from Curriculum Associates’ i-Ready diagnostic assessments (Curriculum Associates, 2019) among students in kindergarten through fifth grade. The purpose of the study was to examine the influence of pandemic-related school closures and remote
instruction on students’ academic performance in reading and math. The researchers compiled data from i-Ready reading and math performance assessments during the 2020-2021 pandemic-influenced school year to determine the differences when compared with prepandemic, historical norms. The sample was comprised of math assessment scores from 357,731 students and reading scores from 255,018 students from 25 states in kindergarten through grade five.

Dorn et al. (2020) examined student growth in math and reading on i-Ready assessments from fall 2019 to fall 2020, then compared the results to historical averages from fall to fall of the previous three prepandemic academic years: 2016-2017, 2017-2018, and 2018-2019. The average of all scores in the pandemic sample revealed that students from kindergarten through fifth grade demonstrated mastery of 67% in math skills and 87% in reading skills when compared to the average growth from prepandemic years (with 100% as equivalent to historical matched averages). In addition, the researchers found differential patterns of reading mastery and growth based on schools with different demographics.

Dorn et al. (2020) concluded that there were disturbing differences between reading mastery scores of students in schools having more than 50% students of color compared to schools’ having 50% or more White students. The authors postulated that the differences may have reflected the differential influence of remote instruction in spring 2020. Students of color may have had less access to appropriate technology-based tools and reliable internet access than White students during remote instruction.

Sawchuck and Sparks (2021) reported that an average of 60% of students in America participated in remote instruction in the spring of 2020. However, the percentages of students in different demographic groups varied, as seen in Figure 1.
Further research into the effectiveness of remote learning and optimal delivery methods is needed to assist educators as they strive to meet the needs of all students. In addition, the relationships between remote instruction and academic achievement are critical areas of research to inform teachers, administrators, and policymakers.
Kuhfeld et al. (2020) conducted a large-scale study of students in Grades 3 through 7 who took the Measures of Academic Progress Assessment (MAP; Northwest Evaluation Association [NWEA], 2021). The purpose of the study was to examine the relationships between pandemic-related school closures, remote instruction, and overall student math and reading achievement. The researchers first investigated existing research on the influence of summer vacations, weather-related school closures, and absenteeism on student achievement, as well as research on absenteeism and its relationships to student achievement. The national sample consisted of five million students in third through seventh grades who took the MAP growth assessments over the course of 2 prepandemic academic years in 2017-2018 and 2018-2019. To investigate three research questions, the researchers wanted to compare actual growth trajectories during a standard school year to expected growth models. The first research question asked, what were students’ learning patterns during the onset of the pandemic in the 2019-2020 school year? Second, how did learning rates vary during the pandemic-related school closures in the spring of 2020? Finally, what was the relationship between the extended school closures and projected learning rates in math and reading in the 2020-2021 school year? Data collection included anonymized longitudinal student achievement data in math and reading from NWEA. The results of the researchers’ analyses revealed that students achieved an overall average of 63% to 68% of expected learning gains in reading in the 2019-2020 academic year compared to the previous two school years. According to Kuhfeld et al., these results supported the existence of learning loss during virtual instruction.

Many teachers and administrators predicted that students would exhibit greater variability in their reading skills when they returned to F2F classes in the fall of 2020. To address this question, Kuhfeld et al. (2020) compared the results of their study to previous studies of learning
loss in reading during summer breaks. The authors conducted comprehensive reviews of the literature and found that students who typically lost reading skills during the summer regained skills during the subsequent fall. In the same manner, the researchers expected many students to experience some level of learning loss after the spring 2020 school closures and the summer 2020 break. However, the authors conceded the difficulties of accurately weighing the influence of the vast range of challenges during the COVID-19 pandemic.

Kuhfeld et al. (2020) came to three main conclusions related to their study’s research questions. First, students might begin the fall of 2020 less academically proficient than in a typical year and with greater variability in reading skills. Therefore, adequate diagnosis and planning would help educators know where to begin review and instruction. Second, students who fell behind during the summer in the past tended to gain when school commenced. Finally, the researchers reported a large degree of variability in learning gains during the summer of 2020. Specifically, the upper third of the distribution of all grade levels under study showed reading gains when tested immediately after the summer 2020 break, with sixth- and seventh-grade students earning the highest gains among all grade levels. Third- through seventh-grade students in the lowest quartile demonstrated the highest levels of learning loss. The researchers also acknowledged that continued research throughout the pandemic was vital to the development of long-term educational goals and planning.

Analysts at most state and national departments of education in the United States collected and compiled available data related to student achievement since school closures due to the pandemic; however, the results of the data analyses are not easily interpreted. For example, in a press release by the Florida Department of Education (2021a), officials reported that third-grade students who demonstrated acceptable levels of reading proficiency dropped from 58% in
2019 to 54% in 2021 on the state’s annual standardized reading assessment, which is administered each spring semester. However, this percentage is likely mediated by the fact that most students in Florida were not required to take the state assessment in spring 2020 after the school closures; the next required administration of the assessment was in spring 2021 after most students had returned to F2F classrooms for the 2020-2021 school year. Finally, the mean reading scores of third-grade students in 2019 may not necessarily be comparable to the mean reading scores of third-grade students in 2021. More explicit interpretation of the data and additional data analyses are needed to guide policymakers as school districts and states return to F2F instruction while implementing the lessons learned from virtual instruction.

In a recent study, Domingue et al. (2021) conducted a large-scale, quantitative study of 111 school districts to determine ORF rates of approximately 100,000 elementary school students from first through fourth grades from September 1, 2018, through September of the 2020-2021 school year. After obtaining permission from more than 100 school districts nationwide that used the Literably reading program (Literably, Inc., 2018) to track students’ reading progress, the researchers collected ORF data from 22 states. Literably utilizes online, periodic assessments of ORF throughout the school year; according to Domingue et al. (2021), online delivery of the assessments allowed for easy transitions to online environments during the pandemic. First, the researchers analyzed first- through fourth-grade students’ longitudinal, within-person (individual) data as measured by Literably ORF assessments from September 2018 to September 2021. Domingue et al. averaged individual students’ scores by grade level and compared the mean growth scores to Literably’s norms for expected progress over time to analyze students’ growth in ORF. The researchers then conducted model estimates with fixed effects; the authors used B-splines to account for nonlinear trends over time. The analyses for
first-grade students are depicted visually in Figure 2.

**Figure 2**

*Growth Curve in Oral Reading Fluency for First-Grade Students in Fall 2018 Through Fall 2021*

![Growth Curve in Oral Reading Fluency for First-Grade Students in Fall 2018 Through Fall 2021](https://files.eric.ed.gov/fulltext/ED612594.pdf)


Based on the norms for acceptable yearly progress developed by Literably, Domingue et al. (2021) identified a decline in ORF scores for students in the first through fourth grades after the onset of the pandemic in the spring of 2020. However, after conducting a moderation analysis, the researchers found no significant differences between the school districts’ mean composite ORF scores from the 2020-2021 school year to ORF scores from the 2019-2020 school year. The authors concluded that ORF rates of students in first through fourth grades
largely returned to expected progress in the fall of 2020. The researchers suggested that increased familiarity with online learning platforms in the fall of 2020 may have influenced ORF scores. However, the researchers also concluded that gains made in the fall of 2020 did not make up for all the ORF losses recorded during the onset of the COVID-19 pandemic.

Kogan and Lavertu (2021) conducted a quantitative study of English language arts (ELA) achievement of all third-grade students in Ohio during the onset of COVID-19; the researchers examined student reading proficiency rates and student participation rates on the state’s standardized reading achievement tests from fall 2019 to fall 2020. Analyses revealed that student participation rates decreased in the fall of 2020 due to increased absenteeism. The researchers used regression models to “generate predicted values of student test scores for each year” (Kogan & Lavertu, 2021, p. 12). Participation rates on the state’s ELA assessment were generated for both years and compared to determine differences between 2019 and 2020. The analyses revealed that the number of third-grade students who demonstrated reading proficiency in Ohio declined from 45% in the fall of 2019 to 37% in the fall of 2020. In addition, participation in the state’s testing program dropped from 95% in the fall of 2019 to 81% in the fall of 2020. According to Kogan and Lavertu (2021), the decline of 0.23 standard deviations was slightly more than one-third of a year’s growth in reading skills. The authors concluded that an examination of Ohio’s ELA assessment results for the spring of 2021 would be necessary to provide a complete picture of the pandemic’s influence on reading achievement.

Gore et al. (2021) conducted an empirical study of student achievement in math and reading for second- and third-grade elementary school students in New South Wales, Australia, to determine the influence of the onset of the pandemic in the spring of 2020. The study began in the 2019 school year as a randomized controlled trial of professional development for teachers
called Quality Teaching Rounds (Teachers and Teaching, 2022). The professional development focused on improving quality teaching, school morale, and school culture. All teachers in all Australian schools in both 2019 (pre-COVID-19) and 2020 (during COVID-19) school years participated in the professional development. The school year in Australia begins in late January and ends in late December; therefore, the onset of the COVID-19 pandemic occurred in the middle of the Australian school year. Gore et al. also collected second- and third-grade students’ math and reading achievement data from the first cohort of 62 public schools in the 2019 school year prior to the pandemic and in the 2020 school year during the pandemic. The total sample of second- and third-grade students with standardized achievement test scores across both school years was 3,030 students. The instrument used to measure student achievement was the Progressive Achievement Tests in math and reading (Australian Council of Educational Research, 2011, as cited in Gore et al., 2021).

Gore et al. (2021) conducted preliminary analyses of the student achievement data. To guard against cohort effects, the researchers used a sample of matched classes to account for a within-school variance for both school years of the study in their analyses. The authors then used the procedure to match subsets of schools according to baseline achievement and demographic variables. Individual samples of matched classes from within schools for second- and third-grade students were developed from the datasets. The final sample size was 1,584 students in 2019 and 1,446 in 2020. Additionally, the combined math and reading achievement test score means were highly correlated in both 2019 and 2020.

Gore et al.’s (2021) data analyses were exploratory in nature. First, linear mixed models were used to compare reading and math achievement outcomes for each cohort for the 2019 and 2020 school years. The researchers then compared test scores for each school year, the time of
the assessment, and year-by-time interactions as categorical fixed effects within each model. In addition, the researchers used linear mixed models to determine the differences between means.

The 2019 school year cohort was used as the comparison group for group-by-time comparisons.

Gore et al. (2021) reported that no significant differences were observed between second- and third-grade math and reading achievement scores during the 2019 and 2020 school years. However, differences emerged when achievement test results of students with differing demographics were compared within each grade level. For example, the scores of second-grade students from socio-economically disadvantaged schools showed the loss of approximately 2 months of growth in math skills. Second-grade students from the mid-level socio-economic group designation gained the equivalent of 2 months of growth in math skills from 2019 to 2020.

The researchers observed no other significant differences in achievement in math or reading from 2019 to 2020.

During the onset of the pandemic in the spring of 2020, schools in New South Wales, Australia, shut down F2F learning and transitioned to online learning for 8 to 10 weeks (Gore et al., 2021). Gore et al. (2021) cautioned readers to carefully consider the contexts of the results of the study because of the varying factors of school districts during the pandemic. For example, in the Northern Territory of Australia, schools closed for only 4 days, while the state of Victoria, Australia, closed for almost 4 months.

Pier et al. (2021) conducted a quantitative study of data from interim assessments in math and reading of approximately 100,000 students in Grades 3 to 8 in California schools after pandemic-related school closures and subsequent transitions to online teaching and learning. The purpose of the study was to compare the differences between rates of student learning before the COVID-19 pandemic to a period during the pandemic. The period of remote instruction spanned
from the fall of 2019 through the winter of the 2021 school year in California. For comparison purposes, achievement data from two prepandemic periods in fall 2017 through winter 2019 were used in the analyses. The researchers compiled data from three different interim assessment instruments: Northwest Evaluation Association’s MAP assessment (Northwest Evaluation Association, 2021), Renaissance Learning’s STAR assessment (Renaissance Learning, 2022), and Curriculum Associate’s i-Ready diagnostic assessments (Curriculum Associates, 2019). The goal of the study was to address the question, how did students progress academically during the COVID-19 time periods compared to time periods before the pandemic?

According to Pier et al. (2021), the majority of local education agencies in California provided distance-only instruction from fall 2019 through winter 2021. The local education agencies subsequently provided interim math and reading assessment data from the three instruments during the pandemic-related shutdown periods and at least one of the two prepandemic periods under study.

Pier et al. (2021) compiled and analyzed interim (fall and winter) math and reading assessment data from 2017 through winter 2021. For the MAP and STAR assessments, the fall term was described as the period between August 1 and November 25; the winter term was defined as the period from November 26 through the end of February. The fall assessment periods for i-Ready differed; fall assessments took place between August first and the last day of October. The i-Ready winter assessment occurred at the end of February, which was identical to the MAP and STAR assessments.

Pier et al. (2021) described their research analysis model as similar to a quasievaluation model in which the prepandemic periods represented the control group and the pandemic-influenced period represented the treatment group. First, researchers computed growth models
for Grades 4 through 8. The reported grade levels corresponded to the grade levels at the time of the posttest, and the reported growth rates were from the previous fall to the following winter. Therefore, the time period was more than one school year. Using a fourth-grade student as an example, the measurement period during the pandemic-influenced period began when the student was in third grade in the fall of 2020 and ended in the winter of 2021 when the student was in fourth grade. The researchers used the growth models to calculate predictions based on students’ pretest scores and demographic factors of ELL, ESE, homeless status, and low socioeconomic status (SES). The researchers then calculated learning change based on the difference between growth in the pandemic-influenced period and the average growth of the prepandemic periods. The authors stated that “learning change is always calculated by comparing growth from year to year for students in the same grade, subject, assessment, and local education agencies; as a result, these measures compare different cohorts of students” (Pier et al., 2021, p.10).

Pier et al.’s (2021) analyses of the reading and math assessment data revealed significant reductions in academic progress during the pandemic-influenced period on all three reading and math assessment instruments in every grade level except the eighth grade. Using inverse-variance weighted means, the researchers averaged the results across grades and assessments. The researchers determined that during the pandemic-influenced period between fall 2019 and winter 2021, students demonstrated a loss of 2.6 months in reading compared to the prepandemic period. According to the authors, the only grade level that did not demonstrate a decline in reading and math academic progress during the pandemic-influenced period was the eighth grade; eighth-grade students actually demonstrated learning acceleration of approximately 5 months in both subjects. The results of the study are visually presented in Figure 3.
Figure 3

Comparisons of Pre-COVID Achievement to Post-COVID Achievement in English Language Arts (ELA) and Mathematics by Grade Level on Three Instruments


Pier et al. (2021) also reported reductions in academic progress during the pandemic for specific student demographic subgroups: low SES, ELL, ESE, and homeless students. Compared to average academic progress across all grades and assessments, low-SES students demonstrated a decline of 3.2 months in reading scores; students without low-SES designations showed a decline of 1.1 months. ELL students demonstrated a decline of 3.8 months in reading scores from the prepanademic years to the pandemic-influenced school years. The results for ESE students
were mixed. The only statistically significant difference between pre- and post-COVID-19 comparisons among ESE students was in eighth-grade math scores on the MAP assessment. The decline in ESE students’ MAP math scores was significantly lower when comparing pre- to posttest (Pier et al., 2021, p. 16). Unfortunately, no $p$-values were included in the study. Pier et al. concluded that third- through eighth-grade students’ academic skills in reading and math progressed more slowly during and after the onset of the COVID-19 pandemic when compared to other assessment periods before the pandemic.

Hammerstein et al. (2021) conducted a systematic review of evidence-based studies on the influence of COVID-19-related school closures and virtual instruction on achievement among primary and secondary students. The researchers concluded that sufficient time had passed to take an inventory of existing evidence of the pandemic’s influence on academic achievement after the onset of the pandemic. The authors asked two research questions. First, what was the overall influence of the COVID-19 pandemic on student achievement after the initial transition to virtual instruction in the spring of 2020? Second, what was the influence of the pandemic on varying student subgroups?

Hammerstein et al.’s (2021) systematic review of literature began with a search of the Web of Science database. The researchers searched for articles published between March 1, 2020, and April 30, 2021. The search began with keywords such as COVID, corona, SARS-CoV-2, learning, test scores, student performance, and achievement. The search was subsequently refined to search within key categories: education, educational research, economics, educational scientific disciplines, multidisciplinary studies, educational psychology, and social sciences interdisciplinary studies. After the initial search, the researchers selected 601 articles as potentially relevant studies. The researchers and two research assistants studied each publication
using three criteria for selection: a focus on pandemic-related school closures, a focus on primary and secondary education, and student achievement scores as the dependent variable in the studies.

As Hammerstein et al. (2021) continued studying and sorting the articles by relevance and quality, the sample narrowed to a total of 11 relevant studies. The researchers then conducted risk of bias assessments on the 11 studies. The highest risks of bias were studies with confounding variables, participant selection, and deviations from intended interventions. However, the researchers concluded that due to the inclusion of relevant controls, the natural formation of participant groups, and the fact that the pandemic-related school closures were not purposeful interventions, the risks of bias in those cases were deemed low. Hammerstein et al.'s synthesis of the 11 studies was guided by the two research questions to examine information related to seven factors: duration of the closure, school type, sample size, subjects, statistical method, general effect, and differential effects. The results of the literature analyses are presented in Table 4.
Table 4

Descriptive Analyses of Eleven Studies of Student Achievement Published Between March 1, 2020, and April 30, 2021

<table>
<thead>
<tr>
<th>Authors (Country)</th>
<th>Duration of closure in weeks</th>
<th>School type</th>
<th>Sample size</th>
<th>Subjects</th>
<th>Statistical method</th>
<th>General effect</th>
<th>Differential effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clark et al. (China)</td>
<td>7</td>
<td>Secondary</td>
<td>1,835</td>
<td>Reading, Mathematics, English, Politics, History</td>
<td>DiD regression</td>
<td>$-0.22\ SD$</td>
<td>Larger effect in girls, larger effect in low-achieving students</td>
</tr>
<tr>
<td>Depping et al. (Germany)</td>
<td>8</td>
<td>Elementary and Secondary</td>
<td>$\approx27,500$</td>
<td>Reading, Mathematics</td>
<td>Difference analyses</td>
<td>$0\ SD$ to $+0.05\ SD$ (reading) $-0.03\ SD$ to $-0.02\ SD$ (math)</td>
<td>$-0.03\ SD$ to $-0.02\ SD$ (math)</td>
</tr>
<tr>
<td>Engzell et al. (Netherlands)</td>
<td>8</td>
<td>Elementary and Secondary</td>
<td>350,000</td>
<td>Reading, Mathematics</td>
<td>DiD regression</td>
<td>$-0.09\ SD$ (reading) $-0.14\ SD$ (math)</td>
<td>+60% learning loss in low-SES students</td>
</tr>
<tr>
<td>Gore et al. (Australia)</td>
<td>8</td>
<td>Elementary</td>
<td>$&gt;4,800$</td>
<td>Reading, Mathematics</td>
<td>Linear mixed models</td>
<td>$+0.04\ SD$ (reading) $+0.06\ SD$ (math)</td>
<td>Grade 3 in math: $-0.16\ SD$ for low school-level SES, $+0.15\ SD$ for medium school-level SES</td>
</tr>
<tr>
<td>Kuhfeld et al. (USA)</td>
<td>8</td>
<td>Elementary and Secondary</td>
<td>$\approx7$ Million</td>
<td>Reading, Mathematics</td>
<td>Change score analyses</td>
<td>$-0.13\ SD$ to $-0.25\ SD$ (math)</td>
<td>Likely differences for different ethnicities</td>
</tr>
<tr>
<td>Maldonado and De Witte (Belgium)</td>
<td>7</td>
<td>Elementary</td>
<td>$&gt;4,000$</td>
<td>Reading, Mathematics, Social Sciences, Science, French, Mathematics</td>
<td>DiD regression</td>
<td>$-0.29\ SD$ (reading) $-0.19\ SD$ (math)</td>
<td>Larger effect in low-SES students</td>
</tr>
<tr>
<td>Authors (Country)</td>
<td>Duration of closure in weeks</td>
<td>School type</td>
<td>Sample size</td>
<td>Subjects</td>
<td>Statistical method</td>
<td>General effect</td>
<td>Differential effects</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>----------</td>
<td>-------------------</td>
<td>----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Meeter (Netherlands)</td>
<td>8</td>
<td>Elementary</td>
<td>~95,000</td>
<td>Reading, Mathematics</td>
<td>ANOVA</td>
<td>+0.20 SD</td>
<td>−</td>
</tr>
<tr>
<td>Schult et al. (Germany)</td>
<td>8</td>
<td>Secondary</td>
<td>~80,000</td>
<td>Reading, Mathematics</td>
<td>Change score analyses</td>
<td>−0.07 SD (reading)</td>
<td>−0.07 to −0.03 SD (math)</td>
</tr>
<tr>
<td>Spitzer and Musslick (Germany)</td>
<td>8</td>
<td>Elementary and Secondary</td>
<td>&gt;2,500</td>
<td>Mathematics</td>
<td>Linear mixed models</td>
<td>−2.43% (error rate)</td>
<td>Larger improvements in low-achieving students</td>
</tr>
<tr>
<td>Tomasik et al. (Switzerland)</td>
<td>8</td>
<td>Elementary and Secondary</td>
<td>26,685</td>
<td>Reading, Mathematics</td>
<td>Growth curve models</td>
<td>−0.37 SD (elementary school)</td>
<td>−0.10 SD (secondary school)</td>
</tr>
<tr>
<td>van Der Velde et al. (Netherlands)</td>
<td>8</td>
<td>Secondary</td>
<td>133,450</td>
<td>French</td>
<td>Linear mixed models</td>
<td>+0.25 SD (correct solutions to open questions)</td>
<td>−</td>
</tr>
</tbody>
</table>

Hammerstein et al.’s (2021) analyses of the 11 studies yielded mixed results with regard to the initial COVID-19 pandemic-related transition to virtual instruction in the spring of 2020. However, most of the 11 studies reported a decline in students’ reading and math achievement. For example, seven studies of math achievement and five studies of reading achievement revealed an overall decline. Hammerstein et al. (2021) stated, “The median for the effect [size] on mathematics and reading [achievement] is comparable when averaging all studies ($d = -0.10$ $SD$ and $-0.09$ $SD$ for mathematics and reading, respectively)” (p. 5). The researchers also reported that three studies revealed a positive increase in students’ math achievement after the transition to virtual instruction. In all three cases, students used the same curricula online as were used during F2F instruction. Two studies showed evidence of learning gains on reading achievement tests. However, the researchers posited the possibility that the reading achievement tests may have been influenced by improved online teaching strategies implemented by teachers and school districts in the fall of 2020 after the initial crisis or rapid transition to online instruction in the spring of 2020.

Lewis et al. (2021) conducted a study of students’ math and reading achievement in third through eighth grades to examine academic progress during the first full academic year after the onset of the COVID-19 pandemic. The purpose of the study was to provide insight and data to educators and policymakers to improve the overall state of education after the worldwide pandemic. Similar to previous researchers, the authors wanted to gather evidence related to two research questions. First, how do academic gains for the 2020-2021 school year compare to trends before the pandemic? Second, how do students’ reading and math achievement in the spring of 2021 compare to prepandemic levels?
Lewis et al.’s (2021) study was a continuation of the NWEA’s body of research on the influence of the COVID-19 pandemic-related transition to virtual instruction during the early stages of the pandemic. The sample size was five million students in Grades 3 through 8. The researchers used NWEA’s MAP growth assessment data, which measures student achievement on the Rasch Unit (RIT) scale. The RIT is an equal-interval scale, allowing scores to be added together to compare achievement scores across students and over time. Results from the MAP assessments for reading and math during the 2018-2019 school year were used as the baseline for comparison to the MAP assessment results from the 2020-2021 school year. RIT scores were used throughout the entire 2020-2021 school year to answer the question regarding growth over a year. The researchers used the average within-student RIT change score to calculate student gains. RIT scores for fall, winter, and spring were compared to determine gains from the baseline school year in 2018-2019. Along with the RIT scale, the researchers also compiled achievement percentile ranks for reading and math achievement. The percentile rank comparisons are depicted in Figure 4.

The results of the percentile rank comparisons revealed small reductions in reading across all grade levels and larger reductions in math across all grade levels.

Lewis et al. (2021) reported that students in Grades 3 through 8 made fewer gains in reading and math during the 2020-2021 school year when compared to the 2018-2019 prepandemic school year. The researchers determined that a comparison of the trajectories of the
prepandemic baseline scores and postpandemic scores revealed similar fall-to-winter trajectories, then diverged during winter-to-spring for both reading and math at all grade levels in the sample. When compared to the growth during the 2018-2019 school year, student reading and math gains stalled in the second half of the 2020-2021 school year.

Additional results of Lewis et al.’s (2021) study revealed lower reading and math RIT scores of students in third through eighth grades in the 2020-2021 school year when compared to the 2018-2019 prepandemic school year. The median percentile ranks for reading assessments for students in third through fifth grades in the spring of 2021 declined by an average of five percentile points from the spring of the 2018-2019 school year. The researchers also found that median scores of reading assessments dropped six percentile points for third graders, from 57% to 51%. The decline for third-grade students on the 2020-2021 math assessment from the prepandemic levels was 12% and 11% for both fourth- and fifth-grade students. The researchers uncovered greater declines for elementary school students in both math and reading compared to secondary students. Map Growth assessment in sixth, seventh, and eighth grades revealed a decline of approximately three percentile points for reading and eight percentile points in math. Unfortunately, the authors did not report whether the differences were statistically significant.

In February 2022, Dawson released a research report by Curriculum Associates detailing the results of a quantitative study of reading and math achievement of third- through eighth-grade students to examine possible differences in academic growth between prepandemic performance scores and during-pandemic performance scores. Reading and math achievement was measured using Curriculum Associate’s i-Ready diagnostic assessments for math and reading, which are administered at the beginning, middle, and end of each academic year. The i-Ready diagnostic tests are criterion-referenced assessments of student performance over time and
are used to guide both teacher-led and computer-based instruction of specific skills in math and reading. Dawson (2022) collected i-Ready diagnostic assessment data from students in kindergarten through eighth grades from seven testing periods from fall 2016 to fall 2021. The researcher used data from students who took each of the seven diagnostic assessments. The only assessment period without sufficient data was spring 2020, when most schools were closed for F2F instruction or when teachers did not administer the assessment. Two cohorts were constructed from a total of 2,039,367 students who took all the reading assessments and 2,432,643 students who took all the math assessments. The first cohort, described as the pre-COVID-19 cohort, consisted of students who took i-Ready assessments in each of the three assessment periods during fall 2016 through fall 2018 and students who took each assessment from fall 2017 through fall 2019. The second cohort, called the COVID-19 cohort, consisted of students who were tested during each of the three assessment periods from fall 2019 to fall 2021.

Curriculum Associates added a question to the math and reading computer-based i-Ready assessments in the fall of 2020 to determine whether the students were taking the test in the school building or at a different location. If the student needed more than one session to complete an assessment, the question regarding the testing location was asked each time the student logged into the assessment. If students reported that they took the test in different locations, the students were removed from the sample during the data-cleaning process. Dawson reported that 92% of testers for the fall 2021 assessment reported taking the test in school. Other variables under study included students’ grade level, school-level demographics, socioeconomic characteristics, and school location.

Dawson (2022) subsequently disaggregated the data from the COVID-19 cohort into four mutually exclusive groups based on the students’ testing location: in school only, mostly in
school (if the student tested remotely for one test), mostly remote (if the student only tested once
in the school building), and remote-learning only. The researcher used responses to the self-
reported testing location question in the fall of 2020 as a proxy to determine the location where
student instruction took place; however, Dawson recognized that the proxy was based on an
assumption and was not a perfect indicator of the location of each student’s instruction.

Dawson (2022) used a three-level, piecewise longitudinal growth model to adjust for the
measurement of non-linear growth across a school year. The results of the analyses revealed that
students in the COVID-19 cohort showed less overall growth in reading and math compared to
the pre-COVID-19 cohort. On average, the researcher reported few differences in math or
reading growth in the COVID-19 cohort between students who tested mostly or only in school
versus students who tested only or mostly at home. However, larger differences in growth
between the pre-COVID-19 cohort and the COVID-19 cohort were observed among students
who were two or more grade levels below expectations at the start of the assessment period. For
example, below-level students in second through fourth grade from the pre-COVID-19 cohort
improved an average of 50 scale score points on the reading assessment from the first to the last
assessment. Below-level students in second through fourth grade in the COVID-19 cohort
demonstrated an average growth rate of 43 scale score points in reading from the first assessment
in the fall of 2019 to the final assessment in the fall of 2021. However, students who were on
grade level at the start of the assessment period in the fall of the 2019-20 school year showed a
mean difference of two scale score points between the prepandemic and the COVID-19 cohorts.
The comparisons of mean reading scale scores are depicted in Figure 5.
Dawson (2022) suggested that a cursory glance at the mean scale scores for the reading and math i-Ready assessments might mask important differences between subgroups within the overall sample of the study. Although most subgroups of students experienced less growth during the pandemic compared to the prepandemic period, students who were two or more grade levels behind at the start of an assessment period demonstrated lower scale scores.

Dawson (2022) also uncovered an interesting discrepancy in growth for the pre- and postpandemic cohorts related to scale score differences based on ethnicity and income. Schools with more than 75% White students and greater than 200% income-to-poverty ratio...
demonstrated a drop in scale score gains on the reading assessment from an average of 51 scale score points during the pre-COVID-19 cohort to an average of 49 scale score points during the COVID-19 cohort testing period. Conversely, schools with fewer than 25% White students and less than 200% income-to-poverty ratio exhibited an average decline of seven scale score points from an average of 49 point gains to an average gain of 42 points on the reading assessment. Based on the lack of growth in both reading and math after the onset of the pandemic, Dawson concluded that school districts need to plan intensive and strategic methods to accelerate academic growth, especially among below-grade-level students and students from certain demographics.

Summary

This review of recent literature on school closures and virtual instruction focused on student achievement and described studies using different methodologies, comparing different populations and samples, and demonstrating mixed results. Studies by Kuhfeld et al. (2020), Kogan and Lavertu (2021), and Pier et al. (2021) provided evidence of differential levels of learning loss based on ethnic and socioeconomic factors in reading and math during the pandemic. However, studies by Domingue et al. (2021), Gore et al. (2021), Hammerstein et al. (2021), and Dawson (2022) provided evidence of either little or no learning loss nor a differential influence of the school closures and remote instruction on growth in reading and math. Taken as a whole, the preponderance of evidence in the literature review substantiates the differential influence of school closures and remote instruction on young students, below-level students, students of color, ELL and ESE students, homeless students, and students from low-income homes. For these students, research-based interventions must be implemented to reduce the learning gaps found in the current literature. The research base on student achievement is
growing and should continue for many years to assist future researchers in their efforts to identify and promote the lessons learned during the pandemic.
III. METHODOLOGY

The purpose of this study was to explore the reading performance scores of elementary school students before and after school closures due to COVID-19 in one school district. This study was designed to examine one research question: What were the reading performance scores of third-grade students before and after the COVID-19 pandemic? The research sample included a single cohort of Grade 3 students as they progressed through the grade levels over time. Reading performance was measured by the students’ scale scores on i-Ready diagnostic tests used by the district to track reading development. The research design was a nonexperimental, exploratory study of quantitative, archival data.

Data Collection

After approval by the Institutional Review Board at Southeastern University, the researcher submitted an official request to the target school district to obtain student-level and district-level reading performance data from the district’s Performance Matters (Unify, 2020) database, which included i-Ready composite scale scores for each student for five successive administrations of i-Ready’s diagnostic assessments. In addition, the researcher requested the students’ demographic data to include ESE designation, ELL designation, and information on whether the students had been retained at any point in their education.
Sample

Approximately 41,000 students in the target school district were enrolled in kindergarten through Grade 12 during the 2020-2021 school year (Florida Department of Education, 2021b). The race and ethnic origins of the target district were identified as 48% White, 20% Black or African American, 25% Hispanic, 3% Asian or Pacific Islander, and 5% other. The percentage of households with a computer was 88%, and 81% of households had broadband internet access in the home. The entire district is designated Title I. The population in this study consisted of all third-grade students in the district \( N = 3,034 \) during spring 2020. The research sample consisted of a single cohort of 2,006 third-grade students who were not designated as ESE or ESL and who had not been retained at any point in the past. Students with incomplete data were eliminated from the research sample. The cohort students’ reading scores were tracked from the prepandemic assessment period of winter 2020 in third grade through fall 2021, which was the first semester of fifth grade.

Reading performance data for the research cohort included individual students’ i-Ready scale scores from five different assessment periods: January 2020 (Grade 3, prepandemic), August 2020 (Grade 4, fall semester), January 2021 (Grade 4, mid-year), April 2021 (Grade 4, end of academic year), and August 2021 (Grade 5, fall semester).

Instrumentation

This study measured reading performance over time using each student’s mean composite reading scale score from i-Ready diagnostic tests from the five designated assessment periods. The diagnostic reading assessments are computer-based adaptive tests that measure overall reading comprehension and student skill levels for each stage of reading and math skill development in Grades K-12. The publisher of the i-Ready program, Curriculum Associates,
created the diagnostic reading assessments to measure elementary students’ discrete skill development in phonological awareness, phonics, high-frequency words, vocabulary, reading comprehension for literature, and reading comprehension for informational text.

Each i-Ready diagnostic assessment provides an overall scale score that represents a comparison of student performance to grade-level norms and grade-level designations for each of the measured stages of reading (Curriculum Associates, 2019). For example, a second-grade student might earn a scale score of 432 in the first assessment period, which represents reading at a first-grade level. Along with the scale score and accompanying grade-level equivalent, the same student might test out of the phonological awareness test and perform at first-grade level for phonics, second-grade level for high-frequency words, and first-grade level for vocabulary, literature comprehension, and comprehension of informational text.

Each elementary student who participates in the i-Ready program in the target district is administered the diagnostic tests at the beginning, middle, and end of each academic year via computer. After completing the first administration, the student receives a scale score that represents a goal of approximately one year’s growth as a summative measure (Curriculum Associates, 2019). Students also receive a stretch goal, which is designed to quantify larger steps for below-level readers to assist them in reaching grade level by the end of the school year. These target scores can assist teachers as they form guided reading groups and individualize reading instruction.

The diagnostic tests are computer-based and adaptable to individual student progress. For example, as students progress through instruction on phonics, the online assessment of that skill decreases in frequency, while comprehension-based skill assessments increase in frequency.
As the student’s reading ability improves throughout the school year, growth is reflected in subsequent assessment results.

**i-Ready Validity and Reliability**

The development of the i-Ready Diagnostic Assessment was led by a committee of experts in educational measurement and computer adaptive technology from members of various institutions of higher education and organizations promoting literacy (Curriculum Associates, 2014). The i-Ready program was designed to align with the Common Core State Standards, Next Generation assessments, and the Partnership for Assessment of Readiness for College and Careers assessment (Colorado Department of Education, 2016). The i-Ready items and assessments were extensively reviewed for content validity and reliability by educational committees in many states, including Colorado, Illinois, Massachusetts, New York, Ohio, Texas, and Virginia (Curriculum Associates, 2014). In addition, content experts reviewed all i-Ready reading passages to determine text complexity.

Curriculum Associates collaborated with MetaMetrics to conduct a Lexile linking study to gather evidence of concurrent validity. As cited in the 2014 report from Curriculum Associates, the Lexile linking the study of reading performance scores of first, third, fifth, and seventh grades revealed strong correlation coefficients ranging from $r = .88$ to $r = .89$ across the four grade levels in reading (Curriculum Associates, 2014, p. 20). The assessments also demonstrated reliability with strong test metrics, including low standard error of means and strong levels of item discrimination for students of different ability levels.

To develop national norms, Curriculum Associates (2014) conducted a large-scale pilot program that assessed more than a million students in kindergarten through eighth grades (Curriculum Associates, 2014, p. 20). After the third year of the pilot program, data from the
diagnostic assessments were collected from a nationally representative sample to create i-Ready’s national norms. According to a report from the Colorado Department of Education (2016), Curriculum Associates conducted the study with independent researchers from the University of Massachusetts to determine annual growth targets and cutoff scores for students (Colorado Department of Education, 2016, p. 7). The national norms used in the comparisons conducted in the current study were derived from the 2018-2019 academic year. The district norms used in comparisons were also from the 2018-2019 academic year.

**Data Analyses**

After receiving the archived datasets from the district, the data were cleaned in an Excel document to identify missing data and any other problems that might compromise the integrity of the datasets. In addition, the data were disaggregated to eliminate students who were identified as ESE and ELL and students who had been retained; these students will be examined in future studies. From the resulting dataset, the reading performance scores of 2,006 students were utilized in the data analyses.

Preliminary analyses were conducted to determine the research sample’s dataset for normality using Q-Q tests. The internal consistency of the dataset was computed using Cronbach’s alpha.

Descriptive statistics were computed to determine the mean, standard deviation, range, standard error of measurement, skew, and kurtosis of each of the five assessment periods under study. In addition, the researcher conducted five separate one-sample *t* tests to compare the mean composite scale score of the research cohort to i-Ready’s mean scale score from the district’s 2018-2019 norm group for the appropriate grade levels and assessment periods (early, middle, and late school year). The same analyses were also conducted to compare the i-Ready scale
scores of the research cohort to the 2018-2019 national i-Ready norms. The one-sample $t$ tests compared each of the five assessment periods: January 2020 (Grade 3 prepandemic), August 2020 (Grade 4 after schools reopened in the fall term), January 2021 (Grade 4 winter assessment period), May 2021 (Grade 4 spring assessment period), and August 2021 (Grade 5 fall term). The reader will note that the spring 2020 administration of i-Ready tests was not completed due to the pandemic, school closures, and the introduction of mandated remote instruction. The results of the data analyses are described in Chapter 4.
IV. RESULTS

The purpose of this study was to explore the reading performance scores of elementary school students before and after school closures due to COVID-19 in one school district. This study was designed to examine one research question: What were the reading performance scores of third-grade students before and after the COVID-19 pandemic? The research design was a nonexperimental, exploratory study of quantitative archival data. The study was designed to explore a single cohort of third-grade students’ i-Ready scale scores in reading from five different assessment periods: January 2020, August 2020, January 2021, April 2021, and August 2021. The reading scores were compared to national norms and target-district norms in reading from the 2018-2019 school year to address the research question.

Methods of Collecting Data

After approval by the Institutional Review Board at Southeastern University, the researcher submitted an official request to the target school district to obtain student-level and district-level reading performance data from the district’s Performance Matters (Unify, 2020) database, which included i-Ready composite scale scores for each student for five administrations of i-Ready’s diagnostic assessments. The researcher requested the students’ demographic data to include ESE designation, ELL designation, and information on whether the students had been retained. In addition, the researcher also requested i-Ready diagnostic norms from the 2018-2019 school year from the target school district and national norms from the
developers of i-Ready, Curriculum Associates.

**Sample**

The population in this study consisted of all third-grade students in the target district \((N = 3,034)\) in the 2019-2020 academic year. The research sample consisted of a single cohort of 2,006 third-grade students who were not designated as ESE or ELL and who had not been retained at any point in the past. Individual students with missing reading scores were eliminated from the research sample. The cohort students’ reading scores were tracked by the district and shared with the researcher to include the prepandemic assessment period of winter 2020 in third grade through fall 2021 when the students were in fifth grade. The majority of the district’s students were back in F2F classrooms in the fall of 2020.

Reading performance data for the research cohort included individual students’ i-Ready scale scores in reading from five different assessment periods: January 2020 (Grade 3, prepandemic), August 2020 (Grade 4), January 2021 (Grade 4), April 2021 (Grade 4), and August 2021 (Grade 5).

**Preliminary Analyses and Results**

After the research data were cleaned, the internal consistency of the results of the cohort’s reading scale score means was established by conducting a Q-Q plot. The residuals of each assessment period of the Grade 3 cohort were found to satisfy the assumptions of normality. In addition, the dependent variables (assessment period outcomes) were all normally distributed (skew \(-/+ 2.0\)) and (kurtosis \(-/+ 7.0\)). Finally, Cronbach’s alpha was computed to determine the internal reliability of the data from the five assessment periods. The result of the Cronbach alpha analysis revealed a high level of internal consistency in the dataset across five assessment periods, \(\alpha = .95\), 95% CI [0.95, 0.95].
Descriptive Analyses and Results

Descriptive statistics were computed to describe the research cohort’s reading performance before and after the school closures during the COVID-19 pandemic. The descriptive results are presented in Table 5.

Table 5

Descriptive Statistics Summary Table

<table>
<thead>
<tr>
<th>Assessment Period/Time</th>
<th>Grade Level</th>
<th>M</th>
<th>SD</th>
<th>SEM</th>
<th>Min</th>
<th>Max</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2020*</td>
<td>3</td>
<td>525.23</td>
<td>43.04</td>
<td>.96</td>
<td>312.00</td>
<td>643.00</td>
<td>−0.33</td>
<td>0.28</td>
</tr>
<tr>
<td>August 2020</td>
<td>4</td>
<td>532.76</td>
<td>46.16</td>
<td>1.03</td>
<td>342.00</td>
<td>668.00</td>
<td>−0.30</td>
<td>0.17</td>
</tr>
<tr>
<td>January 2021</td>
<td>4</td>
<td>545.59</td>
<td>47.26</td>
<td>1.06</td>
<td>298.00</td>
<td>683.00</td>
<td>−0.47</td>
<td>0.72</td>
</tr>
<tr>
<td>May 2021</td>
<td>4</td>
<td>553.55</td>
<td>49.45</td>
<td>1.10</td>
<td>306.00</td>
<td>690.00</td>
<td>−0.51</td>
<td>0.63</td>
</tr>
<tr>
<td>August 2021</td>
<td>5</td>
<td>556.35</td>
<td>48.40</td>
<td>1.08</td>
<td>349.00</td>
<td>691.00</td>
<td>−0.47</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Note. N = 2,006. M = mean; SD = standard deviation; SEM = standard error of measurement; Min = minimum; Max = maximum.

*These data were the baseline scores for the Grade 3 cohort. Students did not take the May 2020 i-Ready assessment, and all were promoted to Grade 4.

Inferential Analyses and Results

The Grade 3 cohort’s i-Ready reading mean scores were compared to the 2018-2019 i-Ready mean scores of the school district’s norms using 5 one-sample t-tests. The results of the comparisons are presented in Table 6.
Table 6

*Mean Scale Score Comparisons of the Study Cohort and Mean i-Ready District Scale Scores by Assessment Period*

<table>
<thead>
<tr>
<th>Assessment Period</th>
<th>Cohort $M$</th>
<th>Cohort SD</th>
<th>District $M$</th>
<th>$t$</th>
<th>$p$</th>
<th>Cohen’s $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2020</td>
<td>525.23</td>
<td>43.04</td>
<td>519.00</td>
<td>6.48</td>
<td>&lt;.001</td>
<td>0.14</td>
</tr>
<tr>
<td>August 2020</td>
<td>532.76</td>
<td>46.16</td>
<td>529.00</td>
<td>3.64</td>
<td>&lt;.001</td>
<td>0.08</td>
</tr>
<tr>
<td>January 2021</td>
<td>545.22</td>
<td>47.26</td>
<td>543.00</td>
<td>2.10</td>
<td>.04</td>
<td>0.05</td>
</tr>
<tr>
<td>May 2021</td>
<td>553.16</td>
<td>49.45</td>
<td>551.00</td>
<td>1.96</td>
<td>.05</td>
<td>0.04</td>
</tr>
<tr>
<td>August 2021</td>
<td>556.35</td>
<td>48.40</td>
<td>551.00</td>
<td>4.95</td>
<td>&lt;.001</td>
<td>0.11</td>
</tr>
</tbody>
</table>

*Note. N = 2,006. M = mean; SD = standard deviation*

As Table 6 indicates, comparisons of each assessment period revealed statistically significant differences between the mean reading scale scores from the research sample and the mean 2018-2019 scale scores from the district norm group. In each assessment period, the research sample slightly outperformed the district’s 2018-2019 norm group. Although the results of each comparison were statistically significant, the actual scale score differences were small, and the effect sizes were small. In general, the results indicated that both groups showed small increments of change over time.

The Grade 3 cohort’s i-Ready reading mean scores were also compared to the 2018-2019 i-Ready’s mean scores on the national norms using 5 one-sample $t$ tests. One-sample $t$ tests were computed to compare scores from a research sample to known scores from the district and national mean score norms. The results of the comparisons are presented in Table 7.
Table 7

Results of One-Sample t-Test Comparisons of Cohort Mean Scale Scores to i-Ready National

Mean Scale Score Norms by Assessment Period

<table>
<thead>
<tr>
<th>Assessment Period</th>
<th>Cohort M</th>
<th>Cohort SD</th>
<th>National Norm M</th>
<th>t</th>
<th>p</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2020</td>
<td>525.23</td>
<td>43.04</td>
<td>552.5</td>
<td>-28.37</td>
<td>&lt; .001</td>
<td>0.60</td>
</tr>
<tr>
<td>August 2020</td>
<td>532.76</td>
<td>46.16</td>
<td>557.5</td>
<td>-24.00</td>
<td>&lt; .001</td>
<td>0.54</td>
</tr>
<tr>
<td>January 2021</td>
<td>545.22</td>
<td>47.26</td>
<td>590.5</td>
<td>-42.91</td>
<td>&lt; .001</td>
<td>0.96</td>
</tr>
<tr>
<td>May 2021</td>
<td>553.16</td>
<td>49.45</td>
<td>616.0</td>
<td>-56.91</td>
<td>&lt; .001</td>
<td>1.27</td>
</tr>
<tr>
<td>August 2021</td>
<td>556.35</td>
<td>48.40</td>
<td>616.0</td>
<td>-55.19</td>
<td>&lt; .001</td>
<td>1.23</td>
</tr>
</tbody>
</table>

Note. N = 2,006. M = mean; SD = standard deviation

Significant differences were observed in each of the five t-test comparisons, including the scale scores for the January 2020 pre-COVID-19 assessment period, which reflected the sample cohort’s baseline data. The mean reading scale scores of the sample were consistently lower than the national norm group in each assessment period. However, the effect sizes ranged from very small in January 2020 to medium in August 2020 to large in January 2021 and very large in May and August 2021.

Summary

The mean scale scores of the research cohort were compared to the school district’s 2018-2019 norms before and after the mandated school closures related to the pandemic. The results of the comparisons revealed that the research cohort’s mean reading scale scores were significantly different from those of the district norm group’s mean scale scores; the research sample’s mean scale scores were slightly higher than the district’s norm group. However, the mean scores and the effect sizes were small.
As Figure 6 indicates, the students in the sample group performed substantially lower in reading than the 2018-2019 national norm group prior to the pandemic in the spring of 2020 and the introduction of remote teaching and learning. Although the sample student group demonstrated progress over time, the rate of progress was not commensurate with the progress of the 2018-2019 national norm group scale scores. The results of the study will be discussed further in Chapter 5.

**Figure 6**

*Mean Scale Score Comparison by Assessment Period: Study Sample vs. i-Ready 2018-2019 Academic Year National Norms*
V. DISCUSSION

The purpose of this study was to explore the reading performance scores of elementary school students before and after school closures due to COVID-19 in one school district. The research design was a nonexperimental, exploratory study of archival quantitative data to examine one research question: What were the reading performance scores of a group of third-grade students before and after the COVID-19 pandemic? The research sample included a single cohort of Grade 3 students as they progressed through the fourth grade and were tested at the beginning of fifth grade. Reading performance was measured by the students’ scale scores on i-Ready diagnostic tests, which were used by the district to track reading development among all elementary students in the district.

Review of Methodology

After approval by the Institutional Review Board at Southeastern University, the researcher submitted an official request to the target school district to obtain individual student-level and district-level reading performance data from the district’s Performance Matters (Unify, 2020) database. The resulting dataset included i-Ready composite reading scale scores for each third-grade student in the district for five administrations of i-Ready’s diagnostic assessments before and after school lockdowns due to COVID-19 and the introduction of online instruction for all students.
The researcher also requested students’ demographic data to include ESE designation, ELL designation, and information on whether the students had been retained at any point during their education. For comparison purposes, the i-Ready diagnostic mean scale scores from the 2018-2019 (prepandemic) school year were obtained from the target school district, and the 2018-2019 national norms were obtained from the developers of i-Ready.

**Research Population and Sample**

In the target school district, approximately 41,000 students were enrolled in kindergarten through Grade 12 during the 2020-2021 school year (Florida Department of Education, 2021). The race and ethnic origins of the target district were identified as 48% White, 20% Black or African American, 25% Hispanic, 3% Asian or Pacific Islander, and 5% other. The percentage of households with a computer was 88%, and 81% of the households had broadband internet access. The entire school district is designated as a Title I district. The target population in this study consisted of all third-grade students in the target district during the 2019-2020 academic year.

Demographic data provided by the school district were disaggregated according to ESE, ELL, and retained students; these datasets will be examined in future studies. After ESE, ELL, and retained students and students with missing data were removed from the target population’s dataset, the final sample size of the research cohort was 2,006 third-grade students whose reading performance was followed during their fourth-grade school year.

**Instrumentation**

Every elementary student in the target district participated in the i-Ready program and was administered the reading diagnostic tests at the beginning, middle, and end of each academic year. The i-Ready diagnostic tests are computer-based and adaptable to individual student progress; for example, as students progress through instruction in phonics, the online assessment
of that skill set decreases in frequency, and comprehension-based assessments increase in frequency. As the student’s reading ability improves throughout the school year, growth is reflected in subsequent assessment results.

Each i-Ready diagnostic assessment provides an overall scale score that represents a comparison of student performance to grade-level norms and grade-level designations for each of the measured stages of reading (Curriculum Associates, 2019). For example, a second-grade student might earn a scale score of 432 in the first assessment period, which represents reading at a first-grade level. The same student might test out of the phonological awareness test, perform at first-grade level for phonics, second-grade level for high-frequency words, and first-grade level for vocabulary, literature comprehension, and comprehension of informational text. These assessment results are subsequently used to inform reading instruction by teachers.

Analyses

After disaggregating the research cohort’s i-Ready reading dataset based on demographic variables and removing students with missing data, the researcher conducted descriptive and inferential analyses of the third-grade cohort’s i-Ready reading scale scores from five different assessment periods: January 2020 (the Grade 3 mid-year assessment before COVID-19), August 2020 (the Grade 4 beginning-of-the-year assessment after schools in the district reopened in August), January 2021 (the Grade 4 mid-year assessment period), April 2021 (the Grade 4 end-of-the-year assessment period, and August 2021 (the Grade 5 beginning-of-the-year assessment). The inferential analyses included 5 one-sample t tests of the cohort’s reading scores over time to compare to the district and national i-Ready norm scores.
Summary of Results

Preliminary Results

Preliminary analyses of the research data revealed no missing data and that the data were normally distributed. Cronbach alpha analysis revealed a high level of internal consistency in the dataset, $\alpha = .95$. Descriptive statistics were computed to describe the Grade 3 cohort’s reading performance before and after the school closures due to the COVID-19 pandemic (see Table 5 in Chapter 4). The students’ mean scale scores after the January 2020 assessment revealed small increments of reading growth over time after school reopened in the fall of 2020.

Inferential Results

The researcher conducted 5 one-sample $t$ tests to compare the research cohort’s mean reading scale scores to the target school district’s norms from the 2018-2019 academic year for the five assessment periods under study. As shown in Table 6 (see Chapter 4), the results of the $t$-test comparisons revealed that the research cohort’s mean reading scale scores were significantly different from those of the district norm group’s mean scale scores for academic year 2018-2019; the research sample’s mean scale scores were slightly higher than the district’s norm group. However, the mean scores, differences, and effect sizes were small.

The researcher also compared the research cohort’s mean reading scale scores to i-Ready’s mean scale scores of the most recent i-Ready national norm group (academic year 2018-2019) for the five assessment periods under study. The results of the $t$-test comparisons are presented in Table 7 in Chapter 4 and displayed in Figure 6 (see Chapter 4). The students in the research cohort performed significantly lower in reading performance when compared to the 2018-2019 national norm group. The differences were observed prior to the pandemic in the January 2020 assessment, just before the introduction of remote teaching and learning due to
COVID-19. In addition, the mean reading scale scores of the research cohort were consistently lower than the 2018-2019 national norm group in each of the five assessment periods. Although the research cohort students demonstrated progress over time, the rate of progress was not commensurate with the progress of the 2018-2019 national norm group.

**Implications**

The descriptive and inferential analyses conducted in this exploratory study yielded interesting results that may serve to inform and guide educational practice and future research. Many educational stakeholders were deeply concerned about learning loss when students returned to classrooms in the fall of 2020. In a meta-analysis of the relationships between summer vacation and student achievement in Grades 1 through 8 Cooper (1996) found that, on average, readers experienced a loss of 1.5 months in reading skills. However, the third-grade students in the current research study were able to maintain their reading skills during remote instruction and the summer months; in fact, the cohort’s mean reading scale score went up slightly from January 2020 to August 2020. Apparently, the three-month period of mandated remote instruction was effective in continuing to develop and maintain reading progress among this group of students. Whether the maintenance of reading skills was related to remote instruction, parental involvement, student maturation, or other factors cannot be known without further study.

After students in the state returned to classroom-based instruction in August 2020, the elementary students in the research cohort made very small levels of progress in reading performance over time. These results align with the results published in a research report for
Curriculum Associates (Dawson, 2022), which described small increments of growth in reading from fall 2020 to fall 2021 as measured by percentages of students who were reading on grade level (see Figure 7).

Figure 7

*i-Ready Diagnostic Reading Assessment Results From Fall 2020 and Fall 2021*

![Bar chart showing students reading on grade level, Fall 2020 and Fall 2021](image)

Adapted from *Student Growth during COVID-19: Grade-Level Readiness Matters*, by M. Dawson, 2022, Curriculum Associates: Research and Efficacy, p. 9


Each year, Curriculum Associates publishes an annual report entitled “State of Student Learning.” The annual reports describe progress in math and reading of large numbers of students in Grades 1 through 8 as measured by i-Ready diagnostic assessments. The 2022 annual report was of particular interest since it compared reading progress longitudinally and served to
assist educators and researchers as they evaluated the influence of the pandemic, school lockdowns, and remote instruction. The results of the longitudinal comparisons in reading performance, typically conducted in May or June of each school year, are represented in Figure 8.

**Figure 8**

*i-Ready Diagnostic Reading Assessment Results from Spring 2021 and Spring 2022*

As indicated in Figure 8, the percentages of students on grade level in Grades 1 and 2 appeared to drop dramatically after the introduction of online instruction in March 2020 and continued after students returned to classroom-based instruction through fall 2022. By third grade, the students’ scores approached historical performance levels by the end of 2021 and
2022, indicating that the students were closing the learning gaps. However, overall percentages of students who were on-level in reading dropped again in Grades 4 through 8. These results point directly to reading development theory.

According to Chall’s (1995) reading development theory, students develop critical reading concepts, skills, and dispositions in an orderly, hierarchical fashion; skillsets developed in early childhood and the primary grades provide the foundation for all future reading development. When the development of the basic skills of reading is interrupted or compromised, especially in early elementary grades, learning gaps can manifest in higher grade levels. Many comprehensive research studies substantiate that students who do not learn to read on grade level by Grade 3 are likely to continue to need reading support throughout their lives (Aber et al., 2013). Critical periods exist during which reading skill development is optimal and maximized; these critical periods build the foundation for continued reading development into adulthood. Theoretically, the possibility of learning loss or reduction in reading performance would be greater for primary-grade students who were still learning the basic concepts and skills of reading and who might have difficulty following instructions and responding appropriately in online formats due to unfamiliarity with keyboards. On the other hand, third-grade readers probably would have already covered and mastered the basics of phonics, decoding, and other elements of reading taught in the primary grades and could be considered independent readers, which would enable them to fully engage in remote instruction.

This theoretical assumption was confirmed in the annual report published by Curriculum Associates (2022), which pointed out the challenges and difficulties of teaching the basic skills of reading remotely to young children. The results of extensive analyses of i-Ready data in the 2022 annual publication revealed that the percentages of first-, second-, and third-grade students
who were below grade level in phonics increased over time (Figure 9) during remote instruction due to the pandemic.
Figure 9

Percentage of Students Below Grade Level in Phonics Skills on i-Ready Diagnostic Assessments


Whether the increases in percentages of below-level students in phonics were directly related to remote instruction is debatable and deserves further investigation. In any case, Chall’s (1995) work on reading development over time provides an important foundation for understanding the COVID-19 research and mitigating the influence on young children’s overall reading development as well as reducing the number of older students who struggle to read.

District Norm Comparisons

In the current study, the district and national norm group comparisons from the 2018-2019 academic year were important barometers of change considering the norms reflected
children’s reading scores during regular classroom instruction. When the research cohort’s mean reading scale scores were compared to the 2018-2019 district-level scores, the cohort students’ scale scores were significantly higher, although the differences and the effect sizes were very small. In the target district, i-Ready’s online instructional tools and assessments were familiar to all elementary students because i-Ready had been implemented as a part of daily supplemental reading and math instruction since the 2015-2016 school year. When all instruction moved to online formats in March 2020, teachers and students at the elementary level continued to use i-Ready supplemental math and reading curricula to maintain and advance skill levels. This familiarity may have enabled teachers and students to adapt more easily to online teaching and learning, which may have influenced the small differences observed between the research cohort’s reading performance scores and the district’s reading performance scores from the 2018-2019 academic year.

Although most students in the state and in the target district returned to classroom-based instruction in August 2020 (Marshall & Bradley-Dorsey, 2020), the cohort’s reading progress continued to increase in very small increments, yet commensurate with the district norms from the previous school year. These results may relate to those of Dawson (2022), who found that mean reading improvement scores on i-Ready assessments were not dramatically different from mean scores prior to COVID-19 among second- through fourth-grade students who were reading on grade level, regardless of instructional delivery method (in school, mostly in school, mostly remote, or remote only). However, Dawson reported that reading improvement scores declined dramatically among students who were reading two or more grade levels below norms, regardless of instructional delivery methods. Unfortunately, the author provided only descriptive
results, and no inferential results were presented for consideration. Dawson’s results were especially pertinent to the target district because the entire district is designated as a Title I district.

**National Norm Comparisons**

The cohort students’ differences in reading performance were most apparent when their scores were compared to 2018-2019 national i-Ready norms. An examination of the comparisons’ results revealed that the two groups were significantly different across all assessment periods, with medium effect sizes in January 2020 and August 2020, large effect sizes in January 2021, and very large effect sizes in May and August 2021. In other words, as the cohort students progressed through fourth grade, the differences between 2018-2019 pre-COVID national norms grew larger and larger. The scale score differences between the cohort’s scores and the 2018-2019 national norms are easily observed in Table 8.

**Table 8**

*i-Ready Scale Score Differences Between Means of Research Cohort and 2018-2019 National Norms*

<table>
<thead>
<tr>
<th>Assessment Period</th>
<th>Grade Level</th>
<th>Cohort Mean</th>
<th>2018-2019 National Mean</th>
<th>Point Value Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2020</td>
<td>3</td>
<td>525.23</td>
<td>552.50</td>
<td>27.27</td>
</tr>
<tr>
<td>August 2020</td>
<td>4</td>
<td>532.76</td>
<td>557.50</td>
<td>24.74</td>
</tr>
<tr>
<td>January 2021</td>
<td>4</td>
<td>545.22</td>
<td>590.50</td>
<td>45.28</td>
</tr>
<tr>
<td>May 2021</td>
<td>4</td>
<td>553.16</td>
<td>616.00</td>
<td>62.84</td>
</tr>
<tr>
<td>August 2021</td>
<td>5</td>
<td>556.35</td>
<td>616.00</td>
<td>59.65</td>
</tr>
</tbody>
</table>

Perhaps these scores would be different if the parents or the district had opted to create opportunities for tutoring or reading camps during the summer of 2020. Unfortunately, many
parents were still concerned about the spread of COVID-19 during that summer and might not have participated if the resources had been offered. In any case, the results of the national comparisons point to the need for further investigation into children’s reading development after radical, unexpected changes related to school interruptions, the introduction of exclusively remote instruction, or other factors. Was a critical reading skillset overlooked or missed during a critical period of reading development during remote instruction that influenced a measurable, cumulative decline in reading performance later among the cohort students?

Mitigating Factors

In a recent qualitative study of teachers who delivered instruction remotely during the 2020 school lockdowns, Ward (2022) identified mitigating factors that may have influenced the academic performance of students: the initial chaos related to the rapid deployment of remote instruction, the teachers’ adjustment to teaching with technology, and inconsistent access to online curricula and technology tools in impoverished and mountainous areas (pp. 80-81). The teachers in the study reported that absenteeism and inconsistent internet access for both teachers and students had an enormous influence on student performance (Ward, 2022, pp. 81-82). One teacher in Ward’s (2022) study stated, “Only one-third of the students attended class on a daily basis” (p. 112). In this relatively small district, school and district administrators and the local community valiantly attempted to provide both technical and curricular support to the teachers and students during the unprecedented changes related to the pandemic. However, all the teachers in the study agreed that they were unprepared to transition rapidly to remote instruction and that student engagement with the curricula and with age peers was radically diminished during remote instruction in the spring of 2020.
The problem of absenteeism in classroom-based instruction has been extensively studied in the past (Marcotte & Hemelt, 2008). Educators know that time on task, attention, and focus are critically important to students’ academic success and overall development as educated people. Online and hybrid instruction cannot be effective if students do not participate. How can a school ensure that students remain engaged with the material, the instructor, and their peers, just as they would in classroom-based instruction? How can parents ensure that their children are present and engaged during online instruction? Educational stakeholders, including district and school administrators, teachers, students, and parents, could collaborate to conduct a review of recent research studies, technological tools, and instructional strategies related to the problems of absenteeism and student disengagement during remote instruction. The result of the collaboration might lead to the development of consistent school-wide guidelines and expectations for participation in remote instruction.

As teachers continue to use remote instructional delivery methods, whether via hybrid or exclusively online formats, students must be accountable for being present and actively engaged during direct instruction, small group interactions, and student projects. Teachers can acquire and polish the skills and dispositions to make online instruction developmentally appropriate, engaging, and challenging. In addition, teachers can create and share effective online material and new ways to assess learning. Schools and districts can make accommodations to teachers’ schedules to promote and provide outstanding and easily accessed professional development to teachers as they continue to develop technical skills and curricular expertise to provide high-quality teaching and learning.

In June 2020, Dorn et al. developed a statistical model to estimate the potential influence of school closures due to the pandemic on students’ achievement. The model was based on
previous research studies related to the effectiveness of remote learning compared to classroom-based learning. The researchers predicted that sixth-grade students would lose 3 to 4 months of learning in math compared to typical in-class teaching and learning. For those students who participated in low-quality remote instruction, the anticipated learning loss in math was 7 to 11 months. Students who chose not to participate in remote education at all could anticipate 12 to 14 months of learning loss. Clearly, the quality of remote instruction is a mediating factor when discussing learning loss during the pandemic. Unfortunately, the rapid transition from F2F instruction to online delivery of instruction in March 2020 afforded little time for planning and implementation. Hopefully, one of the big lessons learned from the pandemic is the need for suitable learning management systems and professional development of teachers so they are prepared for future disruptions in educational programming.

Dorn et al. (2020) also predicted a differential influence of school closures on learning for minority students and those from low-income homes. These predictions were especially pertinent to the target school district, which is a Title I district. The results of the Dorn et al.’s modeling suggested that high-school students who participated in remote instruction from March 2020 to January 2021 were at-risk for 6 to 7 months of learning loss in mathematics. Minority students and low-income students’ risks for learning loss in mathematics were even greater. The authors suggested that the level of learning loss during remote instruction in the past was primarily related to student disengagement, lack of technology, and lack of internet access. After schools reopened in the target state in August 2020, many school districts in the state used stimulus money to acquire laptops, iPads, or tablets for each student to use at school and to take home to promote continued learning in the future.
Dorn et al. (2020) also predicted that high-school dropout rates would dramatically increase after school closures based on previous studies of interrupted educational programming. For example, after Hurricane Maria in 2017, 14 to 20% of high-school students never returned to high school (Declining Enrollment, Shuttered Schools: Puerto Rico’s Education System in Numbers, 2019). If Dorn et al.’s (2020) predictive modeling is found to be accurate in future studies, the influence on the American economy and society could be enormous. The authors offered recommendations to jump-start educational responses to alleviate the deleterious influence of school closures: tutoring programs, summer school, parent education, improvements in online curriculum design and implementation, and professional development for teachers on best practices in remote teaching and learning. School and district administrators can also assist by providing technological tools, technical support, flexible staffing, and different instructional models.

In many classrooms, the predictions of disastrous learning loss in reading and math after school lockdowns have been realized in 2022, especially among young children (Dawson, 2022), minority students (Dorn et al., 2021), and students from low-income homes (Gore et al., 2021). Clearly, much work is needed to ensure that all children and young people overcome any losses during the pandemic. The lessons learned during the pandemic can assist educators as they move away from the chaos and uncertainty during the pandemic and move forward into more orderly, planned change and innovation. In addition, future research studies can continue to inform and assist educators as they make the transitions.

**Recommendations for Future Research**

The worldwide nature of the COVID-19 pandemic and its relation to education presented many opportunities for ongoing research. This study should be replicated using the existing
datasets to examine the influence of the pandemic on ESE and ELL students’ reading performance. In addition, reading performance trends could be tracked by the district or other researchers over a larger number of assessment periods after remote instruction.

Another important study might explore the performance of different grade levels during the same period and for math performance. As mentioned earlier, Chall’s work (1995) on reading development suggests that primary-grade students might experience greater difficulties during remote instruction considering many of them cannot yet read. In a study on frequent school closures and student reading performance, Marcotte and Hemelt (2008) described a correlation between frequent school closures and lower reading performance for students in third grade and below but not for students in fourth through eighth grades (p. 19). A similar study should be replicated and expanded to include a study of kindergarten, first- and second-grade students’ overall reading performance scores after remote instruction and the longitudinal development of reading skills as the students progress through the grade levels. A study of this type might lend greater insight into both remote reading instruction and reading development theory. With regard to reading development theory, a study of K-2 students’ reading performance during remote instruction might provide more understanding of the early steps in the process of learning to read and the influence of technology on normal reading development.

Researchers might also study remote instruction during the initial break from F2F instruction during the spring of 2020 and throughout the pandemic period in different districts and states. In addition, the i-Ready national norms for the 2020-2021 academic year should be updated and used to compare to the norms prior to the pandemic lockdowns and remote instruction.
A study of the correlation between absenteeism and reading performance during the pandemic might provide valuable information to educators. Another recommended study would be to examine the correlation between reading performance scores and varying lengths of online instructional periods for different types of students.

Future studies may lend insight to policymakers, legislators, higher education leaders, and school leaders alike. For example, Sato (2020) studied the educational response to the pandemic in Japan and revealed that “the underdevelopment of online learning [management] systems” was the primary negative policy issue during the onset of the pandemic (p. 64). Consequently, a comprehensive inventory of technical and curricular infrastructure and preparedness for remote education would provide valuable feedback to the stakeholders who help prepare schools, students, and teachers for unexpected emergencies similar to the COVID-19 pandemic. In addition, a school or district analysis of teachers who are adequately trained and sufficiently competent to deliver instruction remotely would help school leaders identify and support teachers as they continue to build their remote teaching skills.

Limitations

All students from kindergarten through eighth grade in the target school district took the i-Ready Diagnostic Reading Assessments; this research study focused only on the reading performance of a single cohort of third-grade students as they progressed through the fourth-grade and began the fifth grade. This exploratory study addressed only five specific i-Ready assessment periods: January 2020 (pre-COVID-19), August 2020 (when schools reopened), January 2021 (mid-year assessment), May 2021 (end of the year assessment), and August 2021 (new school year). In addition, the researcher focused on data from one school district.
The i-Ready program includes math diagnostic assessments; however, this study examined only reading assessment data. The i-Ready diagnostic assessment measures were designed to measure reading performance at different periods during the school year; as such, the instruments can be considered performance-based, formative assessments and should not be considered summative achievement tests (Curriculum Associates, 2019). Finally, exploratory studies are not designed to test specific hypotheses or to predict future development or performance among learners; as such, the results of this study merely describe educational performance during the periods under study and can be used to suggest future areas of research.

**Significance of the Study**

The purpose of this study was to explore the reading performance scores of elementary school students before and after school closures due to COVID-19 in one school district. Third-grade students’ reading performance was tracked over five different assessment periods and compared to the district and national norms. The results of the comparisons revealed small but statistically significant reading progress over time among the students in the research cohort. However, the progress of the research cohort was not commensurate with i-Ready’s national reading norms. This study adds to the body of knowledge related to the influence of school closures and remote instruction during school lockdowns and remote instruction due to COVID-19.
References


Using data from successive grades to estimate learning loss. RTI SharEd.
https://shared.rti.org/content/calculating-educational-impact-covid-19-part-ii-using-data-
successive-grades-estimate

http://www.setda.org/ls2013/wp-content/uploads/sites/8/2014/12/Adaptive-Diagnostic-
Science-SETDA.pdf

https://www.curriculumassociates.com/-/media/main/site/files/i-ready/research-overview-
proven-to-work-brochure-2019.pdf

https://www.curriculumassociates.com/-/media/main/site/files/corporate/state-of-student-
learning-2022.pdf

Curriculum Associates: Research and Efficacy.
https://www.curriculumassociates.com/research-and-efficacy/unfinished-learning-
research

Declining enrollment, shuttered schools: Puerto Rico’s education system in numbers. (2019,
December 4). Education Week. https://www.edweek.org/leadership/declining-enrollment-
shuttered-schools-puerto-ricos-education-system-in-numbers


https://www.census.gov/quickfacts/fact/table/marioncountyflorida/POP010220


https://firescholars.seu.edu/coe/113