ENGLISH-SPEAKING AND ENGLISH LEARNING STUDENTS' EARLY LITERACY PERFORMANCE BEFORE AND DURING THE COVID-19 PANDEMIC

by

KAITLYN M. ROY

A DISSERTATION

Presented to the Department of Special Education and Clinical Sciences
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2022

DISSERTATION APPROVAL PAGE

Student: Kaitlyn M. Roy

Title: English-Speaking and English Learning Students' Early Literacy Performance Before and During the COVID-19 Pandemic

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Special Education and Clinical Sciences by:

Gina Biancarosa	Chairperson
Hank Fien	Chairperson
Patrick Kennedy	Core Member
Geovanna Rodriguez	Core Member
Ilana Umansky	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded September 2022

DISSERTATION ABSTRACT

Kaitlyn Roy

Doctor of Philosophy

Department of Special Education and Clinical Sciences

September 2022

Title: English-Speaking and English Learning Students' Early Literacy Performance Before and During the COVID-19 Pandemic

Student reading skills observed during the COVID-19 pandemic are significantly lower than reading skills observed in same-grade students pre-pandemic (Domingue et al., 2021), and particularly for students from historically marginalized backgrounds (Amplify, 2021). Research examining COVID-related differences in reading trajectories and outcomes for students identified English learning (EL) students in particular, who have historically demonstrated heightened risk for performing well below their non-EL peers (NCES, 2019), is lacking. This study utilizes DIBELS 8 curriculum-based measurement (CBM) data from one cohort of students enrolled in a large, diverse, urban school district during the 2019-20 and 2020-21 school years to examine differences in early literacy skill development before and after onset of COVID-19 related disruptions to instruction. Results from this study demonstrate learning loss in decoding, blending, and sight word reading for all students from the benchmarking periods immediately before to those immediately after the onset of the COVID-19 pandemic. Relative to Non-EL students, ELs demonstrated less substantial learning loss in decoding and blending, but not sight word reading, followed by less accelerated growth in gaps between student groups relative to pre-pandemic trajectories. EL students with more advanced English language proficiency generally demonstrated higher levels of skills than their less-English proficient

peers, however patterns of growth trajectories varied widely across language proficiency groups.

CURRICULUM VITAE

NAME OF AUTHOR: Kaitlyn M. Roy

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Northern Arizona University, Flagstaff

DEGREES AWARDED:

Doctor of Philosophy, School Psychology, 2022, University of Oregon
Master of Science, Special Education, 2020, University of Oregon
Bachelor of Science, Psychological Sciences, 2017, Northern Arizona University

AREAS OF SPECIAL INTEREST:

Data-Based Decision-Making
Evidence-Based Behavioral and Academic Interventions
Equity in Educational Practices
Implementation Science

PROFESSIONAL EXPERIENCE:

School Psychologist, Durango School District 9-R, 2022-Present

School Psychologist Intern, Durango School District 9-R, 2021-2022

Graduate Teaching Assistant, College of Education, 2020-2021

Advanced Practicum Student, Center on Teaching and Learning, 2019-2021

Assessment Annotator, Analytic Measures Inc, 2020

Content Developer, National Center on Improving Literacy, 2019-2020

GRANTS, AWARDS, AND HONORS:

COMTSS State Personnel Development Grant, Durango School District 9-R, 2022-27

Doctoral Student Excellence Award, School Psychology Program, 2021

MTSS Early Career Scholar, I-MTSS Research Network, 2020

Alumni Scholarship Award, College of Education, 2017

ACKNOWLEDGMENTS

This work would not have been possible without my co-chairs, Gina Biancarosa and Hank Fien. Thank you both for always encouraging me to explore my interests and passions and for ensuring I had the opportunities to pursue them. No dream felt like it was too big when I talked with you two. I also owe special thanks to Patrick Kennedy, who spent countless hours on Zoom with me as I gained my bearings with a platform and a methodology that once felt incredibly foreign. I can't thank you and Gina enough for your unwavering support and willingness to meet me where I was, even when I felt so far away from where I wanted to be. Dr. G, thank you for being the teacher, mentor, and role model that I've needed. Your presence in our program, classrooms, and in my advising corner has been invaluable. Special thanks to Ilana Umansky for serving as the Institutional Representative on my committee and for pushing me to think outside of the box in my research. To have your knowledge, expertise, and perspective on my committee is a true honor.

I owe my deepest gratitude to my family and friends who have supported me throughout this journey. I wouldn't be here today if it weren't for your unconditional love, support, and belief in me.

To my current and future colleagues, students, and clients.

To my family and friends; this one’s for you.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

Early Literacy Skills 1

Importance of Reading Skills 1

Reading Skills Development for English Learning Students 2

Impact of School Closures on Reading Outcomes 4

COVID-19 Impacts on Public Education 5

COVID-19-Related Impacts on Reading Development 7

Impacts on Subgroups of Students 9

Research Gap 10

Current Study 11

Statement of Purpose 11

Research Questions 12

II. METHODS 14

Participants 14

Procedures 14

Measures 16

DIBELS 8 Measures 16

DIBELS 8 Nonsense Word Fluency 17

DIBELS 8 Word Reading Fluency 19

ACCESS For ELLs 20

English Language Indicator 21

Analyses ……. 21

Research Question 1 Analysis ……. 22

Research Question 2 Analysis ……. 25

Research Question 3 Analysis ……. 26

III. RESULTS ……. 29

Missing Data ……. 29

Descriptive Statistics ……. 32

Predictors ……. 32

English Learner (EL) Indicator ……. 32

English Language Proficiency …. ……. 32

DIBELS 8 Measures ……. 32

Research Question 1: All Students ……. 35

Decoding Skills ……. 35

Blending Skills ……. 35

Sight Word Reading Skills ……. 38

Research Question 2: EL Students ……. 39

Decoding Skills ……. 39

Blending Skills ……. 41

Sight Word Reading Skills ……. 42

Research Question 3: EL Students by English Proficiency Level ……. 43

Decoding Skills ……. 43

Blending Skills ……. 47

Sight Word Reading Skills ……. 49

IV. DISCUSSION …… 51

Summary of Results & Implications ……..…….…….. 51

Research Question 1: All Students ……..……...…. 51

Research Question 2: EL Students ……..……...…. 52

Research Question 3: EL Students by English Language Proficiency Level ....…. 53

High Performing ELs ……...…..…….……...……. 54

Variance Across Students and Schools ....…..……..…. 55

Study Limitations ……..……..…….…….. 57

Implications for Practice & Research …..……...…….…….. 58

APPENDICES ……. 61

A. OBSERVED PATTERNS OF MISSINGNESS FOR DIBELS 8 BENCHMARKING DATA ……..… 61

B. Q-Q PLOTS ILLUSTRATING SCORE DISTRIBUTIONS FOR ALL DIBELS 8 MEASURES EXAMINED ……....… 62

C. HISTOGRAMS FOR DIBELS 8 MEASURES ACROSS TIMEPOINTS …… 63

D. BOX PLOTS FOR ALL MEASURES BY ENGLISH LEARNER STATUS …… 72

E. RESIDUAL PLOTS FOR PIECEWISE HIERARCHICAL BASE MODELS …. 75

F. EL STUDENT BASE MODEL FIXED EFFECTS ……...… 78

REFERENCES CITED ……...… 79

LIST OF FIGURES

Figure Page

1. Estimated NWF-CLS Piecewise Growth for All Students …… 37

2. Estimated NWF-WRC Piecewise Growth for All Students …..…… 37

3. Estimated WRF Piecewise Growth for All Students ……..…… 38

4. Estimated NWF-CLS Piecewise Growth for EL and Non-EL Students …… 41

5. Estimated NWF-WRC Piecewise Growth for EL and Non-EL Students …..…… 42

6. Estimated WRF Piecewise Growth for EL and Non-EL Students …..……...…… 43

7. Estimated NWF-CLS Piecewise Growth for EL and Non-EL Students Across English Language Proficiency Levels …… 47

8. Estimated NWF-WRC Piecewise Growth for EL and Non-EL Students Across English Language Proficiency Levels …… 48

9. Estimated NWF-WRC Piecewise Growth for EL and Non-EL Students Across English Language Proficiency Levels …… 50

LIST OF TABLES

Tables Page

1. Sample Demographics ……..……..…… 15
2. Time-Varying Covariate Coding Schemes ……..……..……. 22
3. Random and Fixed Effects Included in Piecewise Growth Models …..…… 23
4. Piecewise Growth Models Investigated for Goodness of Fit …… 24
5. Frequency of DIBELS 8 Observations for All Students, ELs, and Non-ELs …… 29
6. Demographic Characteristics of Students Excluded from Study Sample …… 30
7. Frequency of Students with Complete DIBELS 8 Data Across Timepoints …… 31
8. EL Students' ACCESS for ELLs Overall Language Proficiency Levels ……..…… 33
9. Descriptive Statistics for DIBELS 8 Measure Outcomes …… 34
10. Research Question 1: Piecewise Growth Model Results for All Students Across All Measures ……..……...…… 36
11. Research Question 2: Piecewise Growth Model Variance Based on EL Status …… 40
12. Research Question 2: Piecewise Growth Model Fixed Effects Based on EL Status ……..……..……...……. 40
13. EL Student Base Model Random Effect Variance ……..…… 44
14. Research Question 3: Piecewise Growth Model Fixed Effects Based on EL Students' Overall Language Proficiency ……..……..……..…….. 45
15. Research Question 3: Piecewise Growth Model Random Effect Variance Based on EL Students' Overall Language Proficiency ……..…..….. 46

I. INTRODUCTION

Early Literacy Skills

Importance of Reading Skills

Reading is a fundamental academic skill taught as early as pre-kindergarten and continuing through high school. Reading ability in childhood and adolescence is associated with a variety of positive life outcomes including an individual's likelihood of graduating high school and/or attending college, as well as negative outcomes such as likelihood of becoming involved with the criminal justice system (Hernandez, 2011; Lesnick et al., 2010; McIntosh, Girvan, Horner, & Smolkowski, 2014). Unfortunately, many students across the United States struggle to perform at grade-level on standardized measures of reading ability. Data from the National Assessment of Education Progress (NAEP), for example, demonstrate that in 2019, a staggering 65% of fourth grade students performed below proficiency standards in reading, a significant decrease of 2 percentage points from just two years prior (National Center for Education Statistics [NCES], 2019). These data come not as a surprise, however, but rather a continuation of a trend demonstrating little change since 1992, nearly three decades earlier, when 71% of students performed below proficiency standards (NCES, 1992).

Reading experts suggest that early reading, or literacy, instruction should focus on phonemic awareness and the alphabetic principle, building a strong foundation for the development of more advanced skills in regular word decoding, sight word reading, and reading increasingly complex texts with accuracy, fluency, and comprehension (National Reading Panel, 2000). When these fundamental skills are lacking, future reading performance suffers. Research has shown that students' reading performance in early elementary school is strongly predictive of their future reading performance and growth (Duncan et al., 2007; Herbers et al., 2012). For

example, Spira, Bracken, and Fischel (2005) found that students' reading achievement at first and second grade is predictive of reading performance at the end of fourth grade. These findings support the argument that early literacy skills, such as phonemic awareness and the alphabetic principle, are crucial for future reading development.

Difficulties in reading also compound over time, resulting in widening gaps between students who are experiencing or are at risk for experiencing reading difficulty and their proficient peers (Francis et al., 1996). A recent study by Atteberry and McEachin (2021) examined student reading performance longitudinally from first through eighth grade and found that variation in reading achievement grew significantly as students progressed through these grades. Additionally, student characteristics, such as race, ethnicity, and socioeconomic status (SES), are associated with differential reading trajectories. That is, students from low SES and/or racial or ethnic minority backgrounds typically perform lower than their peers in reading and tend to fall further behind their peers over time (Phillip et al., 1998; Morgan et al., 2008).

Reading Skill Development for English Learning Students

As of 2018, approximately one in five individuals aged 5 years and older living in the United States spoke a language other than English in their home environment (U.S. Census Bureau, 2018). English learning (EL) students represent a quickly growing population within the public education system (Office of English Language Acquisition, 2021). Despite significant and rapid growth in this population over the past several decades, only minimal changes in the achievement gap between Non-EL and EL students have been observed.

In 2019, student's currently identified as EL scored, on average, 33 points (.74 standard deviations [SD]) lower than non-EL students on the NAEP assessment of reading for fourth grade (NCES, 2019). This achievement gap widens to differences of 45 (1.11 SD) and 53 points

(1.19 SD) for students in eighth and twelfth grades, respectively. Although these gaps have narrowed over time, the magnitude of these changes have been minimal (e.g., narrowing by 2 to 3 points across all grades over the past 10 years; NCES, 2020). Additionally, while 39% of non-EL students scored at or above NAEP's proficiency standards in the fourth grade, only 10% of EL students performed at or above this criterion (NCES, 2019). These proportions become even smaller at 4% and 3% for EL students in the eighth and twelfth grades, respectfully, indicating that less than one in twenty ELs scored at or above proficiency levels on this assessment. As the largest ongoing assessment of student skills across the United States, these reading data from NAEP indicate a clear need for education researchers to further investigate these gaps in reading achievement scores for EL students (Kieffer & Thompson, 2018).

Complicating matters, multilingual students seldom retain EL classification for the entirety of their education. Per federal law, education agencies are mandated to assess ELs annually in reading, writing, speaking, and listening to determine English proficiency development as well as eligibility for EL services (Every Student Succeeds Act, 2015). A subpopulation of multilingual students never qualify for EL services due to strong initial English proficiency and are thus referred to as Initial Fluent English Proficient (IFEP) rather than EL. Students who are at some point identified as EL but eventually reach adequate English language proficiency are exited from EL services and are subsequently referred to as "Former ELs". By nature of this reclassification, many students who were once classified as ELs may no longer be classified as such just four years later (Kieffer & Parker, 2016). In the context of NAEP assessment results, EL student scores in eighth grade are more likely to reflect eighth graders with newly developing English proficiency, rather than ELs who have been receiving English language services since a very young age.

When NAEP data was analyzed for current ELs, Former ELs, and IFEP students, Kieffer and Thompson (2018) found that multilingual students as a whole demonstrated faster growth over time in comparison to monolingual students. In addition, a longitudinal study that followed EL and non-EL students from kindergarten through eighth grade found that while EL students with limited English proficiency demonstrated far lower reading skills in kindergarten, these skill gaps were remediated by the end of eighth grade (Kieffer, 2011). Other studies have found that ELs demonstrated substantially slower reading growth rates in comparison to non-EL students, however much of these differences were explained by students' SES and/or classroom-level factors (Richardson, Rocconi, & Crewdson, 2020; Kieffer, 2008).

Curriculum-based measures (CBMs) of reading provide another method for examining differences in reading outcomes for EL and non-EL students. CBMs are brief and cost-effective assessments commonly used in schools to identify and monitor the progress of students with or at risk for academic difficulties. Primary schools are recommended to screen all students up to three times per year (Gersten et al., 2009), creating rich data sources for use in investigating patterns in student performance. Further, CBMs have been found, generally, to be valid and reliable indicators of student reading performance when compared to student performance on other measures of reading concurrently and predictively (January & Klingbeil, 2020). When used with multilingual students, CBMs maintain strong predictive and external validity (e.g., Fien et al., 2008; Vanderwood, Linklater, & Healy, 2008; Muskens, Betts, Lau, & Marston, 2009).

Impact of School Closures on Reading Outcomes

In addition to trends in student reading performance across multiple years, researchers have also studied within-year patterns in student rates of learning. While gaps in student reading

abilities generally grow wider over time, these gaps tend to become more pronounced over summer months, when students are not in school, than during the school year (Downey et al., 2004; von Hippel et al., 2018; Atteberry & McEachin, 2021). One recent study found that some students maintain the same rate of learning throughout the entire year, despite school closures during the summer, while others lose nearly one third of their school-year reading growth in the absence of formal instruction (Atteberry & McEachin, 2021). Although some students who lose their reading gains over the summer may demonstrate larger rates of growth upon returning to school in the fall, these gains are often not large enough to catch them up to their more proficient peers (Kuhfeld, Soland, et al., 2020).

These differential patterns of within-year learning rates are not random. As with multi-year trajectories of reading growth, student socioeconomic status (Downey et al., 2004) and race and ethnicity (von Hippel & Hamrock, 2019) are both associated with differential impacts of summer vacation on within-year rates of reading growth. Although there are some concerns in the field about the validity of the measures used in previous studies examining summer learning loss (see von Hippel et al., 2018), more recent studies (von Hippel & Hamrock, 2019; Atteberry & McEachin, 2021) demonstrate that even when more sound research methodology is utilized, summer learning loss occurs for many students, especially for students from historically marginalized backgrounds (i.e., low SES, students of color, etc.).

COVID-19 Impacts on Public Education

The quick and wide spread of the novel COVID-19 disease in the early months of 2020 led to abrupt changes in the daily lives of individuals across the United States, including those of at least 55.1 million U.S. children whose schools closed their doors in March due to the global pandemic (Education Week, 2020). Schools scrambled to piece together virtual

instruction delivery models to provide some amount of continued schooling and supports to their students. By early April 2020, 83% of parents surveyed by a Gallup poll indicated that their children were participating in an online learning program delivered by their school (Brenan, 2020), but many expressed concern about the quality of that instruction (Rapaport et al., 2020). In response to a survey of teacher perceptions of the impacts of COVID-19 on student learning, 22.5% of teachers reported that they were "Very Concerned" about the achievement of their students due to online learning, with an additional 31.1% of teachers reporting concern for their EL students' achievement in particular (Catalano, Torff, & Anderson, 2021). The impact of COVID-related school closures quickly became a clear concern to educators and families alike.

However, schools and students across the nation were not equally impacted by the COVID-19 pandemic. Data collected through the COVID-19 Educational Response Longitudinal Survey (C-ERLS) showed that, in May 2020, only 20% of included districts offered remote-learning opportunities that met the study's standards of rigorous instructional practices (i.e., offered at least partially synchronous instruction, utilized a mechanism for encouraging student engagement such as grades and/or attendance, and engaged in direct correspondence between teachers and students; Malkus, 2020). Further, schools that were considered low-poverty and/or higher-achieving were more likely to meet the study's standards, highlighting district or school-level differences in instructional practices that may contribute to post-pandemic achievement gaps for students.

Maintaining student engagement in a virtual environment also posed unique challenges. In one survey, teachers reported that approximately one in five students were unreachable during remote instruction in Spring of 2020 (Sawchuk & Samuels, 2020). Another survey found

that only 60% of students were regularly engaged in remote learning activities, with notable differences in engagement reported by teachers who served high proportions of students from low-income or racial minority backgrounds (Kraft & Simon, 2020). This variability in remote instruction delivery methods and student engagement with online learning raises concerns about the quality of instruction students received during the pandemic and the impact that COVID-19-related school closures have had on student learning outcomes as a result.

COVID-19-Related Impacts on Reading Development

As the COVID-19 pandemic continued through the entirety of the 2020-21 school year, researchers have worked to predict the impact of COVID-19-related closures on student learning outcomes, with mixed results. For instance, Kuhfeld, Soland, and colleagues (2020) utilized data from studies examining seasonal learning, weather-related closures, and student absenteeism to predict that students would begin the 2020-21 school year with only 63-68% of the learning gains in reading typically demonstrated under normal conditions. Notably, these data assume a complete absence of instruction (i.e., predictions based on data from summer vacation, snow days, student absence from school), so study authors cautioned that these estimates may represent lower bounds of expected growth due to many schools' use of virtual or hybrid instruction models. Similarly, a study conducted by Bao and colleagues (2020), predicted that from the beginning of 2020 through September 2020, kindergarten students would experience approximately 31% less reading gains than those demonstrated under more typical instructional circumstances.

Similarly, a recent meta-analysis of studies examining COVID-19-related impacts on educational outcomes found that students in primary education were more severely impacted by the COVID-19 pandemic than were students in secondary education (König & Frey, 2022). One

study examining reading achievement for elementary-aged students pre- and -post pandemic found that students demonstrated up to 3-4 months of learning loss in Fall 2020, with the largest impact observed in late elementary grades (Bielinski, Brown, & Wagner, 2021). Kuhfeld, Tarasawa, and colleagues (2020) utilized MAP Growth (NWEA, 2019) assessment data for grades 3-8 to examine student reading scores in Fall 2020 compared to scores collected in the previous academic year. Unlike other students, the authors found that student reading scores on the MAP Growth assessment in fall 2020 were similar to those of same-grade students who completed the assessment in fall 2019. Analyses also found that for all but two grades, on average, students made academic gains in reading that were similar to gains made by students in the previous year. Another study using the same sample examined attrition of students from fall 2019 to fall 2020 and found that most students who did not complete the MAPS assessment in fall 2020 were members of ethnic/racial minority groups, demonstrated lower achievement in fall 2019, and/or attended schools with higher proportions of low SES students (Johnson & Kuhfeld, 2020).

Although Kuhfeld, Tarasawa, and colleagues' (2020) findings suggest that students did not demonstrate reading learning loss as a result of the COVID-19 pandemic, bias in their study sample may be masking potential learning loss of students from backgrounds that historically perform lower in reading (Johnson & Kuhfeld, 2020). For instance, Domingue and colleagues (2021) found that for students in the second and third grade, growth in Oral Reading Fluency stalled in Spring 2020 and then re-accelerated in Fall 2020. Despite this reacceleration, students performed about 30% lower than typical performance for their respective grade levels pre-pandemic. Similarly, a recent study (Furjanic, Ives, Fainstein, Kennedy, & Biancarosa, revision in progress) found that, for a national sample of students in grades one through five,

students' average Oral Reading Fluency (ORF) scores in fall 2020 were lower in comparison to students in previous cohorts and even more discrepant from previous cohorts' scores in winter 2021. Another study conducted by Amplify Education, Inc (2021a) showed that observed COVID-related reading losses during the 2020-21 year were most substantial for students in grades K and 1, and especially for students who are Black or Hispanic.

Impacts on Subgroups of Students

Researchers have attempted to estimate the impacts of COVID-19 on pre-existing achievement gaps. Bailey and colleagues (2020) surveyed over 200 education researchers across the United States to collect their forecasts of the impact of COVID-19-related closures on student achievement gaps. Compared to reading achievement gaps pre-pandemic, median expert forecasts estimated that gaps between low- and high-SES students would be 0.25 standard deviations wider in Spring 2021 and 0.20 standard deviations wider in Spring 2022. These estimates suggest that experts anticipated that achievement gaps would be smaller in Spring 2022 than in Spring 2021, but only by a very marginal difference of 0.05 standard deviations. Expert forecasts were highly variable, however, and one quarter of respondents reported forecasts that were larger in Spring 2022 than Spring 2021, suggesting that some researchers anticipated that COVID-19 related changes to these achievement gaps may widen over time. Results from this survey demonstrate that education researchers expected COVID-19 to disproportionally impact students from low-SES backgrounds, those students who are already at heightened risk for experiencing academic difficulty.

Kuhfeld, Soland, and colleagues' (2020) predictions, discussed in more detail above, also suggest that lower-achieving students were expected to have experienced larger losses in reading relative to their higher-achieving peers, some of whom may experience reading gains

despite COVID-19 changes to schooling. In a study conducted by Domingue and colleagues (2021), also discussed above, students attending higher achieving schools were observed to demonstrate faster rates of growth than students attending lower achieving schools. A study conducted by Pier and colleagues (2021) using curriculum-based measures found that students from low-SES backgrounds, English learners, and Latinx students experienced greater learning lag than their peers during the 2020-21 school year.

Research Gap

Although research on the observed impact of COVID-19 has progressed quickly, research has not yet examined the impacts on academic achievement for EL students in particular. More specifically, although researchers are beginning to examine the impacts of COVID-19 on reading development for subgroups of students (i.e., students from low-income and/or racial minority backgrounds), research on the development of specific reading skills is lacking for this group of students. Additionally, EL students' English proficiency has yet to be examined within the context of COVID-19 disruptions to education to determine potential differential impacts on the COVID-19 pandemic for subgroups of ELs, such as those whose English language proficiency was limited at the time of school entry.

COVID-19 has also impacted schools' ability to assess students. By Spring 2021, at least 14 states required schools to offer in-person learning opportunities to some or all students, while 36 states required either partial closures or left decision-making power to individual school districts (Education Week, 2021). Thus, many students in the United States attended school at least partially in-person, changes to educational environments and limitations to standardized testing procedures due to the pandemic may have limited opportunities to empirically examine the impacts of COVID-19 on student outcomes and accurately and reliably

estimate current levels of student performance. Nevertheless, some school districts have continued to collect CBM benchmarking data for all students as part of their typical education practices, providing a unique opportunity for researchers and education stakeholders to examine the impact of the ongoing COVID-19 pandemic on student outcomes.

Current Study

Statement of Purpose

This study examines the impacts of COVID-19 related school closures on the reading outcomes of students attending a large, urban school district as measured by CBM benchmarking data collected as part of typical education practices. Specifically, the study utilized student performance data from two DIBELS 8 (University of Oregon, 2020a) subtests: Nonsense Word Fluency (NWF), a measure of student decoding abilities that provides measures of correct letter sounds and words recoded correctly; and Word Reading Fluency (WRF), a measure of regular and irregular word reading that provides one score representing the number of words read correctly. All data were drawn from benchmarking assessments administered by school district personnel. Relationships between student language proficiency and reading outcomes before and after the onset of COVID-19 disruptions were examined using district-provided indicators of EL or non-EL status, as well as individual EL student scores on the WIDA ACCESS for ELLs test of language proficiency from kindergarten.

This study responds to the need for empirical analyses of the impacts of the COVID-19 related changes to education on student academic outcomes, particularly in reading (Bailey et al., 2021). Additionally, this study aimed to answer important questions about potentially differential impacts for EL students, who, as a group, performed much poorer than their non-EL peers in reading prior to the pandemic (e.g., NCES, 2020). By utilizing data collected as part of

typical educational practices, this study reports empirical findings that education stakeholders can incorporate during decision-making processes.

Research Questions

Using a sample of students who were in first and second grade during the 2019-20 and 2020-21 school years, respectively, this study answered three research questions:

1. Are there observed differences in student decoding, blending, and sight word reading abilities and growth before and after the onset of the COVID-19 pandemic? Consistent with researcher predictions and forecasts (Bailey et al., 2021; Kuhfeld, Soland, et al., 2020) as well as preliminary studies examining COVID-19 impacts on reading outcomes (Kuhfeld, Tarasawa, et al., 2020), I hypothesized that there would be significant impacts of the COVID-19 pandemic on second grade student reading abilities across decoding, blending, and sight word reading abilities.

2. Is there a relationship between: a) any observed differences in decoding, blending, and sight word reading before and after the onset of the COVID-19 pandemic; and b) student status as an English learning (EL) student? Given evidence demonstrating large achievement gaps between EL and non-EL students in reading (e.g., NCES, 2019), I hypothesized that there would be significant differences in the impact of the COVID-19 pandemic on second grade reading outcomes across all reading domains examined (decoding, blending, and sight word reading).

3. Is there a relationship between: a) any observed differences in EL student decoding, blending, and sight word reading before and after the onset of the COVID-19 pandemic; and b) EL students' language proficiency level? I hypothesized that student language proficiency levels would be associated with differential impacts of the COVID-19

pandemic on each domain of reading examined (decoding, blending, and sight word reading). To progress through English language proficiency levels, students must demonstrate development in skills in listening, speaking, reading, and writing in English. Therefore, students who demonstrate higher levels of English language proficiency on a standardized assessment of language may be more likely to perform better on measures of reading ability due to the inherent similarities in reading assessment tasks.

II. METHODS

Participants

This study used extant early literacy benchmarking and language proficiency data from a large, urban school district collected during the 2019-20 and 2020-21 school years. Students who were in first grade during the 2019-20 school year (i.e., before the onset of the COVID-19 pandemic) and were promoted to second grade during the 2020-21 school year (i.e., after the onset of the COVID-19 pandemic) were included in this study's sample. The sample was limited to students who had early literacy data from at least one benchmarking period in each school year and who had complete literacy assessment data from each of the benchmarking periods in which they participated. Applying these inclusionary criteria, the sample was narrowed from 5,362 to 4,149 students from 93 schools. Of the students excluded, the majority were missing data from the 2019-20 school year (see Results and Appendix A for more details about the students excluded from this sample). Student demographic information, including race/ethnicity, placement in special education, English learner status, initial English language proficiency (i.e., WIDA ACCESS data), and school of enrollment were provided by the district as of the beginning of the 2019-20 school year and consequently do not capture potential changes in these characteristics across time. See Table 1 for a summary of demographic characteristics for students included in this study.

Procedures

The participating district regularly conducted benchmarking assessments for all students during the Fall, Winter, and Spring benchmarking periods. Student performance on the DIBELS 8 Nonsense Word Fluency (NWF) and Word Reading Fluency (WRF) measures (see the Measures section for a detailed description of these assessments) in Fall and Winter of

Table 1

Sample Demographics

	All Students (n = 4,149)		Non-ELs (n = 3,541)		ELs (n = 608)	
Variable	*n*	%	*n*	%	*n*	%
English Language Proficiency						
Non-English Learner (Non-EL)	3,541	85.3	–	–	–	–
Never English Learner	3,520	84.8	3,520	99.4	–	–
Exited English Learner	21	0.5	21	0.6	–	–
English Learner (EL)	608	14.7	–	–	–	–
Gender						
Male	2,120	51.1	1,805	51.0	315	51.8
Female	2,029	48.9	1,736	49.0	293	48.1
Race/Ethnicity						
Black/African American	3,057	73.7	3,029	85.5	28	4.6
Hispanic/Latino	723	17.4	174	4.9	549	90.3
White	251	6.0	238	6.7	13	2.1
Multiracial	78	1.9	†	–	†	–
Asian	25	0.6	†	–	†	–
Native Hawaiian/Pacific Islander	10	0.2	†	–	†	–
American Indian/Alaska Native	†	–	†	–	†	–
Special Education Status						
None	3,584	86.4	3,017	85.2	567	93.3
Special Education	446	10.7	415	11.7	31	5.1
Exited Special Education	63	1.5	†	–	†	–
504 Plan	54	1.3	†	–	†	–
Exited 504 Plan	†	–	†	–	†	

Note. ELs = English Learners. Non-ELs = Non-English Learners. Never English Learner refers to students who have never been classified as an English Learner. Exited English Learner refers to students who were once classified as English Learners but had since discontinued receiving English language services.

† data suppressed due to subgroup $n < 10$

school year 2019-20 and Fall, Winter, and Spring of 2020-21 were included in this study's analyses. Due to the onset of the COVID-19 pandemic in March 2020 and subsequent school closures, data from the Spring 2020 benchmarking period were not collected by the district.

Measures

DIBELS 8 Measures

The Dynamic Indicators of Basic Early Literacy Skills, Eighth Edition (DIBELS 8; University of Oregon, 2020a) is a battery of brief measures of early literacy skills for students in kindergarten through eighth grade. DIBELS 8 assessments are used by schools across the United States to screen students for reading difficulty, conduct benchmark assessments of student skills across multiple time periods, and monitor the progress of students' early literacy skill development. Students within the participating district completed DIBELS 8 assessments using Amplify Education, Inc's mCLASS system, a computer-administered version of DIBELS 8 measures.

Cut scores are established for each measure at each benchmarking period to reflect students' likelihood of demonstrating grade-level proficiency at the end of the school year (University of Oregon, 2020a). Scores that fall within the *At Risk* or *Intensive Support* range indicate performance that is well below the benchmark and demonstrates high risk for reading difficulty. Scores within the *Some Risk* or *Strategic Support* range indicate performance that is below benchmark and demonstrates some risk for reading difficulty. Scores within the *Minimal* or *Negligible* risk (i.e., *Core Support*) range indicates performance that is above or well above the benchmark and demonstrates high likelihood of continuing to achieve at or above benchmark through the end of the academic year.

This study examines student performance on the Nonsense Word Fluency (NWF) and Word Reading Fluency (WRF) measures during their first and second grade years. Beginning and end of year correlations between NWF and WRF scores across first and second grades are strong (.75-.84), indicating that the two assessments measure distinct, but interrelated, skills (University of Oregon, 2020b).

DIBELS 8 Nonsense Word Fluency. Nonsense Word Fluency (NWF) is a one-minute timed, standardized, individually administered measure of student knowledge and application of the alphabetic principle (Kaminski & Good, 1996). When administered NWF, students are presented with a list of decodable pseudowords and instructed to either produce the individual sounds within each word or to read the whole word. NWF yields two scores: a Correct Letter Sounds (CLS) score, which represents the number of correct letter sounds the student produced within one minute; and a Words Read Correctly (WRC) score, which represents the number of words that were blended and read correctly. Together, these two scores capture student skills in both letter-sound correspondence (i.e., CLS score) and blending (i.e., WRC score), which requires advanced skills in letter-sound correspondence (National Reading Panel, 2001).

Reliability and validity coefficients are reported separately for NWF-CLS and NWF-WRC scores. Median test-retest reliability coefficients for NWF-CLS in first and second grade were .81 and .75, respectively (University of Oregon, 2020b). NWF-CLS alternate form reliability coefficients were similar for first (.79) and second grades (.83; University of Oregon, 2020b), suggesting that reliability of this measure is strong over time and when alternate forms are utilized. Across benchmarking periods, the concurrent criterion validity correlations of first grade NWF-CLS ranged from .71 to .85 with the DIBELS Next Composite, .57 to .65 for the IOWA Total Reading, and .34 to .54 for the IOWA Word Analysis (University of Oregon,

2020b). Second grade concurrent criterion validity of NWF-CLS ranged from .44 to .80 for DIBELS Next Composite, .62 to .68 for IOWA Total Reading, and .32 to .51 with IOWA Word Analysis (University of Oregon, 2020b). Predictive validity coefficients comparing Fall and Winter NWF-CLS to end of year performance on criterion measures ranged from .55 to .76 for DIBELS Next Composite, .50 to .66 for IOWA Total Reading, and .47 to .70 for IOWA Word Analysis in first and second grade (University of Oregon, 2020b). NWF-CLS from DIBELS 8 appears to be more strongly correlated with reading measures administered at a later time rather than concurrently.

Median test-retest coefficients for NWF-WRC are .78 for first grade, and .72 for second grade (University of Oregon, 2020b). Alternate form reliability for NWF-WRC in first and second grade were .72 and .81, respectively (University of Oregon, 2020b). Concurrent criterion validity of NWF-WRC in first grade ranged from .61 to .86 for DIBELS Next Composite, .52 to .63 for IOWA Total Reading, and .26 to .56 for IOWA Word Analysis (University of Oregon, 2020b). For second grade, concurrent criterion validity of WRF ranged from .48 to .79 relative to the DIBELS Next Composite, .60 to .70 for IOWA Total Reading, and .62 to .73 for IOWA Word Analysis (University of Oregon, 2020b). Predictive validity coefficients comparing Fall and Winter NWF-WRC to end of year performance on criterion measures in first and second grade ranged from .59 to .75 for the DIBELS Next Composite, .46 to .67 for IOWA Total Reading, and .40 to .70 for IOWA Word Analysis (University of Oregon, 2020b). Overall, reliability and validity coefficients across NWF-CLS and -WRC are largely similar and demonstrate that this measure has strong utility relative to other measures of reading ability and remains psychometrically sound when used repeatedly, as designed.

DIBELS 8 Word Reading Fluency. Word Reading Fluency (WRF) is an individually administered standardized assessment of students' skills with the alphabetic principle and fluency with reading real words. When administered this assessment, students are presented with a list of grade-appropriate words and asked to read them aloud for one minute. The student's score represents the number of words read correctly within that minute, with no penalty for words read incorrectly.

Median test-retest reliability coefficients for DIBELS 8 WRF are .92 and .95 for first and second grade, respectively (University of Oregon, 2020b). Additionally, alternate form reliability coefficients for WRF benchmarking probes demonstrate strong reliability in first (.96) and second grade (.97; University of Oregon, 2020b). Concurrent validity correlations for first grade WRF ranged from .69 to .88 in relation to the DIBELS Next Composite, .51 to .79 with the IOWA Total Reading, and .67 to .84 with the IOWA Word Analysis (University of Oregon, 2020b). Second grade WRF concurrent validity ranged from .77 to .91 with the DIBELS Next Composite, .62 to .83 with the IOWA Total Reading, and .60 to .82 with the IOWA Word Analysis. Predictive validity coefficients comparing first and second grade WRF in Fall and Winter to end of year performance on criterion measures ranged from .72 to .85 for the DIBELS Next Composite, .58 to .78 for the IOWA Total Reading, and .41 to .78 with the IOWA Word Analysis (University of Oregon, 2020b). Overall, reliability and validity coefficients for WRF indicate that, similar to NWF-CLS and -WRC, WRF has strong assessment utility for capturing students' reading ability relative to other measures of reading skills and may be administered repeatedly.

ACCESS for ELLs

The Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS for ELLs) assessment is a standardized assessment of student English language proficiency developed by the WIDA Consortium. The WIDA Consortium includes 40 states in the U.S. and in the 2018-19 school year, 1.8 million EL students completed the ACCESS for ELLs assessment (Wisconsin Center for Education Research, n.d.). ACCESS for ELLs assesses students' English proficiency in four domains: listening, speaking, reading, and writing. Student performance is reported using both scale scores and proficiency levels for each individual domain as well as an Overall Proficiency composite. Domain and composite scores range from 100 to 600 and levels range from 1.0 to 6.0, where larger numbers represent higher levels of proficiency. There are six Overall Proficiency levels: Entering (1.0-1.9), Emerging (2.0-2.9), Developing (3.0-3.9), Expanding (4.0-4.9), Bridging (5.0-5.9), and Reaching (6.0). The overall proficiency composite for kindergarten places more weight on the reading and writing domains (i.e., each makes up 35% of the composite score/level) than on the listening and speaking domains (i.e., each makes up 15% of the composite score/level; Porter, Cook, & Sahakyan, 2019). Reliability of the kindergarten Overall Proficiency score was .97 (WIDA, 2020). Additionally, classification accuracy of the Overall Proficiency score for grades one and two was .82 and .83, respectively (WIDA, 2020), indicating that the ACCESS ELLs Overall Proficiency score is largely accurate at classifying an individual's language proficiency level.

Federal law requires the use of an English proficiency assessment, such as ACCESS for ELLs, for monitoring student progress in developing English language proficiency and determining continued access to school-based English language support services (Ragan &

Lesaux, 2006). This study's participating school district administers the ACCESS for ELLs assessment to all EL students every spring and requires that students achieve an overall proficiency level of 4.5 or greater to exit English language support services. Given the utility of the ACCESS for ELLs overall proficiency level for decision-making in schools, this study utilizes the overall proficiency level as an indicator of EL students' English language proficiency in analyses. More specifically, this study utilizes language proficiency data from students' kindergarten year (i.e., the 2018-19 school year) as a baseline measure of language proficiency.

English Language Indicator

Students in this sample were identified by the district as either English learners (i.e., EL), not English learners (i.e., Non-EL), or Exited English Learner. To be classified as an EL in this district, the student must have both 1) been identified by their family as speaking a non-English language in the home, and 2) demonstrate an English language proficiency score below the cutoff (i.e., 4.5 for ACCESS for ELLs). Students who were at one point identified as an EL but no longer participate in English language services were classified as exited English learners, referred to as Former-ELs in this study.

Analyses

To answer the proposed research questions, a series of three-level piecewise growth models were fit for each DIBELS 8 measure of interest (i.e., NWF-CLS, NWF-WRC, and WRF) using the following three hierarchical levels: Time (level 1), Students (level 2), and Schools (level 3). Models were fit using the lme4 package within the R platform (Bates et al., 2015; R Core Team, 2021). Student growth across time was modeled using five time-varying covariates, summarized in Table 2 (McCoach & Kaniskan, 2010). Each time-varying covariate

provides a unique coding scheme for time, allowing for discontinuity in growth rates across time within the model. Five unique models with various combinations of time-varying covariates and random effects were run for each DIBELS 8 outcome measure and compared for goodness of fit using the AIC, BIC, log likelihood, and Deviance. Table 3 provides an overview of each model investigated. As shown in Table 4, Model 1A was the model of best fit for all DIBELS 8 outcome measures and thus was utilized for all multilevel analyses conducted.

Table 2

Time-Varying Covariate Coding Schemes

Time	Time Linear	Time Instruction	Alt. Time Instruction	COVID Summer	COVID Instruction	Spring 2021
BOY 2019-20	0	0	0	0	0	0
MOY 2019-20	1	1	1	0	0	0
BOY 2019-20	2	1	1	1	0	0
MOY 2020-21	3	2	2	1	1	0
EOY 2020-21	4	3	2	1	1	1

Note. BOY = Beginning of Year. MOY = Middle of Year. EOY = End of Year.

Research Question 1 Analysis. To explore potential differences in the early literacy skills (i.e., decoding, blending, and sight word reading) of all students before and after the onset of the COVID-19 pandemic, piecewise hierarchical growth models were run for each DIBELS 8 outcome measure of interest (i.e., NWF-CLS, NWF-WRC, and WRF) using Model 1A. This model is represented by the following equation:

$$Y_{tij} = \beta_{000} + \beta_{100}(Time\_Instructional) + \beta_{200}(COVID\_Summer) + \beta_{300}(COVID\_Instruction) + \beta_{400}(Spring21) + u_{00j} + u_{10j} + u_{20j} + u_{30j} + r_{0ij} + r_{1ij} + r_{2ij} + r_{3ij} + e_{tij}$$

In this equation, Y_{tij} represents the score for student *i* in school *j* at time *t*. β_{000} is the intercept and represents the average score for the average student in the average school at the beginning

Table 3

Random and Fixed Effects Included in Piecewise Growth Models

	Random Effects					Fixed Effects				
Model	Time Instruction	Alt. Time Instruction	COVID Summer	COVID Instruction	Spring 21	Time Instruction	Alt. Time Instruction	COVID Summer	COVID Instruction	Spring 21
Model 1A	X		X	X		X		X	X	X
Model 1B	X			X	X	X		X	X	X
Model 1C	X		X		X	X		X	X	X
Model 2A		X	X	X			X	X	X	X
Model 2B		X		X	X		X	X	X	X

Table 4

Piecewise Growth Models Investigated for Goodness of Fit

Model by Measure	npar	AIC	BIC	logLik	Deviance
NWF-CLS					
Model 1A	26	162,751	162,955	-81,349	162,699
Model 1B	26	163,281	163,485	-81,615	163,229
Model 1C	26	163,024	163,228	-81,486	162,972
Model 2A	26	163,961	164,165	-81,954	163,909
Model 2B	26	164,129	164,333	-82,039	16,4077
NWF-WRC					
Model 1A	26	126,375	126,579	-63,162	126,323
Model 1B	26	126,894	127,097	-63,421	126,842
Model 1C	26	126,598	126,801	-63,273	126,546
Model 2A	26	127,170	127,374	-63,559	127,118
Model 2B	26	127,746	127,950	-63,847	127,694
WRF					
Model 1A	26	138,127	138,331	-69,038	138,075
Model 1B	26	139,100	139,303	-69,524	139,048
Model 1C	26	138,306	138,510	-69,127	138,254
Model 2A	26	138,392	138,596	-69,170	138,340
Model 2B	26	139,517	139,720	-69,732	139,465

Note. npar = nonparametric test. AIC = Akaike information criterion. BIC = Bayesian information criterion. logLik = loglikelihood.

of the 2019-20 school year. β_{100} is the academic year slope and represents the average change in scores between each benchmarking occasion during the academic year for the average student in the average school. β_{200} is the change in scores associated with school closures due to the COVID-19 pandemic as well as summer effects for the average student in the average school. β_{300} is the difference in growth observed from the beginning to middle of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average student in the average school. β_{400} is the difference in growth from the middle to end of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average student in the average school. u_{00j} represents the variance between schools around β_{000}. u_{10j}

represents the variance between schools around β_{100}. u_{20j} represents the variance between schools around β_{200}. u_{30j} represents the variance between schools around β_{300}. r_{0ij} represents the variance between students within schools around β_{000}. r_{1ij} represents the variance between schools around β_{100}. r_{2ij} represents the variance between schools around β_{200}. r_{3ij} represents the variance between schools around β_{300}. Lastly, e_{tij} represents the residual variance within the student across time.

Research Question 2 Analysis. To explore the relationship between student EL status (i.e., EL or Non-EL) and potential differences in observed early literacy skills (i.e., decoding, blending, and sight word reading) before and after the onset of the COVID-19 pandemic, students' indicator of EL status (i.e., either EL or Non-EL, which included both Never-EL and Former-EL students) at the beginning of the 2020-21 school year was entered into Model 1A as a student-level (Level 2) predictor as a moderator of each time-varying covariate. The mixed-model equation for Research Question 2 is represented in equation 2 below:

$$\begin{aligned} Y_{tij} = {} & \beta_{000} + \beta_{010}(EL) + \beta_{100}(Time\_Instruction) + \beta_{110}(Time\_Instruction * EL) \\ & + \beta_{200}(COVID\_Summer) + \beta_{210}(COVID\_Summer * EL) \\ & + \beta_{300}(COVID\_Instruction) + \beta_{310}(COVID\_Instruction * EL) \\ & + \beta_{400}(Spring21) + \beta_{410}(Spring21 * EL) + u_{00j} + u_{10j} + u_{20j} + u_{30j} \\ & + r_{0ij} + r_{1ij} + r_{2ij} + r_{3ij} + e_{tij} \end{aligned}$$

In this equation, β_{000} is the intercept and represents the average score for the average student who is not an EL in the average school. β_{010} represents the average difference in the intercept associated with student EL status. β_{100} is the academic year slope and represents the average change in scores between each benchmarking occasion during the academic year for the average non-EL student in the average school. β_{110} represents the average change in scores between

each benchmarking occasion during the academic year for the average EL student in the average school. β_{200} is the change in scores associated with school closures due to the COVID-19 pandemic as well as summer effects for the average non-EL student in the average school. β_{210} is the change in scores associated with school closures due to the COVID-19 pandemic as well as summer effects for the average EL student in the average school. β_{300} is the difference in growth observed from the beginning to middle of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average non-EL student in the average school. β_{310} is the difference in growth observed from the beginning to middle of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average EL student in the average school. β_{400} is the difference in growth from the middle to end of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average non-EL student in the average school. β_{410} is the difference in growth from the middle to end of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average EL student in the average school.

Similar to equation 1, u_{00j} in this equation represents the variance between schools around β_{000}. u_{10j} represents the variance between schools around β_{100}. u_{20j} represents the variance between schools around β_{200}. u_{30j} represents the variance between schools around β_{300} . r_{0ij} represents the variance between students within schools around β_{000}. r_{1ij} represents the variance between schools around β_{100}. r_{2ij} represents the variance between schools around β_{200}. r_{3ij} represents the variance between schools around β_{300}. Lastly, e_{tij} represents the residual variance within the student across time.

Research Question 3 Analysis. To explore the relationship between EL students' English language proficiency level and potential differences in observed early literacy skills (i.e., decoding, blending, and sight word reading) before and after the onset of the COVID-19 pandemic, students' English language proficiency level from kindergarten was entered into Model 1A as a student-level (Level 2) predictor. Additionally, this model included only students who were identified as ELs at the beginning of the 2020-21 school year (n = 608). The mixed-model equation for research question 3 is represented in equation 3 below:

$$
\begin{aligned}
Y_{tij} = {} & \beta_{000} + \beta_{010}Emerging + \beta_{020}Developing + \beta_{030}Expanding \\
& + \beta_{100}(Time\_Instruction) + \beta_{110}(Time\_Instruction * Emerging) \\
& + \beta_{120}(Time\_Instruction * Developing) \\
& + \beta_{130}(Time\_Instruction * Expanding) + \beta_{200}(COVID\_Summer) \\
& + \beta_{210}(COVID\_Summer * Emerging) \\
& + \beta_{220}(COVID\_Summer * Developing) \\
& + \beta_{230}(COVID\_Summer * Expanding) + \beta_{300}(COVID\_Instruction) \\
& + \beta_{310}(COVID\_Instruction * Emerging) + \beta_{320}(COVID\_Instruction \\
& * Developing) + \beta_{330}(COVID\_Instruction * Expanding) \\
& + \beta_{400}(Spring21) + \beta_{410}(Spring21 * Emerging) + \beta_{420}(Spring21 \\
& * Developing) + \beta_{430}(Spring21 * Expanding) + u_{00j} + u_{10j} + u_{20j} + u_{30j} \\
& + r_{0ij} + r_{1ij} + r_{2ij} + r_{3ij} + e_{tij}
\end{aligned}
$$

In this equation, β_{000} is the intercept and represents the average score for the average EL with Entering language proficiency in the average school. β_{010}, β_{020}, and β_{030} represent the average difference in the intercept associated with the Emerging, Developing, and Expanding language proficiency levels, respectively. β_{100}, β_{110}, β_{120}, and β_{130} represent the academic year slope or the average change in scores between each benchmarking occasion during the academic year for the average EL student with Entering, Emerging, Developing, and Expanding language proficiency in the average school. β_{200} is the change in scores associated with school closures due to the COVID-19 pandemic as well as summer effects for the average EL student with Entering proficiency in the average school. β_{210}, β_{220}, and β_{230} represent the change in scores

associated with school closures due to the COVID-19 pandemic as well as summer effects for the average EL student with Emerging, Developing, and Expanding language proficiency in the average school. β_{300} is the difference in growth observed from the beginning to middle of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average EL student with Entering language proficiency in the average school. β_{310}, β_{320}, and β_{330} represent the difference in growth observed from the beginning to middle of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average EL student with Emerging, Developing, and Expanding language proficiency in the average school. β_{400} is the difference in growth from the middle to end of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average EL student with Entering proficiency in the average school. β_{410} is the difference in growth from the middle to end of the 2020-21 school year compared to the beginning to middle of the 2019-20 school year for the average EL student with Emerging, Developing, and Expanding language proficiency in the average school.

In this equation, u_{00j} represents the variance between schools around β_{000}. u_{10j} represents the variance between schools around β_{100}. u_{20j} represents the variance between schools around β_{200}. u_{30j} represents the variance between schools around β_{300}. r_{0ij} represents the variance between students within schools around β_{000}. r_{1ij} represents the variance between schools around β_{100}. r_{2ij} represents the variance between schools around β_{200}. r_{3ij} represents the variance between schools around β_{300}. Lastly, e_{tij} represents the residual variance within the student across time.

III. RESULTS

Missing Data

Table 6 summarizes the demographic characteristics of students who did not meet the study's inclusionary criteria. Of the 1,212 students excluded from the study sample, 910 students had benchmark assessment data from the 2019-20 school year only, 280 had benchmark assessment data from the 2020-21 school year only, 4 had only partial benchmark assessment data during at least one benchmarking period, and 18 students had no benchmark assessment data from any benchmarking period. Nearly two thirds of students in the study sample had complete DIBELS 8 data across all five benchmarking periods of interest. On average, students in this sample participated in 4.49 of the five benchmarking periods examined (*SD* = 0.77; Table 5). The number of students with data at each time point is shown in Table 7, and Appendix A provides a comprehensive overview of patterns of student participation across benchmarking periods.

Table 5

Frequency of DIBELS 8 Observations for All Students, ELs, and Non-ELs

	Group Frequency		Students Per Observation Frequency			
Students	*n*	*M* (*SD*)	Two	Three	Four	Five
All Students	4,149	4.49 (0.77)	54 (1.3)	535 (12.9)	877 (21.1)	2,683 (64.7)
Non-ELs	3,541	4.47 (0.77)	43 (1.2)	487 (13.8)	780 (22.0)	2,231 (63.0)
ELs	608	4.63 (0.71)	11 (1.8)	48 (7.9)	97 (16.0)	452 (74.3)

Note. EL = English Learner. Non-EL = Non-English Learner. Never-EL = Never English Learner. Exited EL = Exited English Learner.

Table 6

Demographic Characteristics of Students Excluded from Study Sample

Demographic Variable	*n*	%
English Language Proficiency		
Non-English Learner (Non-EL)	892	73.6
Never English Learner	882	72.8
Exited English Learner	10	0.8
English Learner (EL)	734	14.7
None	126	10.4
Gender		
Male	2,120	51.1
Female	2,029	48.9
Race/Ethnicity		
Black/African American	3,057	73.7
Hispanic/Latino	723	17.4
White	251	6.0
Multiracial	78	1.9
Asian	25	0.6
Native Hawaiian/Pacific Islander	10	0.2
American Indian/Alaska Native	†	–
Special Education Status		
None	892	73.6
Special Education	101	8.3
Exited Special Education	14	1.2
504 plan	11	0.9
Exited 504	0	0.0

Note. ELs = English Learners. Non-ELs = Non-English Learners. Never English Learner refers to students who have never been classified as an English Learner. Exited English Learner refers to students who were once classified as English Learners but had since discontinued receiving English language services.

† data suppressed due to subgroup $n < 10$

Table 7

Frequency of Students with Complete DIBELS 8 Data Across Timepoints

			EL Status		EL Overall Language Proficiency Level				
Year	Time	All Students	ELs	Non-ELs	Entering	Emerging	Developing	Expanding	Unknown
2019-20	BOY	4,014	568	3,446	279	69	64	27	129
	MOY	4,024	584	3,440	277	69	64	26	148
2020-21	BOY	3,820	572	3,248	270	63	62	23	154
	MOY	3,532	554	2,978	264	62	61	25	142
	EOY	3,246	536	2,710	262	61	56	21	136
Total *n*		**4,149**	**608**	**3,541**	**284**	**69**	**64**	**27**	**164**

Note. ELs = English Learners. Non-ELs = Non-English Learners. BOY = Beginning of Year. MOY = Middle of Year. EOY = End of Year.

Descriptive Statistics

Predictors

English Learner (EL) Indicator. Because the number of Former-EL students in the study sample was very small relative to the size of the total sample (n = 21), Former-ELs were included in the Non-EL group for all analyses[1]. ELs attended 48 of the 93 schools in this study's overall sample. The majority of EL students within this sample spoke Spanish at home (n = 402), although a total of 21 non-English home languages were reported.

English Language Proficiency. Of the 444 EL students with available ACCESS for ELLs data from kindergarten, approximately 64% scored within the *Entering* range for Overall Proficiency (Overall Proficiency score of 1.0-1.9; Table 8), 11.3% scored within the *Emerging* range (Overall Proficiency score of 2.0-2.9), 10.5% scored within the *Developing* range, and 4.4% scored within the *Expanding* range. Approximately 27% of EL students within this sample (n = 164) did not have ACCESS for ELLs data from kindergarten and were categorized as *Unknown* proficiency.

DIBELS 8 Measures

DIBELS 8 descriptive statistics are presented in Table 9 for all students, disaggregated by EL status, and for EL students by Overall Language Proficiency Level. Visual analysis of Q-Q plots shows abnormal distributions of scores within this sample across all DIBELS 8 measures examined (see Appendix B). Further, visual analysis of histograms reveals positive skew in the distribution of scores and floor effects for all students during the BOY 19-20 timepoint for all DIBELS 8 measures studied (i.e., NWF-CLS, NWF-WRC, WRF; see Appendix C). Histograms illustrating NWF-WRC and WRF scores from all students

[1] Including Former ELs within the EL subsample was considered. Kindergarten WIDA ACCESS scores available for these students indicate English language proficiency near or above the threshold for EL classification, therefore most Former EL students within this sample are unlikely to have received EL services.

demonstrate a continued floor effect across all timepoints, although the severity of that effect lessens at the later time points in each school year. The distribution of scores for NWF-CLS appears to normalize during both MOY timepoints as well as EOY for 20-21. Histograms for EL student data show floor effects across all timepoints for NWF-CLS, while Non-EL data demonstrates floor effects during BOY timepoints only. Disaggregated plots for NWF-WRC and WRF demonstrate floor effects for both EL and Non-EL data across all timepoints. Additionally, visual analysis of box plots showing median scores across time for all participating schools demonstrate cross-school differences in initial skill levels and changes in skill over time for non-EL and EL students (see Appendix D).

Table 8

EL Students' ACCESS for ELLs Overall Language Proficiency Levels

Proficiency Level	Overall Composite Score	*n*	%
Entering	1.0 - 1.9	284	46.7
Emerging	2.0 - 2.9	69	11.3
Developing	3.0 - 3.9	64	10.5
Expanding	4.0 - 4.9	27	4.4
Bridging	5.0 - 5.9	0	0.0
Reaching	6.0	0	0.0
Unknown	–	164	27.0

Note. EL = English Learner; Unknown = ACCESS for ELLs scores unavailable.

Table 9

Descriptive Statistics for DIBELS 8 Measure Outcomes

					EL Status				EL Students' Overall Language Proficiency Level									
			All Students		Non-ELs		ELs		Entering		Emerging		Developing		Expanding		Unknown	
Variable	Year	Time	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
NWF-CLS	2019-20	BOY	25.1	20.2	26.2	20.3	17.9	18.1	15.3	12.8	22.9	10.9	35.9	24.8	43.3	23.8	6.8	13.0
		MOY	47.8	28.9	49.3	28.8	39.1	27.7	36.1	19.1	48.7	21.9	65.0	34.7	74.3	33.5	22.8	24.4
	2020-21	BOY	34.3	26.6	35.5	26.6	27.7	25.3	23.9	18.6	37.0	21.8	51.2	32.9	63.7	33.6	15.8	19.1
		MOY	50.5	31.3	51.5	30.9	45.5	32.8	41.7	27.9	52.0	25.5	71.5	39.5	81.7	36.2	30.7	26.7
		EOY	60.3	35.7	61.1	35.0	56.0	38.7	51.7	34.6	65.4	31.2	86.4	41.9	98.4	35.0	41.0	36.3
NWF-WRC	2019-20	BOY	3.9	6.9	4.1	7.1	2.6	5.6	1.4	3.2	3.1	3.9	7.4	9.9	9.2	10.5	1.1	3.6
		MOY	12.4	11.5	13.0	11.6	9.1	10.6	7.3	8.0	12.2	9.3	19.0	13.5	22.8	13.1	4.6	8.5
	2020-21	BOY	8.2	9.2	8.5	9.3	6.5	8.5	5.0	6.2	9.6	7.9	13.5	12.0	19.2	11.6	3.2	6.0
		MOY	13.9	11.5	14.2	11.5	12.3	11.7	10.5	10.3	15.0	9.5	23.2	13.3	25.2	13.4	7.5	9.1
		EOY	17.2	12.9	17.5	12.8	15.7	13.5	14.0	12.2	19.4	11.5	26.4	14.2	31.0	13.2	10.6	12.2
WRF	2019-20	BOY	10.3	11.6	11.1	11.9	5.7	8.2	3.4	4.5	8.1	6.9	15.2	11.7	19.0	10.3	1.9	5.3
		MOY	20.3	17.8	21.4	18.1	13.8	14.3	10.6	8.5	19.5	12.6	28.1	16.4	41.7	21.1	6.0	9.3
	2020-21	BOY	17.0	17.7	18.1	17.9	10.6	14.6	7.0	10.0	16.8	16.1	26.7	17.2	35.3	17.7	4.2	8.4
		MOY	25.5	20.6	26.9	20.7	17.9	18.8	13.9	15.3	23.8	16.8	37.0	20.6	46.8	16.1	9.7	13.1
		EOY	30.0	22.6	31.6	22.7	21.5	20.0	17.8	16.8	28.7	18.1	41.1	21.1	49.4	19.2	12.9	15.8

Note. ELs = English Learners. Non-ELs = Non-English Learners. NWF-CLS = Nonsense Word Fluency - Correct Letter Sounds. NWF-WRC = Nonsense Word Fluency - Words Read Correct. WRF = Word Reading Fluency. BOY = Beginning of Year. MOY = Middle of Year. EOY = End of Year.

Research Question 1: All Students

Decoding Skills

Table 10 presents results from the piecewise hierarchical growth model fit to investigate Research Question 1, described above. All fixed effects were statistically significant for NWF-CLS at the $p < .001$ level. Overall, variance accounted for at the school level was 23.5%, while 59% was at the student level[2]. Results estimate that students scored, on average, approximately 24.26 correct letter sounds (CLS) on an NWF probe at the beginning of the 2019-20 school year. Students were estimated to grow by an additional 21.96 letter sounds between the BOY and MOY timepoint during that school year, resulting in an estimated average score of 46.21 CLS (see Figure 1). Model estimates indicate that upon returning to school after COVID-19 related disruptions (i.e., BOY 2020-21), students lost an average of 12.18 CLS (-55% of growth estimated from BOY 2019-20 to MOY 2019-20) and began the year scoring an estimated average of 34.03 CLS during that time. Positive, but significantly slower growth was observed for the 2020-21 school year, with a more substantial slowdown of growth observed between the MOY and EOY timepoints.

Blending Skills

All fixed effects for NWF-WRC were statistically significant at $p < .001$ (Table 10). School-level differences accounted for 25.3% of the variance in student NWF-WRC scores, while 56.4% of the variance was explained by differences across students nested within schools. Results indicate that on average, students scored 3.74 NWF-WRC at the BOY 2019-20 timepoint. Students were estimated to grow by an average of 8.19 NWF-WRC between BOY and MOY of the 2019-20 school year, during which they demonstrated the largest growth of all

[2] See Appendix E for base model Residual Plots.

Table 10

Research Question 1: Piecewise Growth Model Results for All Students Across All Measures

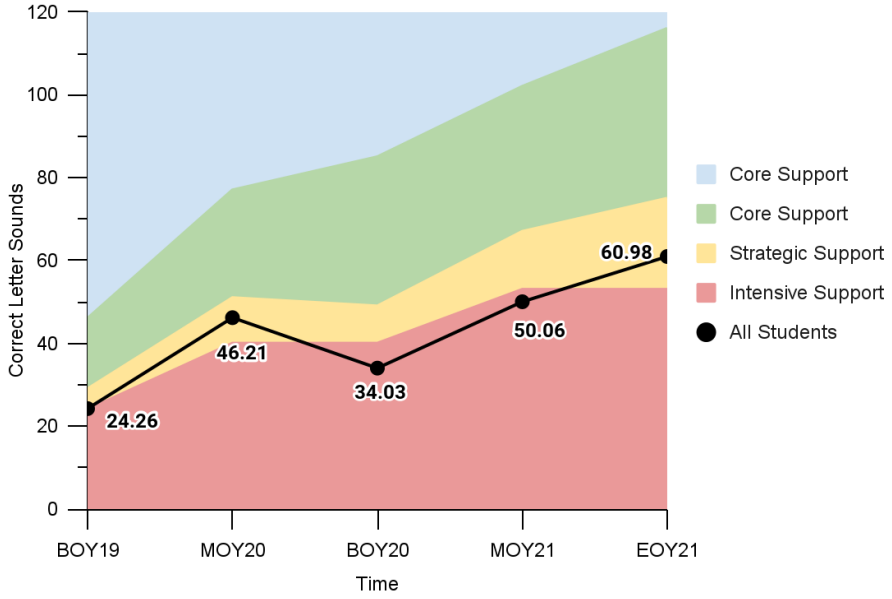
	NWF-CLS			NWF-WRC			WRF		
Fixed Effects	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*
Intercept	24.26 (0.78)	31.20***	–	3.74 (0.25)	14.94***	–	10.07 (0.41)	24.70***	–
Time_Instruction	21.96 (0.62)	35.32***	–	8.19 (0.26)	32.13***	–	9.64 (0.26)	36.82***	–
COVID_Summer	-12.18 (0.80)	-15.24***	-0.42	-3.76 (0.31)	-12.10***	-0.34	-2.36 (0.30)	-7.83***	-0.14
COVID_Instruction	-5.93 (1.07)	-5.56***	-0.20	-2.58 (0.44)	-5.80***	-0.23	-1.21 (0.42)	-2.91**	-0.07
Spring21	-11.03 (0.42)	-26.44***	-0.38	-4.30 (0.16)	-26.57***	-0.39	-3.97 (0.19)	-20.97***	-0.24
Random Effects	Variance (*SD*)			Variance (*SD*)			Variance (*SD*)		
Student Level									
Intercept	274.17 (16.56)			31.92 (5.65)			118.12 (10.87)		
Time_Instruction	64.28 (8.02)			13.88 (3.73)			22.83 (4.78)		
COVID_Summer	110.39 (10.51)			14.05 (3.75)			44.29 (6.66)		
COVID_Instruction	55.01 (7.42)			9.19 (3.03)			29.22 (5.41)		
School Level									
Intercept	43.97 (6.63)			4.28 (2.07)			11.31 (3.36)		
Time_Instruction	25.83 (5.08)			4.40 (2.10)			4.10 (2.03)		
COVID_Summer	46.96 (6.85)			7.18 (2.68)			5.42 (2.33)		
COVID_Instruction	84.37 (9.19)			15.05 (3.88)			11.28 (3.36)		
Residual Variance	149.38 (12.22)			22.49 (4.9)			30.57 (5.53)		

Note. NWF-CLS = Nonsense Word Fluency - Correct Letter Sounds. NWF-WRC = Nonsense Word Fluency - Words Read Correct. WRF = Word Reading Fluency.

*** $p < .001$, ** $p < .01$

Figure 1

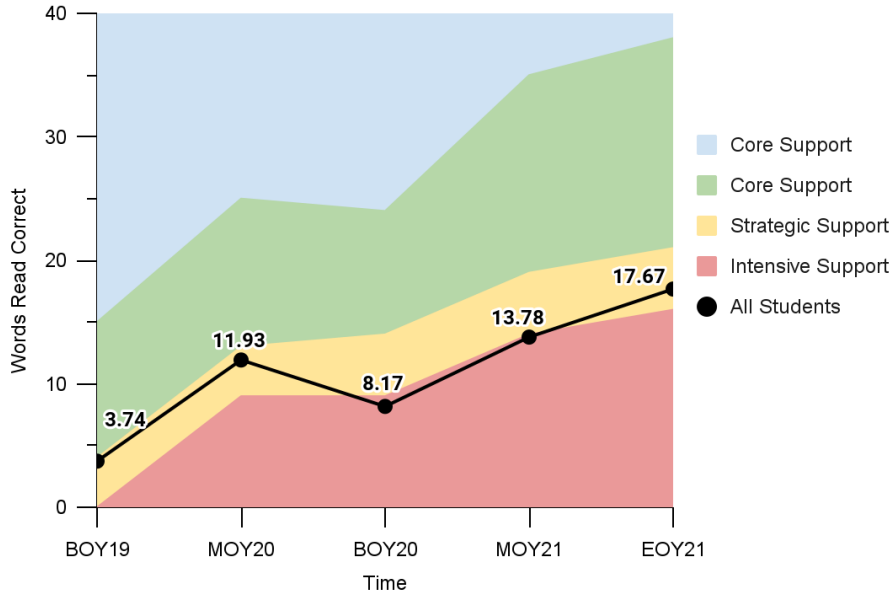
Estimated NWF-CLS Piecewise Growth for All Students



Note. Support bands determined by the University of Oregon (2020c).

Figure 2

Estimated NWF-WRC Piecewise Growth for All Students



Note. Support bands determined by the University of Oregon (2020c).

timepoints observed (see Figure 2). Overall, students were estimated to score 3.13 WRC lower during BOY 2020-21 (i.e., after COVID-19 related disruptions; -46% of growth estimated from

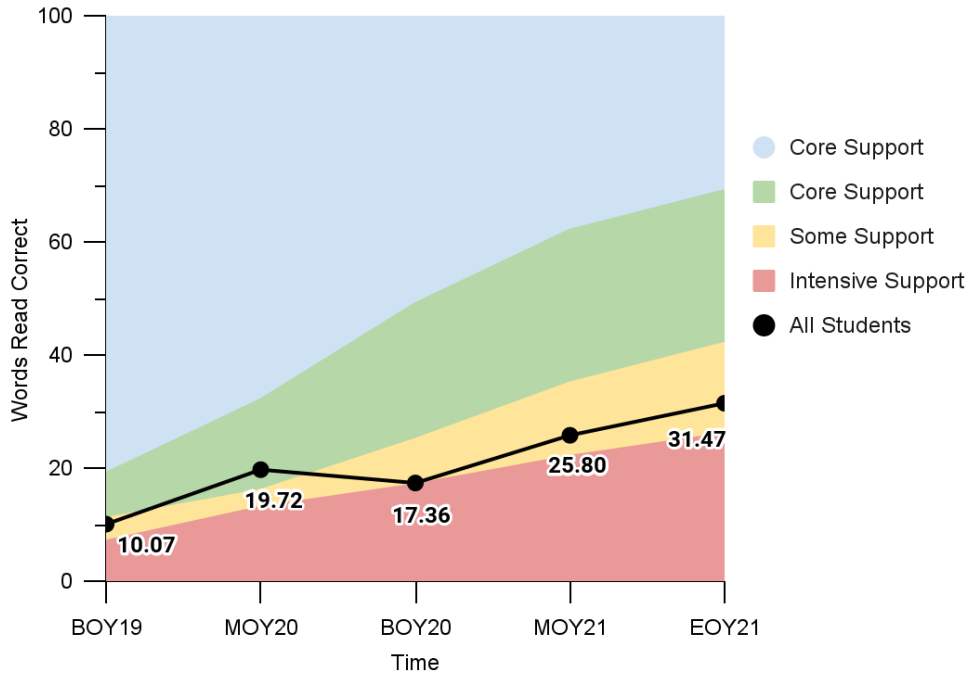
BOY to MOY 2019-20), before demonstrating positive, but slower growth through the remainder of the school year.

Sight Word Reading Skills

All fixed effects for WRF were statistically significant at $p < .01$ (Table 10). School-level differences accounted for 11.6% of the variance in student WRF scores, while differences between students nested within school accounted for 56.4% of the variance. Overall, students were estimated to begin the 2019-20 school year reading 10.07 real words correctly, on average, and grow by an additional 9.64 words by MOY 2019-20 (see Figure 3). Students lost an average of 2.36 words between MOY 2019-20 and BOY 2020-21 (-24% of the growth demonstrated from BOY to MOY 2019-20), followed by positive but slightly slower growth from BOY to MOY 2020-21 and even slower growth to EOY 2020-21.

Figure 3

Estimated WRF Piecewise Growth for All Students



Note. Support bands determined by the University of Oregon (2020c).

Research Question 2: EL Students

Decoding Skills

Results from piecewise hierarchical growth models using student EL status as a level 2 predictor are provided in Tables 11 and 12. For NWF-CLS, all fixed effects from the baseline model remained statistically significant at $p < .001$. Fixed effect differences for ELs relative to Non-ELs were statistically significant for all terms in the model at $p < .05$ with the exception of Spring21, which represents the estimated growth from MOY to EOY 2020-21. Approximately 58.6% of the variance in student NWF-CLS scores is attributable to differences between students nested in schools, while 23.7% of the variance is attributable to school-level differences (Table 11). Model results indicate that at the BOY 2019-20 timepoint, EL student scores were, on average, 10.17 CLS lower than those of Non-ELs students (15.22 and 25.39 CLS, respectively, see Figure 4). EL students continued to score lower than their Non-EL peers throughout all timepoints observed with differences in means ranging from 9.47 at MOY 2020-21 to 12.62 at MOY 2019-20. Non-EL students demonstrated a larger decrease in scores after the onset of the COVID-19 pandemic (-12.44 from MOY 2019-20 to BOY 2020-21) compared to their EL student peers (-10.24 CLS from MOY 2019-20 to BOY 2020-21). Non-EL students also showed a more substantial deceleration in growth between BOY to MOY 2020-21 (72% of growth estimated from BOY to MOY 2019-20) in comparison to their EL peers (85% of growth estimated from BOY to MOY 2019-20). Differential learning loss and deceleration in growth across groups allowed the performance gap between EL and Non-EL students, which was initially widening from BOY to MOY 2019-20 to close slightly at the beginning of the 2020-21 school year and maintain through the remainder of the school year.

Table 11

Research Question 2: Piecewise Longitudinal Growth Model Variance Based on EL Status

Random Effects	NWF-CLS		NWF-WRC		WRF	
	Variance	*SD*	Variance	*SD*	Variance	*SD*
Student Level						
Intercept	266.00	16.31	31.74	5.63	115.27	10.74
Time_Instruction	64.01	8.00	13.75	3.71	22.58	4.75
COVID_Summer	110.43	10.51	14.20	3.77	44.34	6.66
COVID_Instruction	54.07	7.35	8.95	2.99	29.41	5.42
School Level						
Intercept	42.51	6.52	4.33	2.08	10.20	3.19
Time_Instruction	26.04	5.10	4.21	2.05	4.05	2.01
COVID_Summer	47.03	6.86	7.16	2.68	5.39	2.32
COVID_Instruction	84.28	9.18	14.85	3.85	11.25	3.35
Residual Variance	149.33	12.22	22.40	4.73	30.56	5.53

Note. NWF-CLS = Nonsense Word Fluency - Correct Letter Sounds. NWF-WRC = Nonsense Word Fluency - Words Read Correct. WRF = Word Reading Fluency.

Table 12

Research Question 2: Piecewise Longitudinal Growth Model Fixed Effects Based on EL Status

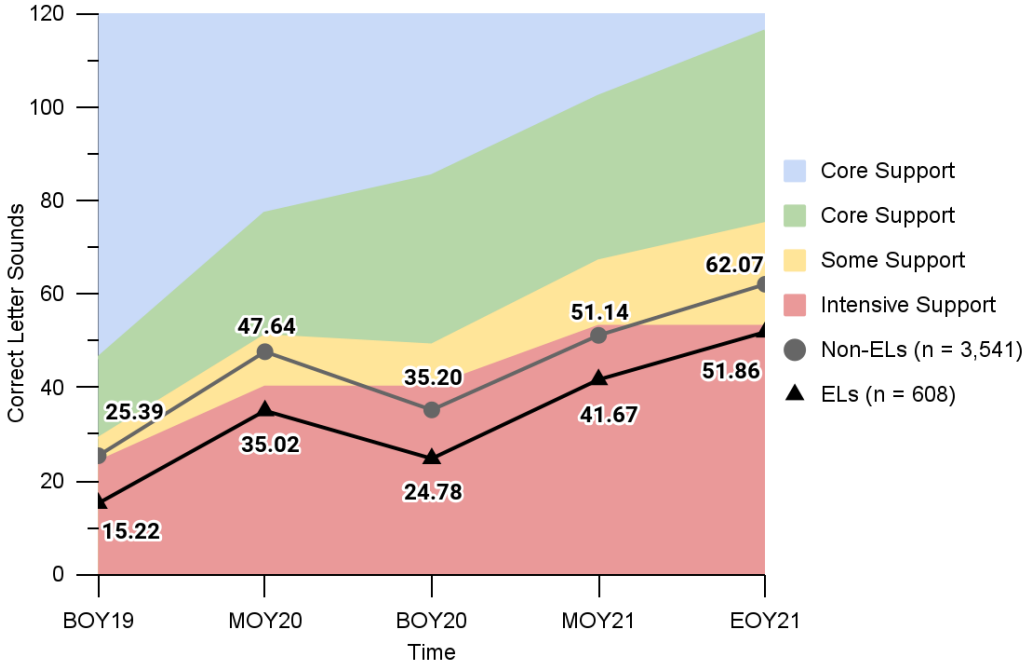
Fixed Effects	ELs			Non-ELs		
	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*
NWF-CLS						
Intercept	-10.17 (1.12)	-9.11***	-0.35	25.39 (0.78)	32.71***	–
Time_Instruction	-2.45 (1.00)	-2.44*	-0.08	22.25 (0.64)	35.04***	–
COVID_Summer	2.21 (1.12)	1.98*	0.08	-12.45 (0.81)	-15.35***	-0.43
COVID_Instruction	3.39 (1.46)	2.32*	0.12	-6.30 (1.08)	-5.84***	-0.22
Spring21	1.71 (1.15)	1.49	0.06	-11.32 (0.45)	-24.93***	-0.39
NWF-WRC						
Intercept	-2.04 (0.40)	-5.09***	-0.18	3.95 (0.26)	15.48***	–
Time_Instruction	-2.34 (0.40)	-5.82***	-0.21	8.49 (0.26)	33.25***	–
COVID_Summer	1.44 (0.42)	3.38***	0.13	-3.94 (0.31)	-12.53***	-0.36
COVID_Instruction	2.25 (0.57)	3.95***	0.20	-2.84 (0.45)	-6.36***	-0.26
Spring21	2.01 (0.44)	4.52***	0.18	-4.62 (0.18)	-26.24***	-0.42
WRF						
Intercept	-6.10 (0.65)	-9.40***	-0.37	10.79 (0.40)	27.05***	–
Time_Instruction	-1.98 (0.48)	-4.14***	-0.12	9.87 (0.27)	36.96***	–
COVID_Summer	-0.28 (0.55)	-0.50	-0.02	-2.33 (0.31)	-7.58***	-0.14
COVID_Instruction	0.72 (0.69)	1.03	0.04	-1.27 (0.42)	-2.99**	-0.08
Spring21	0.31 (0.52)	0.60	0.02	-4.03 (0.21)	-19.53***	-0.24

Note. ELs = English Learners. Non-ELs = Non-English Learners. NWF-CLS = Nonsense Word Fluency - Correct Letter Sounds. NWF-WRC = Nonsense Word Fluency - Words Read Correct. WRF = Word Reading Fluency.

*** $p < .001$, ** $p < .01$, * $p < .05$

Figure 4

Estimated NWF-CLS Piecewise Growth for EL and Non-EL Students



Note. Support bands determined by the University of Oregon (2020c).

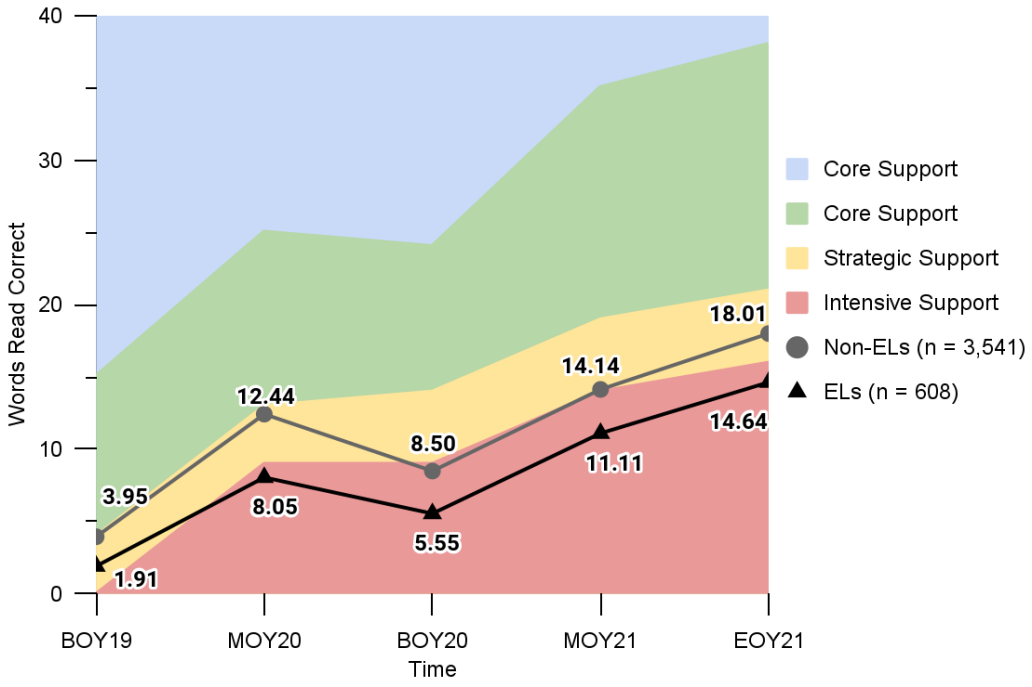
Blending Skills

Similar to NWF-CLS, all base model fixed effects for NWF-WRC remained statistically significant at $p < .001$ (Table 12). In addition, all fixed effects for ELs differed from those for Non-ELs at $p < .001$. Student-level differences accounted for 77.4% of the variance in NWF-WRC scores, while school-level differences accounted for an additional 11.6% of the variance (Table 11). On average, EL students were estimated to begin the 2019-20 school year with scores approximately 2 fewer WRC than their non-EL peers, which increased to a difference of 4.39 WRC between groups at MOY 2019-20 (Figure 5). Non-EL students demonstrated more substantial learning loss than did EL students between the MOY 2019-20 and BOY 2020-21 timepoints (-3.94 and -2.95, respectively), followed by greater deceleration in scores post-pandemic onset (i.e., 2020-21 school year) relative to EL students. Differential

learning loss and deceleration in growth across groups allowed the performance gap between EL and Non-EL students to close slightly at the beginning of the 2020-21 school year, before widening slightly through the remainder of the school year.

Figure 5

Estimated NWF-WRC Piecewise Growth for EL and Non-EL Students



Note. Support bands determined by the University of Oregon (2020c).

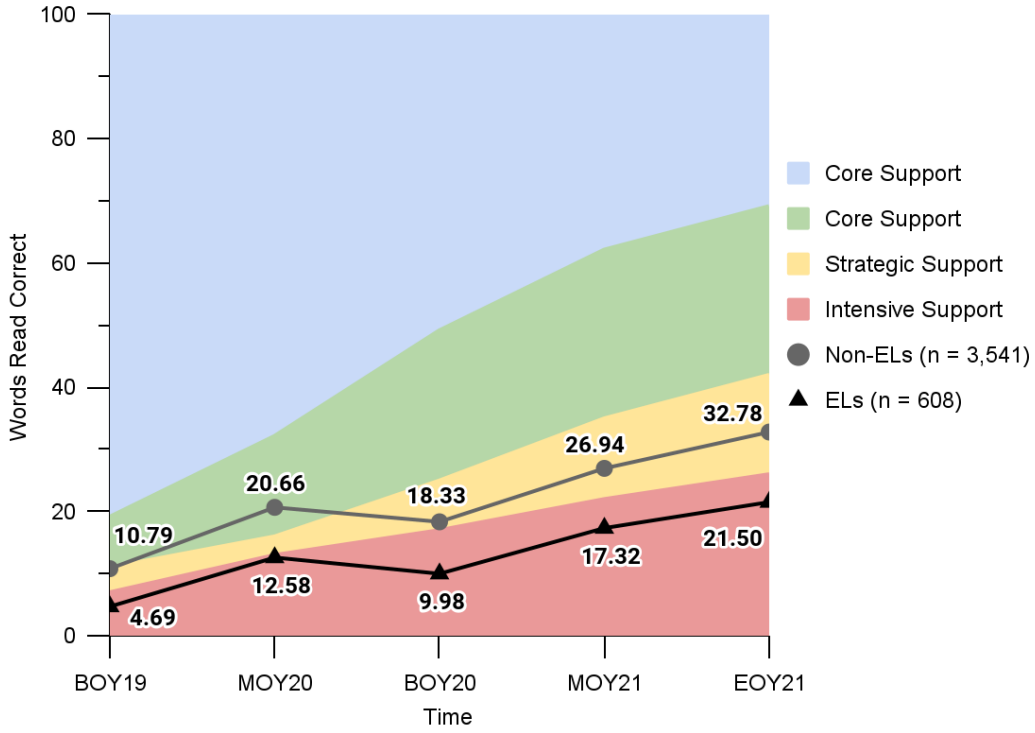
Sight Word Reading Skills

For WRF, all base model fixed effects remained statistically significant at $p < .01$ (see Table 12). For EL students, only the Intercept and Time_Instruction fixed effect, which represents growth estimated from BOY to MOY 2019-20, were statistically significant ($p < .001$). Differences between schools accounted for 11.3% of the variance in student WRF scores and differences between students nested within schools accounted for an additional 77.5% (Table 11). EL students were estimated to score, on average, 6.1 points lower than their Non-EL peers at the beginning of the 2019-20 school year. Differences in learning loss across groups

from MOY 2019-20 to BOY 2020-21, as well as growth post-pandemic onset, were not statistically significant. While growth trajectories do not differ significantly after BOY 2020-21, visual analysis of growth trajectories reveal that discrepancies between EL and Non-EL scores on WRF appear to grow over time, resulting in a gap of 11.28 words per minute by EOY 2020-21 (Figure 6).

Figure 6

Estimated WRF Piecewise Growth for EL and Non-EL Students



Note. Support bands determined by the University of Oregon (2020c).

Research Question 3: EL Students by English Proficiency Level

Decoding Skills

Random effect variances for the EL student base model are provided in Table 13 (see Appendix F for EL base model fixed effect coefficients). When including EL students' Overall Language Proficiency Level as a level 2 predictor for NWF-CLS scores, intercept terms remained statistically significant for all groups at $p < .001$ and the Time_Instruction interaction

term remained statistically significant at $p < .05$ for all groups except for students with Emerging proficiency (Table 14). All other interaction terms remained statistically significant for the Entering proficiency group, as well as the COVID_Instruction term for students in the Unknown proficiency group. Variance in NWF-CLS scores continued to be observed at the school-level after EL students' language proficiency was introduced into the model. Overall, 64.5% of variance in EL students' NWF-CLS scores was attributable to differences across students nested within schools, and 10.1% was attributable to school-level differences (see Table 15). Adding EL students' English language proficiency level into the EL base model reduced the total variance observed between schools by 26% and between students by 23%.

Table 13

EL Student Base Model Random Effect Variance

	Variance (*SD*)		
Random Effects	NWF-CLS	NWF-WRC	WRF
Student Level			
Intercept	218.61 (14.79)	20.78 (4.56)	58.10 (7.62)
Time_Instruction	78.29 (8.85)	14.09 (3.75)	20.23 (4.50)
COVID_Summer	110.37 (10.51)	11.38 (3.37)	29.76 (5.46)
COVID_Instruction	73.92 (8.60)	11.40 (3.38)	15.19 (3.90)
School Level			
Intercept	39.41 (6.28)	3.62 (1.90)	9.08 (3.01)
Time_Instruction	19.22 (4.38)	3.16 (1.78)	4.29 (2.07)
COVID_Summer	11.67 (3.42)	1.51 (1.23)	1.79 (1.34)
COVID_Instruction	7.92 (2.81)	3.89 (1.97)	1.69 (1.30)
Residual Variance	147.17 (12.13)	20.58 (4.54)	22.27 (4.72)

NWF-CLS = Nonsense Word Fluency - Correct Letter Sounds. NWF-WRC = Nonsense Word Fluency - Words Read Correct. WRF = Word Reading Fluency.

Table 14

Research Question 3: Piecewise Longitudinal Growth Model Fixed Effects Based on EL Students' Overall Language Proficiency

	Entering			Emerging			Developing			Expanding			Unknown		
Fixed Effects	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*
NWF-CLS															
Intercept	15.56 (1.38)	11.24***	–	7.55 (2.24)	3.37***	0.31	19.82 (2.30)	8.60***	0.83	27.81 (3.34)	8.34***	1.16	-9.32 (1.73)	-5.38***	-0.39
Time_Instruction	20.62 (1.38)	14.97***	–	4.72 (2.58)	1.83	0.20	7.67 (2.65)	2.89**	0.32	10.05 (3.89)	2.58*	0.42	-4.85 (2.02)	-2.40*	-0.20
COVID_Summer	-11.80 (1.37)	-8.63***	-0.49	-0.32 (2.78)	0.12	-0.01	-0.30 (2.83)	-0.11	-0.01	2.61 (4.27)	0.61	0.11	5.58 (2.06)	2.71**	0.23
COVID_Instruction	-3.53 (1.68)	-2.10*	-0.15	-5.39 (3.57)	-1.51	-0.22	-1.80 (3.63)	-0.50	-0.08	-10.46 (5.43)	-1.93	-0.44	1.33 (2.72)	0.63	0.06
Spring21	-10.50 (1.49)	-7.04***	-0.44	-3.05 (3.41)	-0.90	-0.13	-1.67 (3.48)	-0.48	-0.07	4.89 (5.29)	0.93	0.20	5.06 (2.61)	1.94	0.21
NWF-WRC															
Intercept	1.65 (0.49)	3.36**	–	1.70 (0.80)	2.13*	0.19	5.89 (0.82)	7.16***	0.67	7.80 (1.19)	6.54***	0.88	-0.38 (0.62)	-0.62	-0.04
Time_Instruction	5.93 (0.53)	3.36***	–	2.97 (0.97)	3.06**	0.34	5.43 (1.00)	5.43***	0.62	7.78 (1.47)	5.31***	0.88	-2.58 (0.76)	-3.39***	-0.29
COVID_Summer	-2.36 (0.50)	-4.74***	-0.27	-0.41 (1.00)	-0.41	-0.05	-2.78 (1.02)	-2.74**	-0.32	-1.10 (1.54)	-0.71	-0.12	1.05 (0.74)	1.41	0.12
COVID_Instruction	-0.47 (0.72)	-0.65	-0.05	-2.70 (1.34)	-2.02*	-0.31	-1.39 (1.36)	-1.02	-0.16	-7.76 (2.02)	-3.83***	-0.88	1.26 (1.02)	1.24	0.14
Spring21	-2.18 (0.55)	-3.95***	-0.25	2.65 (1.26)	-2.10*	-0.30	-4.20 (1.29)	-3.26**	-0.48	-2.39 (1.96)	-1.22	-0.27	2.12 (0.97)	2.20*	0.24
WRF															
Intercept	3.90 (0.65)	6.03***	–	4.69 (0.98)	4.78***	0.43	11.40 (1.01)	11.29***	1.04	15.32 (1.46)	10.48***	1.39	-1.84 (0.75)	-2.45*	-0.17
Time_Instruction	7.47 (0.57)	13.08***	–	4.11 (1.02)	11.29***	0.37	5.48 (1.05)	5.22***	0.50	15.47 (1.54)	10.06***	1.41	-3.05 (0.80)	-3.82***	-0.28
COVID_Summer	-3.77 (0.60)	-6.24***	-0.34	0.53 (1.18)	0.45	0.05	3.15 (1.21)	2.61**	0.29	-1.94 (1.81)	-1.07	-0.18	2.09 (0.88)	2.39*	0.19
COVID_Instruction	-0.44 (0.65)	-0.67	-0.04	-2.13 (1.37)	-1.55	-0.19	-1.89 (1.39)	-1.36	-0.17	-10.82 (2.09)	-5.18***	-0.98	1.40 (1.05)	1.34	0.13
Spring21	-3.00 (0.57)	-5.28***	-0.27	-3.50 (1.30)	-2.70**	-0.32	-2.56 (1.32)	-1.93	-0.23	12.94 (2.01)	-6.43***	-1.18	2.46 (0.99)	2.48*	0.22

Note. NWF-CLS = Nonsense Word Fluency - Correct Letter Sounds. NWF-WRC = Nonsense Word Fluency - Words Read Correct. WRF = Word Reading Fluency.

*** $p < .001$, ** $p < .01$, * $p < .05$

Table 15

Research Question 3: Piecewise Longitudinal Growth Model Random Effect Variance Based on EL Students' Overall Language Proficiency

	NWF-CLS		NWF-WRC		WRF	
Random Effects	Variance	*SD*	Variance	*SD*	Variance	*SD*
Student Level						
Intercept	120.73	10.99	13.93	3.73	29.96	5.47
Time_Instruction	67.63	8.22	10.72	3.27	13.75	3.71
COVID_Summer	111.98	10.58	12.09	3.48	30.88	5.56
COVID_Instruction	70.04	8.37	10.20	3.19	12.08	3.48
School Level						
Intercept	25.06	5.01	2.94	1.72	6.44	2.54
Time_Instruction	15.24	3.90	2.42	1.56	3.18	1.78
COVID_Summer	8.54	2.91	1.36	1.17	1.96	1.40
COVID_Instruction	9.09	3.01	4.03	2.01	1.72	1.31
Residual Variance	146.44	12.10	20.12	4.49	21.13	4.60

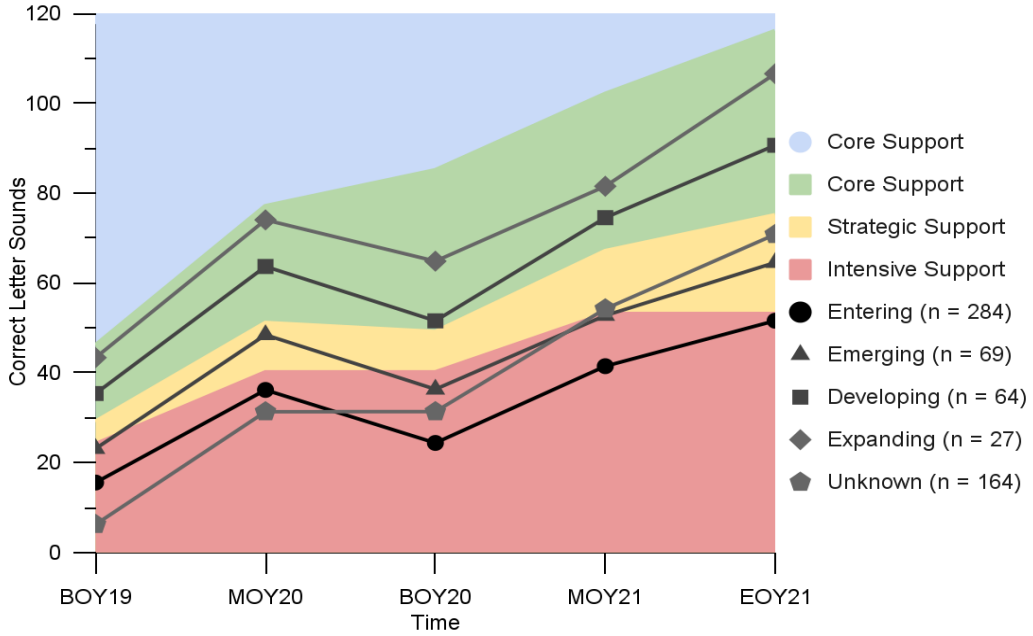
Note. NWF-CLS = Nonsense Word Fluency - Correct Letter Sounds. NWF-WRC = Nonsense Word Fluency - Words Read Correct. WRF = Word Reading Fluency.

Figure 7 illustrates estimated student growth in NWF-CLS scores by language proficiency group. At the BOY 2019-20 timepoint, EL students with Unknown English language proficiency scored the lowest of all EL groups. Students with Emerging proficiency scored higher than students in the Entering group across all timepoints. Students in the Expanding proficiency group scored higher, on average, than all other proficiency groups across all timepoints, followed by students with Developing proficiency. With the exception of the Unknown proficiency group, who demonstrated very slight growth of 0.02 CLS, all groups demonstrated learning loss from the MOY 2019-20 to BOY 2020-21 timepoint (i.e., after COVID-19 related disruptions and before instruction resumed). This lack of learning loss for EL students with Unknown levels of English proficiency, coupled with the largest growth of any group across four of the five timepoints, resulted in predicted scores for students with Unknown

proficiency surpassing predicted scores for students with Entering proficiency at BOY 2020-21, as well as those for the Emerging proficiency group at MOY 2020-21.

Figure 7

Estimated NWF-CLS Piecewise Growth for EL Students Across English Language Proficiency Levels



Note. Support bands determined by the University of Oregon (2020c).

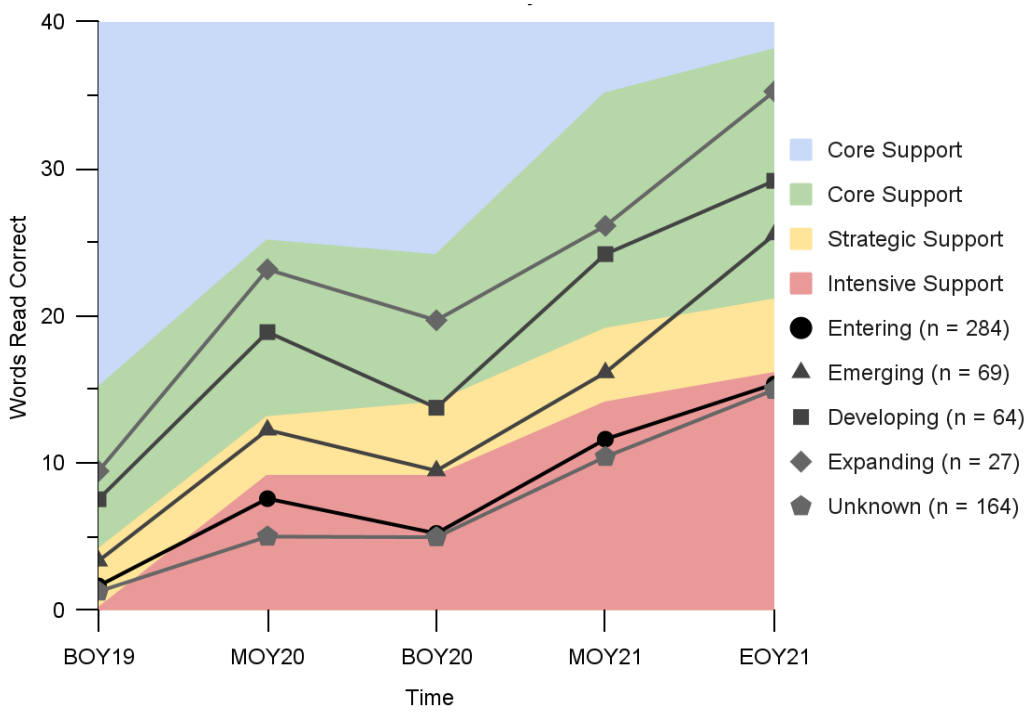
Blending Skills

Statistical significance of fixed effects for NWF-WRC varied substantially across English proficiency groups. The Time_Instruction and intercept fixed effects remained significant across all proficiency groups at $p < .05$, with the exception of the intercept for the Unknown proficiency group (see Table 14). Differences between students nested within schools accounted for 60.3% of variance in student NWF-WRC scores, while school-level differences accounted for 13.8% of variance. Entering English language proficiency level into the EL student base model as a level 2 predictor reduced the total variance observed between schools by 12% and the total variance between students within schools by 19% (see Tables 13 and 15).

Similar to their performance on NWF-CLS, EL students with Unknown language proficiency scored the lowest during the BOY 2019-20 timepoint. For all groups across all timepoints, student groups with more advanced language proficiency scored higher than student groups with relatively lower proficiency levels. Students with Developing proficiency showed the most pronounced learning loss between MOY 2019-20 and BOY 2020-21 (i.e., post-pandemic onset; -5.14 WRC), but demonstrated the largest quantitative growth of all groups between BOY and MOY 2020-21 (+6.67 WRC; see Figure 8). EL students with Emerging and Expanding proficiency levels both demonstrated more substantial increases from MOY to EOY 2020-21 than they did from BOY to MOY 2020-21. Students with Unknown proficiency scored very similarly to students with Entering proficiency across all timepoints, with the largest discrepancy occurring during MOY 2019-20 (2.58 WRC).

Figure 8

Estimated NWF-WRC Piecewise Growth for EL Students Across English Language Proficiency Levels



Note. Support bands determined by the University of Oregon (2020c).

Sight Word Reading Skills

For WRF, the Intercept terms and Time_Instruction fixed effects were statistically significant across all language proficiency levels at $p < .05$ (see Table 14). Statistical significance of other fixed effects varied across proficiency groups. Differences between schools accounted for 11% of the variance in EL students' WRF scores, while differences between students nested within schools accounted for 71.6% of the variance. Including English language proficiency level as a level 2 predictor reduced the total variance observed at the student- and school-levels by 30% and 21%, respectively (see Tables 13 and 15).

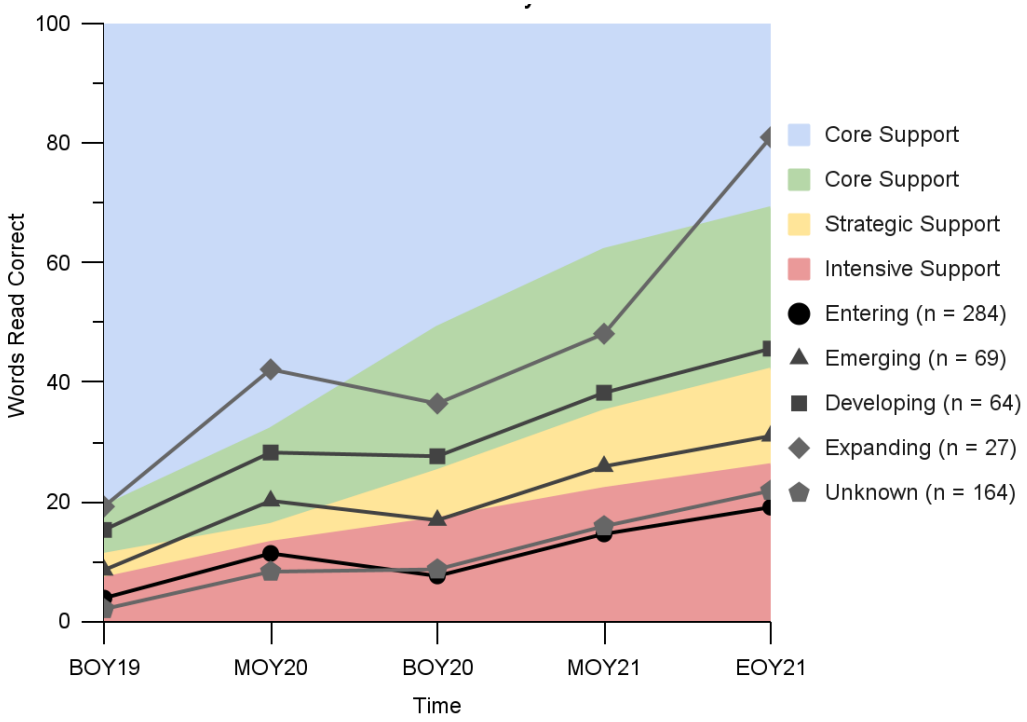
Patterns of BOY 2019-20 scores across proficiency groups were consistent with NWF-CLS and WRC estimates, in that the student groups with more advanced language proficiency scored higher, on average, than those groups with relatively lower proficiency and higher than students with Unknown proficiency (see Figure 9). EL students with Entering, Emerging, or Expanding proficiency levels demonstrated a loss of 3+ words from MOY 2019-20 to BOY 2020-21, while students with Unknown or Developing proficiency demonstrated minimal differences in scores between MOY 2019-20 to BOY 2020-21 (+0.38 and -0.62 words, respectfully). This learning gain, along with relatively substantial growth between the MOY 2019-20 and EOY 2020-21 time periods, allowed predicted scores for students with Unknown proficiency to exceed those for students with Entering proficiency across all timepoints of the 2020-21 school year.

Students who scored in the Expanding proficiency range demonstrated an estimated growth pattern much different from those of students in other proficiency groups. Despite demonstrating the largest growth between BOY and MOY 2019-20 (+22.94 words), their growth appeared much more similar to other groups from BOY to MOY 2020-21. From MOY

to EOY 2020-21, however, students with Expanding proficiency demonstrated an estimated average gain of 32.88 words, resulting in far higher predicted EOY 2020-21 scores than the predicted scores for all other language proficiency groups.

Figure 9

Estimated WRF Piecewise Growth for EL Students Across English Language Proficiency Levels



Note. Support bands determined by the University of Oregon (2020c).

IV. DISCUSSION

Summary of Results & Implications

Using early literacy benchmarking data from a large, urban school district, this study examined the impacts of COVID-19 related school closures on the reading outcomes of district students who were in first grade when COVID-19 related school closures occurred. It also sought to explore how those impacts differed for EL students in comparison to their Non-EL peers, and for EL students of varying English language proficiency levels. This study contributes to a small, but growing body of empirical evidence examining the impacts of COVID-19 related disruptions to education on student reading development during early elementary and for EL students in particular (Bailey et al., 2021). Early literacy in this study was measured via students' scores on DIBELS 8 Nonsense Word Fluency Correct Letter Sounds (NWF-CLS) and Word Reading Fluency (NWF-WRC) measures, as well as on the Word Reading Fluency (WRF) measure to examine how decoding, blending, and sight word reading skills changed over five measurement occasions from 2019 to 2021.

Research Question 1: All Students

For all outcome measures, students' overall estimated average scores demonstrated the greatest improvement between the beginning of year (BOY) and middle of year (MOY) during the 2019-2020 academic year, before the onset of the COVID-19 pandemic. Significant learning loss between MOY 2020 (i.e., pre-pandemic) and BOY 2020 (i.e., after COVID-19 pandemic onset) was observed for all reading skills investigated in this study. Students' observed learning loss in decoding and blending equated to nearly half of the improvement they made between BOY and MOY in the 2019-2020 year (55.5% of previous improvement on NWF-CLS and 45.9% on WRC), and approximately one quarter of improvement from BOY 2019 to MOY

2020 for sight word reading (-25% of improvement on WRF). Despite this substantial learning loss, students' estimated scores at the MOY 2021 demonstrated sufficient improvement from BOY 2020 to recover from these previous losses. Relative to BOY to MOY in 2019-2020, improvement was substantially less from BOY to MOY in 2020-2021, and even less from MOY to EOY. Previous empirical investigation of CBM growth patterns across grades suggest that variability in improvement rates may be typical both within and across grades (Nese et al., 2013). Because comparative DIBELS 8 benchmarking data from district students pre-pandemic was not utilized in this study, the extent to which the observed patterns of slower improvement are similar to typical patterns of reading skill development or may be attributable to COVID-19 related factors is unclear.

Research Question 2: EL Students

EL students began first grade performing about one third of a standard deviation lower than their Non-EL peers in decoding, blending, and sight word reading. In 2019-2020, from BOY to MOY, EL students improved about 2 items less in decoding, blending, and sight word reading than Non-EL students, but these differences amounted to much smaller effects, ranging from a low of -0.08 for decoding skills to -0.21 for blending skills. Surprisingly, Non-EL students demonstrated more substantial learning loss in decoding and blending than did EL students between MOY 2020 and BOY 2020, which resulted in reducing the gap between groups by 2.2 and 1.44, respectively, but did not differ significantly in the impact on sight word reading. At MOY 2021, ELs continued to show less negative changes in growth trends on both decoding and blending compared to their Non-EL peers and no difference in growth trends on sight word reading. By EOY 2021, blending skills were the only area where EL students continued to show less negative changes to growth trends than non-EL students.

These patterns combine to reveal that while EL students continued to perform below their Non-EL peers across all timepoints for all skills investigated, they exhibited progressively shrinking gaps in performance for blending, some reduction in gaps in decoding, but widening gaps in sight word reading. In fact, growth trajectories after the onset of the COVID-19 pandemic demonstrate that EL students were no longer falling further behind their Non-EL peers in decoding and blending, in sharp contrast to the widening gaps evidenced during the first half of the 2019-20 school year. While NWF and WRF both measure early literacy skills, WRF requires students to read real words, many of which have irregular spelling patterns and cannot be decoded (University of Oregon, 2020a). Since sight word reading relies on students' previous exposure to irregular words, differences in WRF trajectories for EL and Non-EL students may potentially reflect differences in students' exposure to and opportunities to practice reading irregular words. District and school practices during the COVID-19 pandemic, such as instructional modalities (i.e., virtual and/or in-person), curricula used for academic and ELL intervention, and/or procedures for allowing all and/or subgroups of students to return to in-person learning, may have impacted EL and non-EL students' growth post-pandemic onset.

Research Question 3: EL Students by English Language Proficiency Level

As hypothesized, differences in growth trajectories and patterns were observed across EL students' English language proficiency levels for all outcome reading skills investigated. That is, EL students who demonstrated higher levels of English proficiency in kindergarten (i.e., more advanced reading, writing, listening, and speaking skills) demonstrated greater skill and growth than their lower-proficient peers across multiple time periods and reading constructs. EL students with the most advanced English proficiency skills in kindergarten (i.e., the Expanding proficiency level) began first grade with more developed decoding, blending, and sight word

reading skills than their less English proficient peers and demonstrated more substantial improvement in blending and sight word reading than any other proficiency group through the end of second grade. EL students whose English proficiency was Unknown in kindergarten performed lower than all other EL proficiency groups at the beginning of first grade for all reading skills investigated, but demonstrated more substantial improvement in decoding than ELs with Entering and Emerging English proficiency. Perhaps most shockingly, ELs with Unknown proficiency were the only group estimated to demonstrate an increase in scores between the MOY 2020 and BOY 2020 timepoints (i.e., immediately pre- and post-pandemic onset) and on a measure of sight word reading (+0.38 words on WRF).

By definition, ELs with higher levels of English proficiency on the ACCESS for ELLs assessment are highly likely to demonstrate higher levels of literacy-related English language skills, such as speaking and reading. Thus, it is not surprising that ELs students with Developing and Expanding proficiency, the two highest proficiency levels observed within the EL student group, performed better than their lower English proficient peers on measures of early literacy skills. Growth trajectories for EL students with missing English proficiency data represent a group of students that may be highly heterogeneous in their English language proficiency, experience with English instruction, and/or exposure to English text, all of which should be considered when the interpreting these trajectories alongside other groups of ELs for whom baseline proficiency is available.

High Performing ELs. In multiple cases, EL students with higher levels of English proficiency were estimated to score, on average, higher than their non-EL peers. At BOY 2019, or the start of students' first grade year, EL students with Developing or Expanding language proficiency scored higher, on average, than their Non-EL peers in decoding and blending. By

EOY 2021, the subgroups of ELs who, on average, outscored Non-ELs expanded to include students with Emerging and Unknown proficiency. The same pattern was observed for sight word reading, with the exception of students with Unknown proficiency who were estimated to score lower than their Non-EL peers at the end of second grade. On a measure of sight word reading, ELs with Developing or Expanding proficiency outscored Non-EL students at both the beginning of first grade and the end of second grade. Sample sizes for EL language proficiency groups should be taken into consideration, however, particularly for the Expanding proficiency group which consisted of substantially fewer students than other EL proficiency groups and the Non-EL student group.

Variance Across Students and Schools

Despite attending schools within the same school district, up to 25.3% of the variance in student scores across the DIBELS 8 measures investigated were explainable by differences between schools. Such a large proportion of variance explained at the school-level is striking and raises questions about the potential differences between schools that may have influenced students' early literacy performance. Only slight changes in variance explained was observed when student EL status was entered into the model, which suggests that other factors are contributing to these differences in scores. Interestingly, EL students attended just under half of the total schools in the district. Baseline models including EL students only (and consequently their narrowed sample of schools) explained 48-61% of school-level variance originally observed in the model that included all students from all schools. That is, by simply narrowing the group of students to ELs and their schools of attendance, approximately half of the variance observed in student scores in the original model was explainable by differences between the

schools that EL students attended. Perhaps schools that serve EL students are more similar to each other than other within-district schools that did not serve EL students in this cohort.

The addition of EL students' English proficiency levels into the hierarchical model further reduced both between-school and between-student variance in scores, resulting in approximately 10-14% of variance observed that was explainable by school-level differences. While school-level differences explain less variance in scores after accounting for student EL status and EL students' English proficiency, factors beyond student characteristics appear to continue to influence student scores within this district. Investigation of the concentration of EL students within specific schools, as well as the concentration of EL students with less developed English proficiency across schools within the district were beyond the scope of this study, but may help explain differences observed in EL students' scores that remain unexplained.

A variety of school-level factors such as school size, instructional practices utilized, training and experience of staff, and/or student population demographics (e.g., neighborhood effects) may have contributed to the variability of skills demonstrated by students in different schools. Additionally, students' reading performance after the onset of the COVID-19 pandemic may have been impacted by family-based factors, such as parental availability to support at-home or hybrid learning, parental English language proficiency, or access to consistent and stable learning environments in the home. Thus, differences in scores observed immediately before and after the onset of the pandemic may be best understood as the combined impacts of school-based *and* family characteristics and practices and how these factors supported continued learning despite disruptions to typical education practices.

Study Limitations

Results from this study provide empirical evidence of students' reading development in one school district during the school years immediately before and after the onset of the COVID-19 pandemic. All students attended a single large, urban school district that serves a student body that is rich in cultural and linguistic diversity. While the demographics of this sample presented a valuable opportunity to examine the reading development of English learners, it is not representative of demographics observed at the national level and may differ from samples observed in other contexts (e.g., state-level, other districts, etc.). Additionally, because this sample included students from one school district only, and a variety of district-level factors no doubt contributed to students' reading development, such as the district's response to the onset of COVID-19 pandemic, the instructional practices and modalities employed during resumption of instruction in Fall 2020 (e.g., in-person vs. remote instruction), and the instructional models and curricula employed. As such, caution should be taken in generalizing this study's findings to other populations, even where locale and local demographics may be very similar.

It is important to note that students' benchmark scores on DIBELS measures represent raw scores, or the total number of correct responses provided by the student on a grade-level probe (University of Oregon, 2020). Because this study utilizes benchmarking data from both first and second grade, the complexity and difficulty of words that students were asked to read was not equal across all timepoints. Consequently, differences in scores observed between MOY 2020 (first grade) and BOY 2020 (second grade) may overestimate the learning loss experienced by students due, in part, to the increased difficulty of the second grade measures relative to those for first grade. Nevertheless, ross-grade comparisons are frequently made when

analyzing longitudinal CBM data, and differences in measure difficulty across grades are not anticipated to fully account for the differences observed after the transition to second grade. Without investigating comparison data from students in prior years, it remains unclear to what extent the slower improvement in early literacy skills during and within the 2020-21 school year are attributed to COVID-19 related disruptions to education or represent growth patterns that are typical during the second grade.

This study utilized EL students' Overall Proficiency Levels from the ACCESS for ELLs assessment administered during Spring of their kindergarten year as a baseline metric of language proficiency before those students received English language supports. Students' language proficiency levels during their first and second grade years were not accounted for in analyses. ELs may vary in the rate which they develop English language proficiency, so it cannot be assumed that students with similar language proficiency levels at kindergarten continued to be similar in first and second grade (Slama, 2012). It is also possible that ELs with more advanced English proficiency in kindergarten may have exited EL programs after being re-assessed in the Spring of subsequent years, and thus did not retain their EL indicator throughout the entirety of this study's timeframe.

Implications for Practice & Research

Although on average, students demonstrated sufficient improvement between BOY 2020 and MOY 2021 to recover from the learning loss that occurred between MOY 2020 and BOY 2020, what this study's growth trajectories do not convey is the missed opportunity to grow between the MOY 2020 and missing EOY 2020 timepoint where instruction was most severely disrupted. Due to this disruption, it is impossible to know how much improvement these students would have demonstrated had they instead received instruction that was more

consistent with typical educational practices. One study found that compared to a previous cohort of students, third grade students demonstrated approximately 20% less improvement in 2020-21 (Kogan & Lavertu, 2022). Additionally, Kogan and Lavertu (2022) determined that over half of the learning loss observed by Spring 2021 could be attributed to differences in learning during the 2019-20 school year. In the context of this study, although student improvement during the 2020-21 school year was observed in relation to students' estimated average scores during BOY 2020, it is unclear how this growth compares to growth observed in second graders whose 1st grade instruction was not disrupted, nor is it clear whether the actual improvement observed is sufficient for students to "catch up" to the achievement of previous cohorts. DIBELS 8 benchmark criteria, however, suggest that students progressed more slowly than expected post-pandemic onset, moving away from the "core support" range and instead falling within the "intensive" or "strategic support" range, for example, in many cases.

Observation and analysis of student growth trajectories may aide school officials in better understanding students' current skills in decoding, blending, and sight word reading to provide instructional supports that are better aligned to the needs of students who experienced a disruption to instruction during the COVID-19 pandemic. Additionally, disaggregated growth trajectories that account for EL status and English language proficiencies may aide in determining differential supports for supporting subgroups of students with limited English proficiency. Perhaps most important is the finding that ELs do not represent a homogenous group whose needs can be met without differentiation. In fact, EL students with more advanced levels of English language proficiency appeared to benefit more substantially from the instructional supports they received during first and second grade than did students with lower levels of language proficiency. Consideration of the supports implemented for EL students and

continued monitoring of progress by subgroups may aid in determining methods for effectively scaffolding instructional supports for students who may be at increased risk for demonstrating difficulty relative to their peers. Former-ELs, or students who were formerly identified as ELs and have since been exited from EL programs, should also be monitored to better understand how this group differs from the general student population and their EL peers and how COVID-19 disruptions impacted their learning (Hopkins, Thompson, Linquanti, Hakuta, & August, 2013).

Finally, future research should continue to investigate differences in student learning that may be associated with the COVID-19 pandemic across grades K-12 and for subgroups of students that are at increased risk for performing lower than their peers, such as ELs. The long-term impacts of COVID-19-related disruptions may continue to change as school systems adapt to the ever-changing needs of their students, and the adaptation of educational practices will continue to be a critical need as this generation of students progresses through their school career. Future research should examine differences in achievement pre- and post-pandemic students in other grades and for other academic constructs, including but not limited to other reading-related skills (e.g., Oral Reading Fluency).

APPENDIX A

Observed Patterns of Missingness for DIBELS 8 Benchmarking Data

Included Patterns	CLS-WRC-WRF	*n*	Excluded Patterns	CLS-WRC-WRF	*n*
Five Timepoints	XXXXX-XXXXX-XXXXX	2,683	Missing 2019-20	OOOOX-OOOOX-OOOOX	39
Four Timepoints	OXXXX-OXXXX-OXXXX	80		OOOXO-OOOXO-OOOXO	11
	XOXXX-XOXXX-XOXXX	54		OOOXX-OOOXX-OOOXX	48
	XXOXX-XXOXX-XXOXX	129		OOXOO-OOXOO-OOXOO	19
	XXXOX-XXXOX-XXXOX	187		OOXOX-OOXOX-OOXOX	13
	XXXXO-XXXXO-XXXXO	427		OOXXO-OOXXO-OOXXO	23
Three Timepoints	OXOXX-OXOXX-OXOXX	9		OOXXX-OOXXX-OOXXX	127
	OXXOX-OXXOX-OXXOX	6	Missing 2020-21	XXOOO-XXOOO-XXOOO	648
	OXXXO-OXXXO-OXXXO	17		XOOOO-XOOOO-XOOOO	202
	XOOXX-XOOXX-XOOXX	10		OXOOO-OXOOO-OXOOO	60
	XOXOX-XOXOX-XOXOX	11	No Benchmarking Data	OOOOO-OOOOO-OOOOO	18
	XOXXO-XOXXO-XOXXO	19	Partial Benchmarking Data	XXOOO-XXOOO-XXOXO	1
	XXOOX-XXOOX-XXOOX	70		XXOOX-XXOOX-XXOXX	1
	XXOXO-XXOXO-XXOXO	88		XXXXO-XXXXO-XXXXX	1
	XXXOO-XXXOO-XXXOO	305		XXXXX-XXXXX-XXXOX	1
Two Timepoints	OXOXO-OXOXO-OXOXO	8	**TOTAL**		**1,212**
	OXXOO-OXXOO-OXXOO	15			
	XOOOX-XOOOX-XOOOX	7			
	XOOXO-XOOXO-XOOXO	8			
	XOXOO-XOXOO-XOXOO	16			
TOTAL		**4,149**			

Note. X = data available. O = missing data.

APPENDIX B

Q-Q Plots Illustrating Score Distributions for All DIBELS 8 Measures Examined

Figure B.1

Q-Q Plot for Nonsense Word Fluency - Correct Letter Sounds (NWF-CLS)

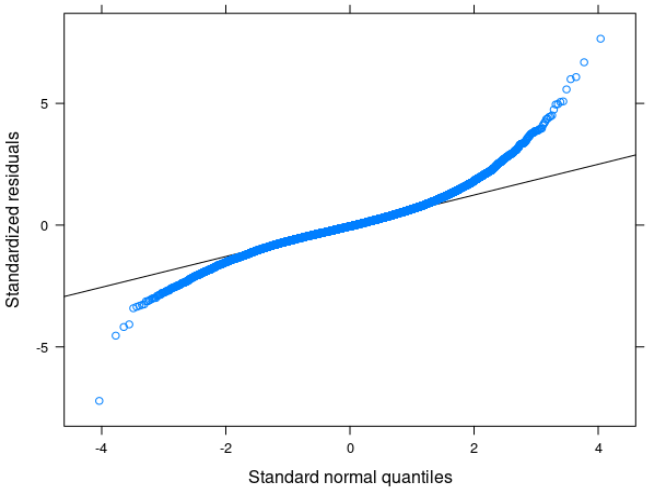


Figure B.2

Q-Q Plot for Nonsense Word Fluency - Words Read Correct (NWF-WRC)

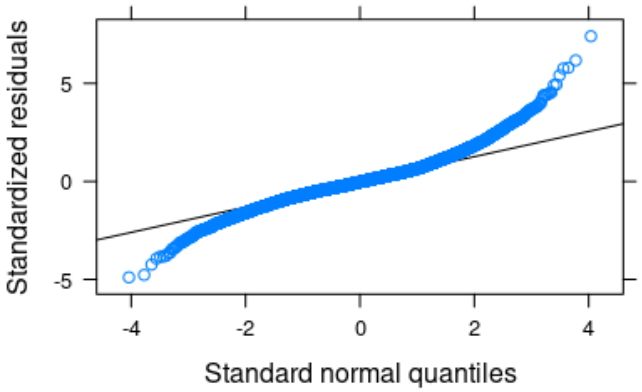
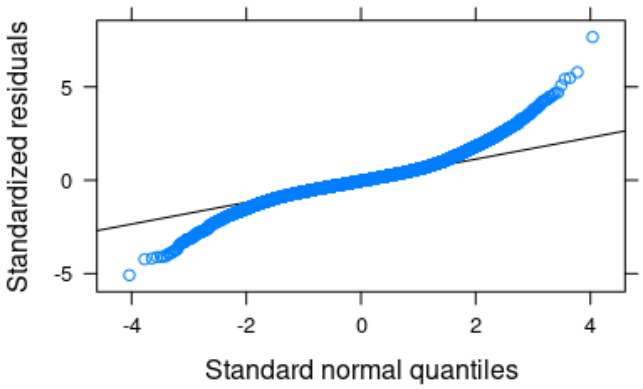


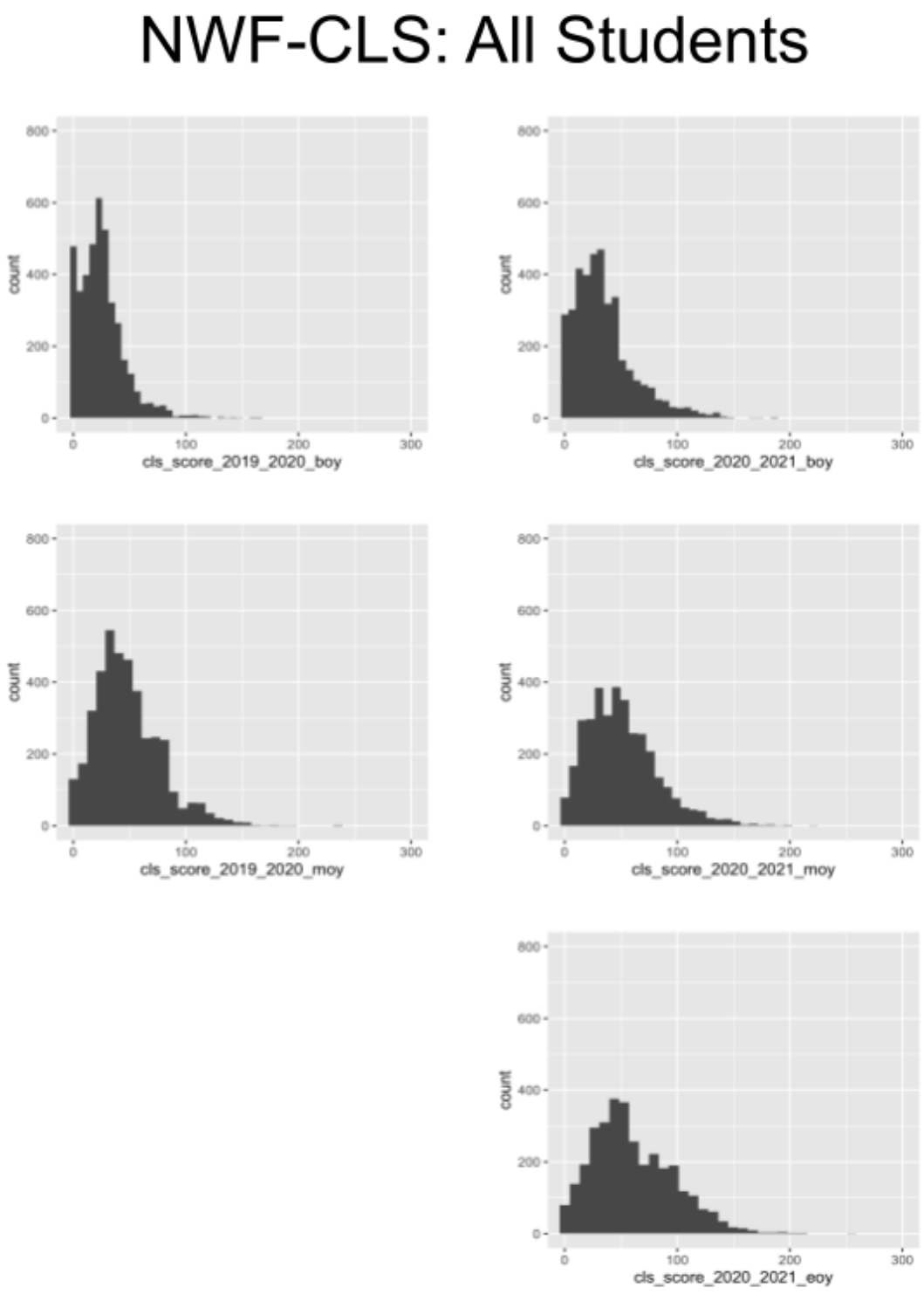
Figure B.3

Q-Q Plot for Word Reading Fluency

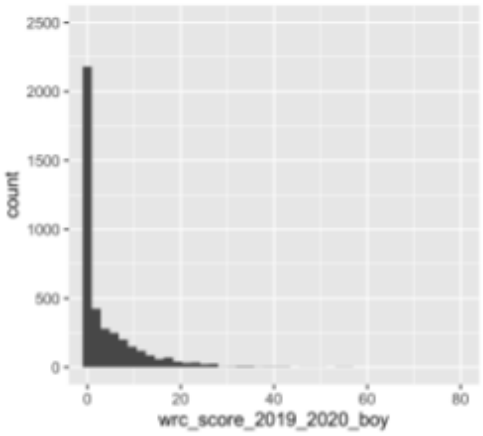


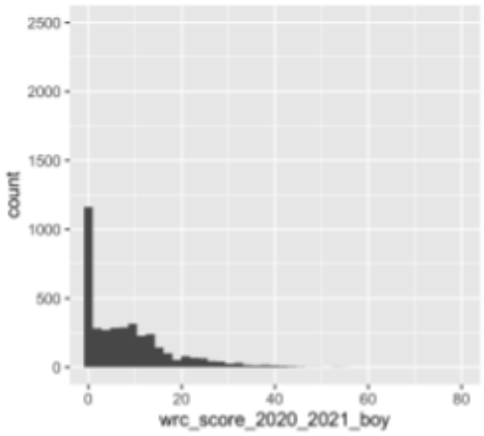
APPENDIX C

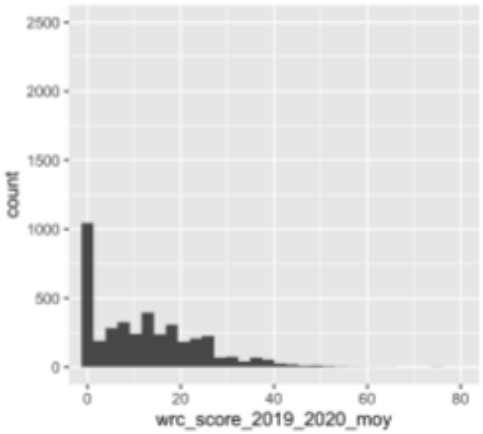
Histograms Illustrating Distribution of Scores on DIBELS 8 Measures Across Timepoints

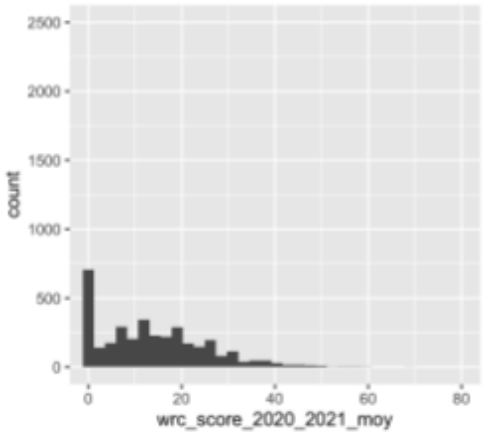


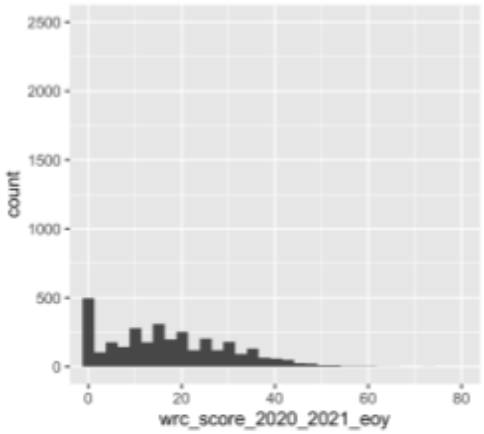
NWF-WRC: All Students



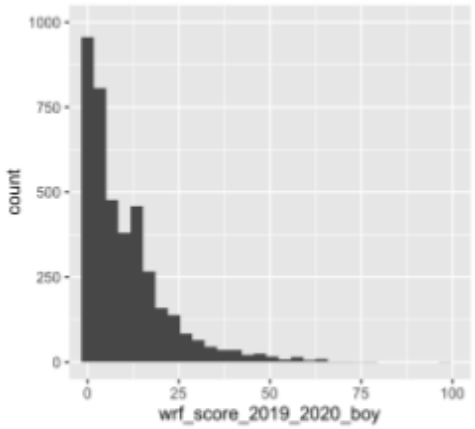


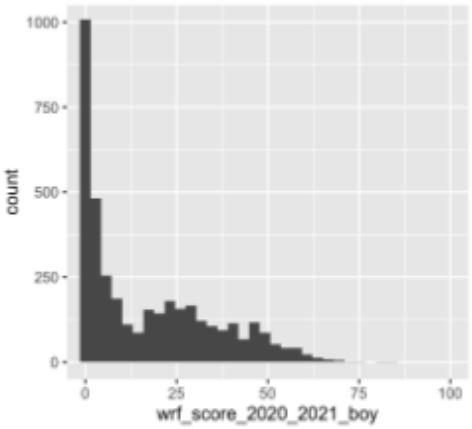


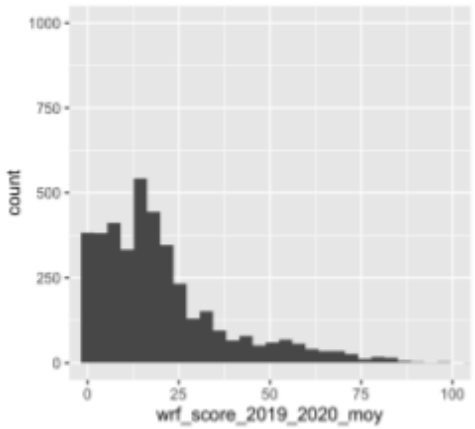


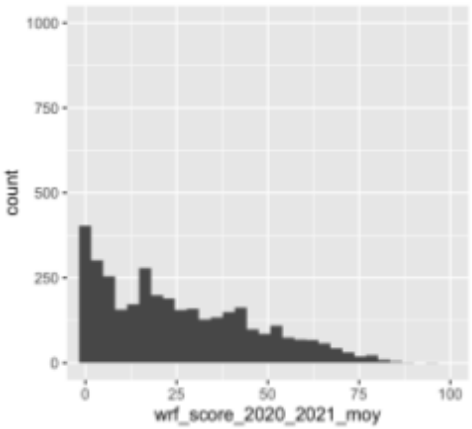


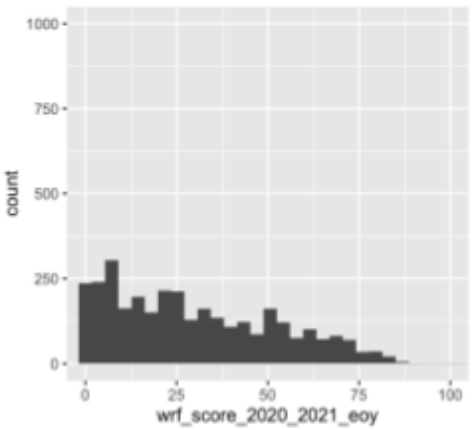
WRF: All Students



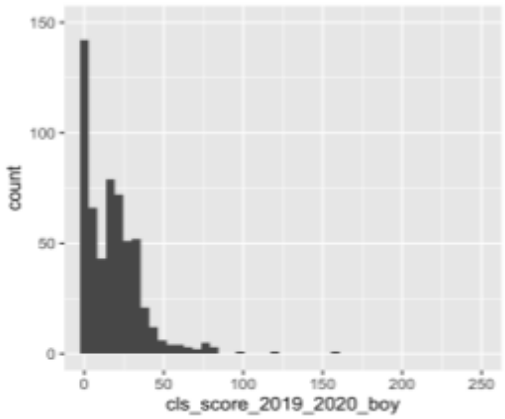


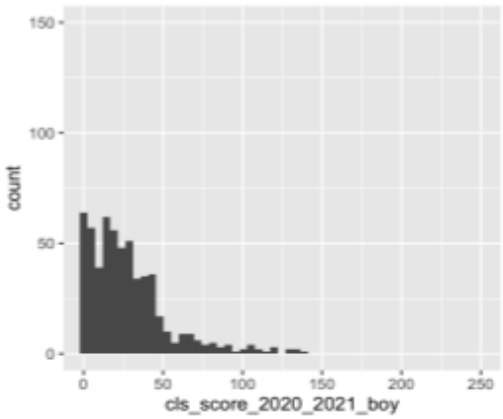


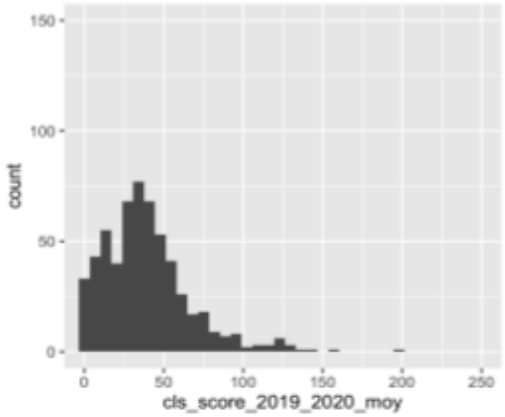


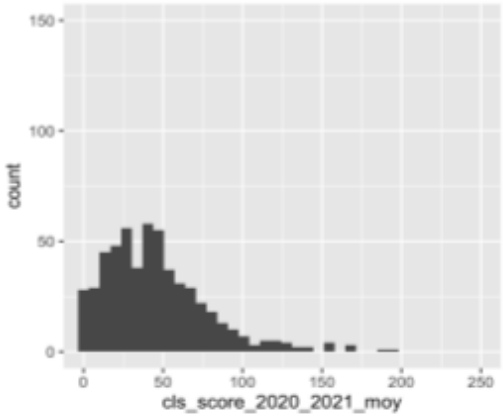


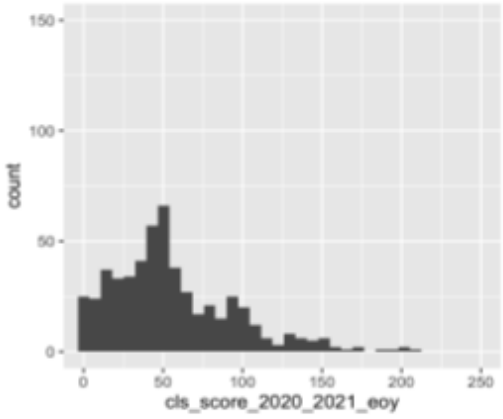
NWF-CLS: ELs


count
cls_score_2019_2020_boy

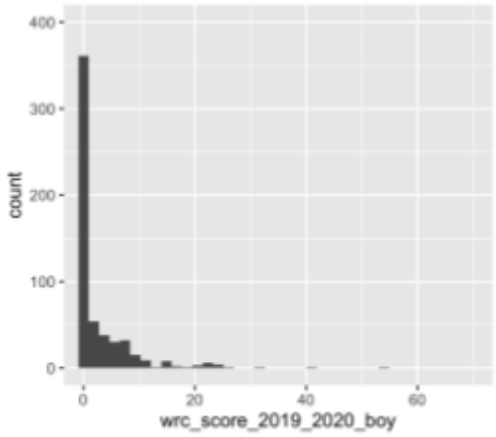

count
cls_score_2020_2021_boy

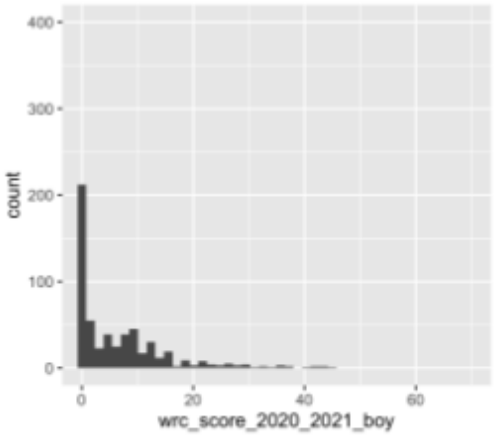

count
cls_score_2019_2020_moy

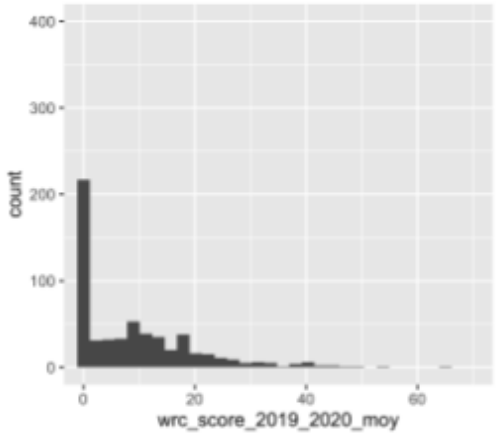

count
cls_score_2020_2021_moy

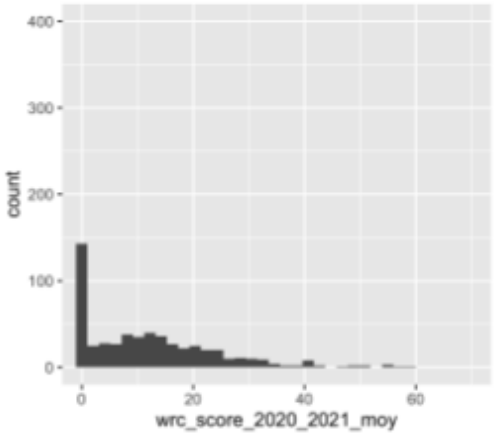

count
cls_score_2020_2021_eoy

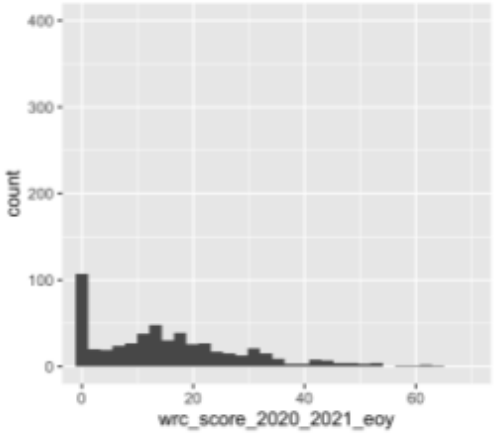
NWF-WRC: ELs


count
400
300
200
100
0
0
20
40
60
wrc_score_2019_2020_boy

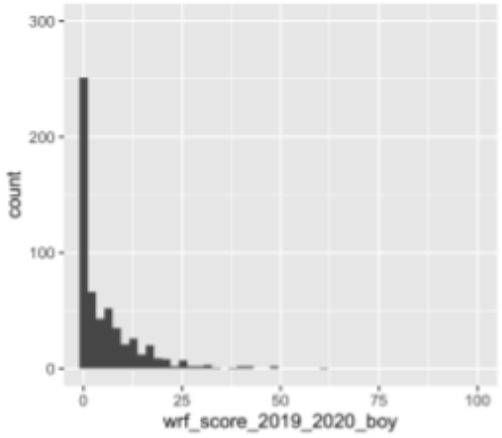

count
400
300
200
100
0
0
20
40
60
wrc_score_2020_2021_boy

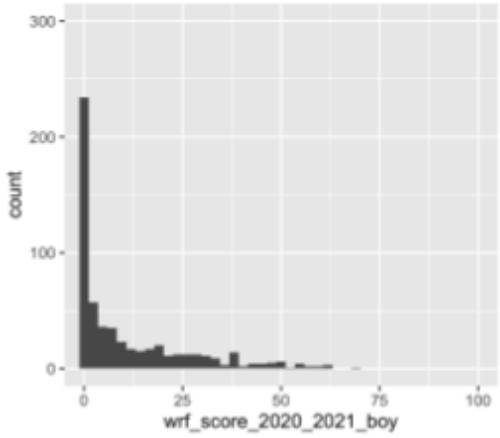

count
400
300
200
100
0
0
20
40
60
wrc_score_2019_2020_moy

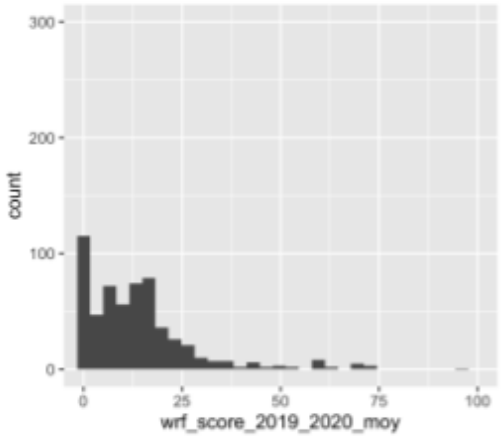

count
400
300
200
100
0
0
20
40
60
wrc_score_2020_2021_moy

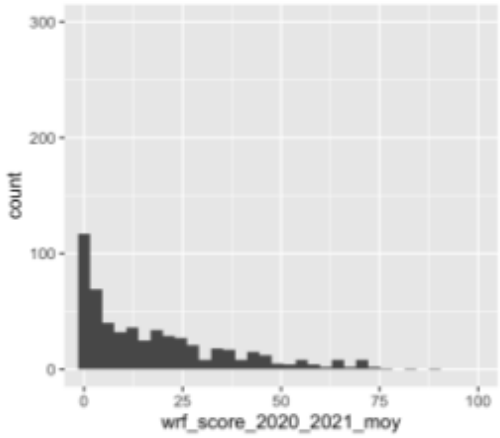

count
400
300
200
100
0
0
20
40
60
wrc_score_2020_2021_eoy

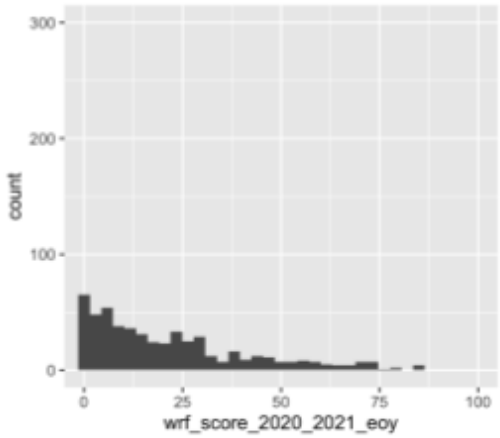
WRF: ELs


count
300
200
100
0
0
25
50
75
100
wrf_score_2019_2020_boy

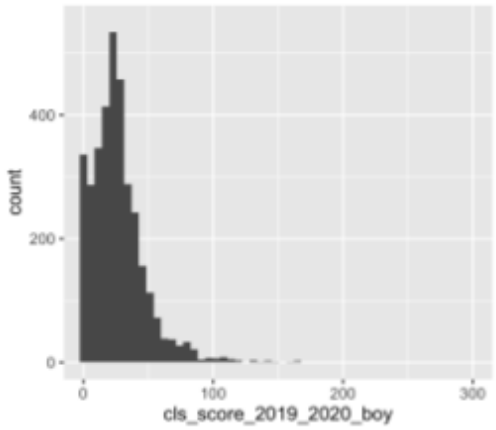

count
300
200
100
0
0
25
50
75
100
wrf_score_2020_2021_boy

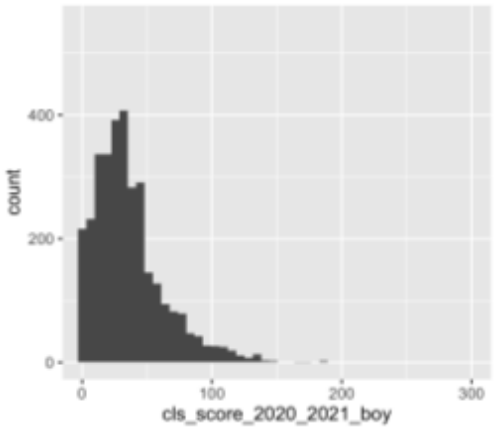

count
300
200
100
0
0
25
50
75
100
wrf_score_2019_2020_moy

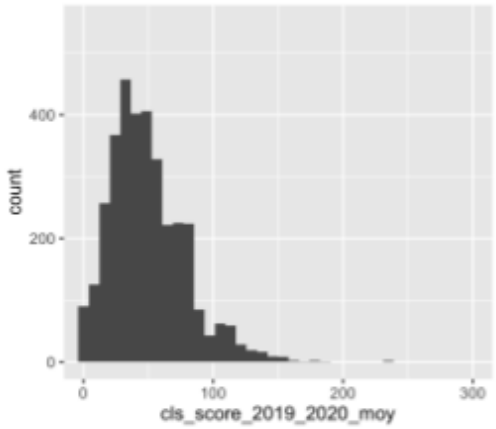

count
300
200
100
0
0
25
50
75
100
wrf_score_2020_2021_moy

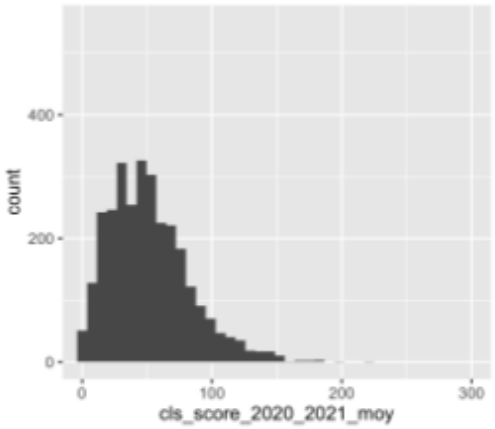

count
300
200
100
0
0
25
50
75
100
wrf_score_2020_2021_eoy

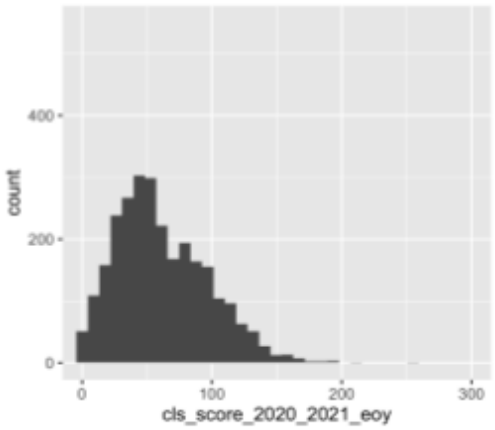
NWF-CLS: Non-ELs



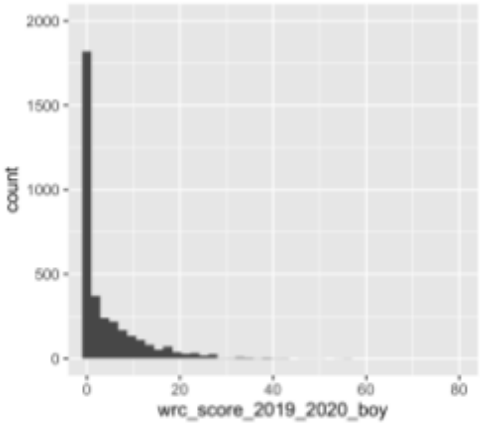


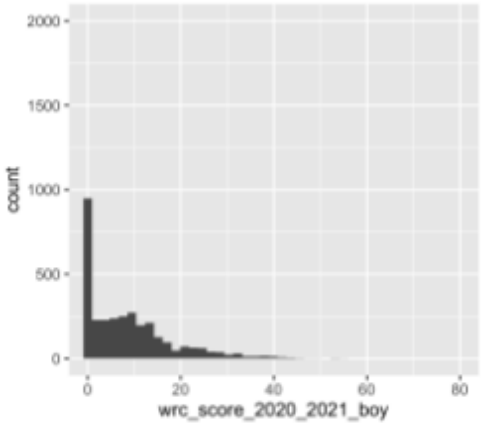


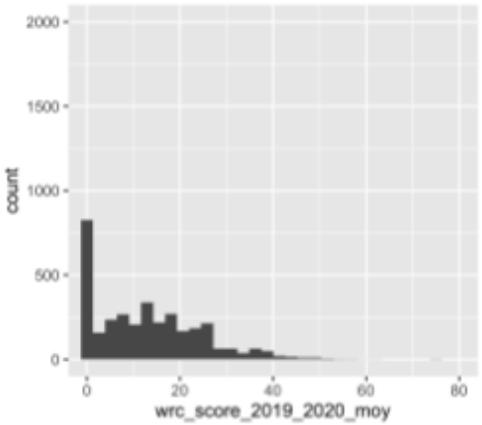


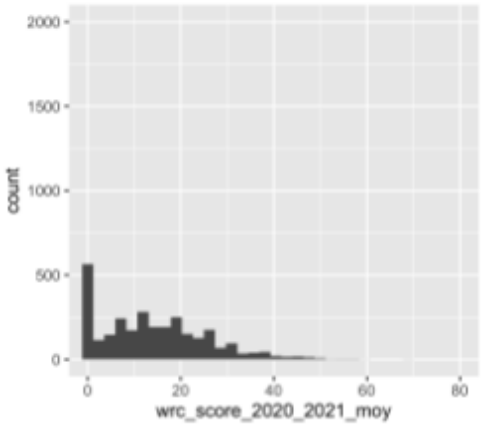


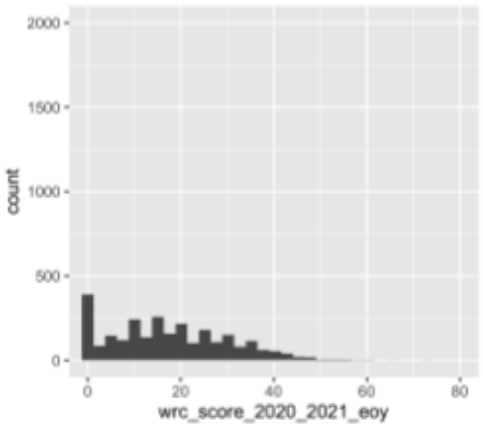
NWF-WRC: Non-ELs



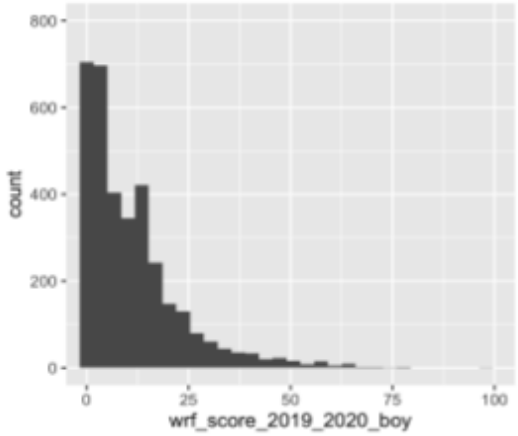


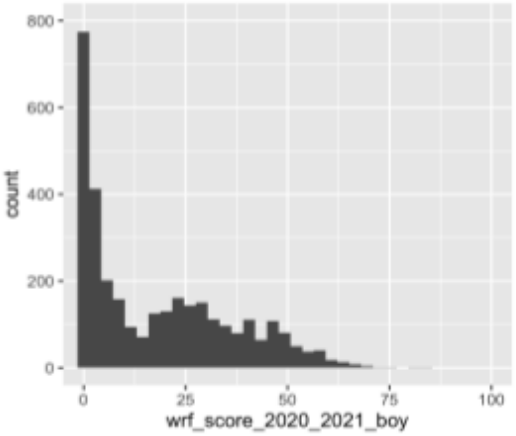


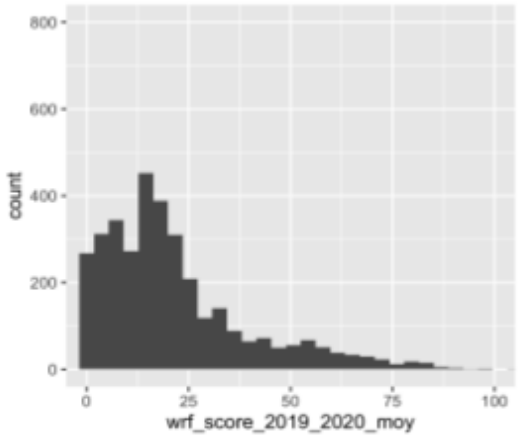


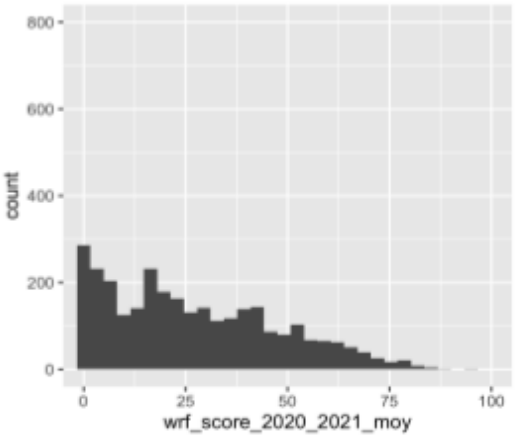


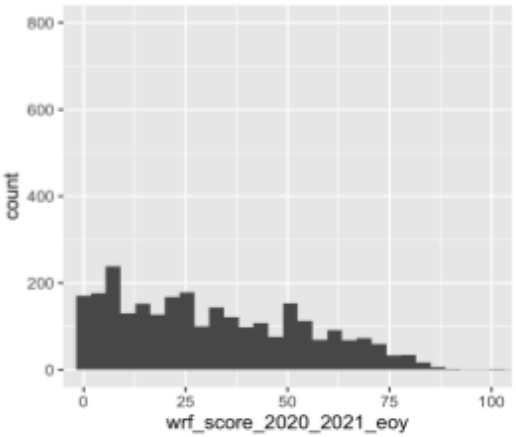
WRF: Non-ELs


count
800
600
400
200
0
0
25
50
75
100
wrf_score_2019_2020_boy


count
800
600
400
200
0
0
25
50
75
100
wrf_score_2020_2021_boy


count
800
600
400
200
0
0
25
50
75
100
wrf_score_2019_2020_moy


count
800
600
400
200
0
0
25
50
75
100
wrf_score_2020_2021_moy


count
800
600
400
200
0
0
25
50
75
100
wrf_score_2020_2021_eoy

APPENDIX D

Box Plots for All Measures by English Learner (EL) Student Status

Figure D.1

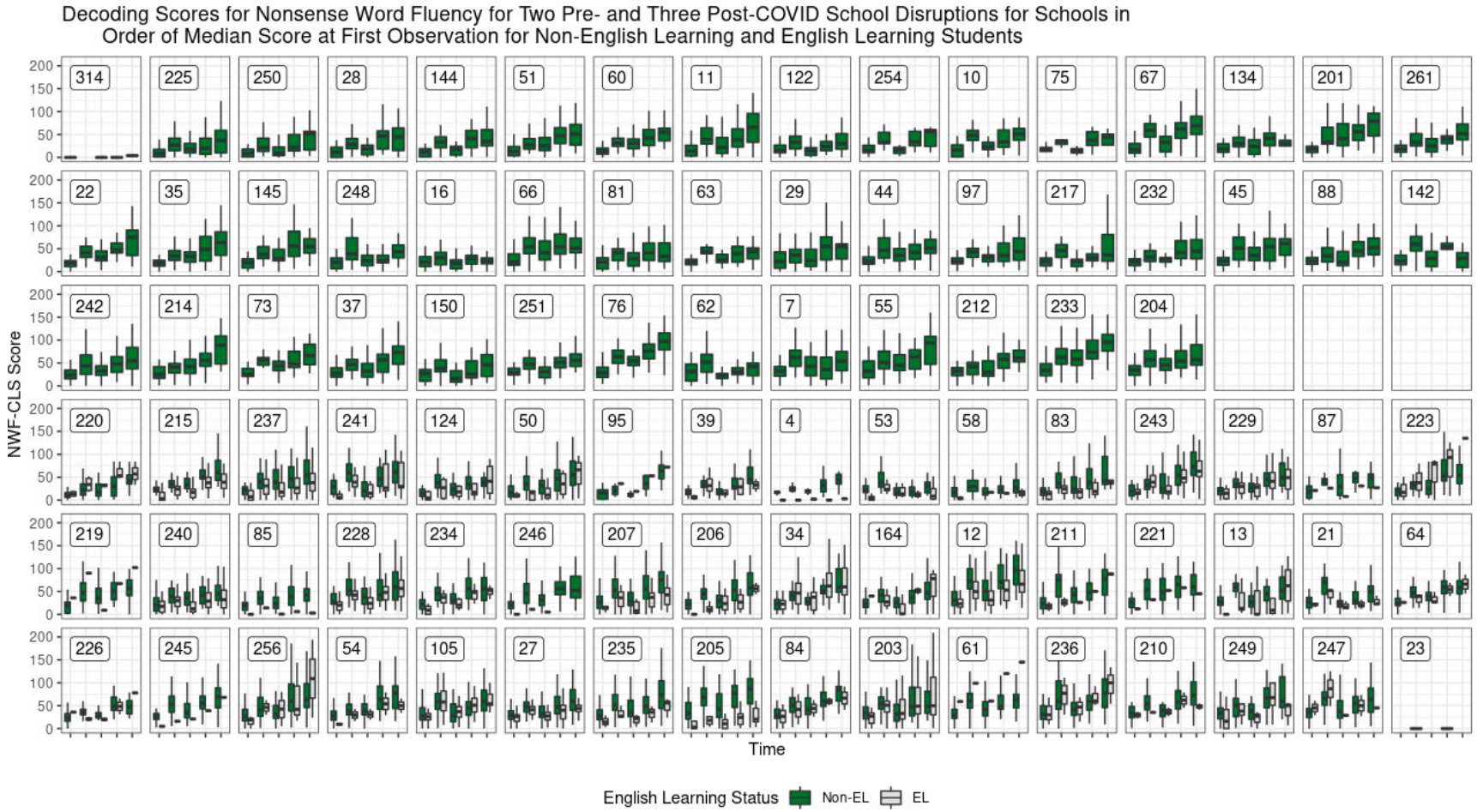


Figure D.2

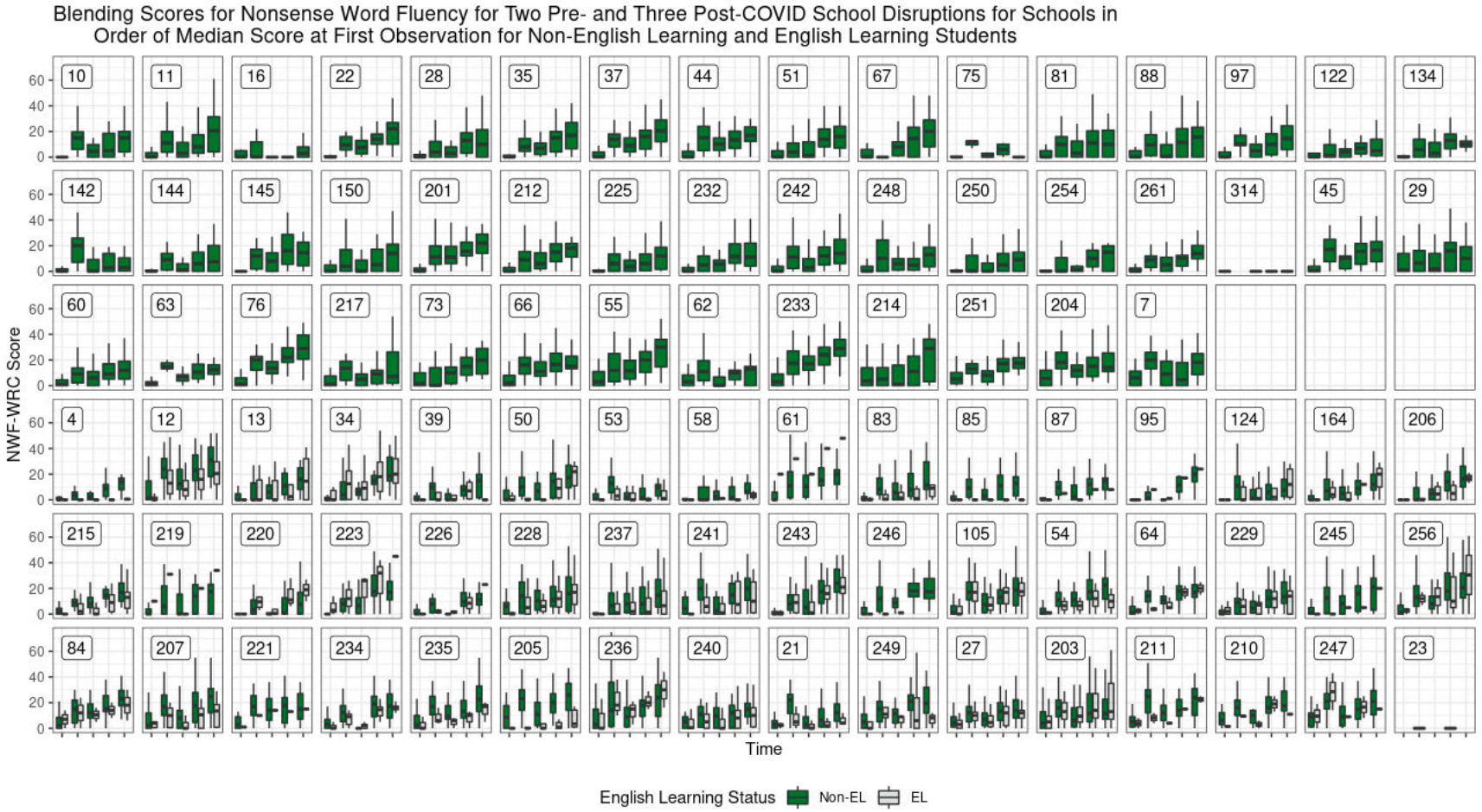
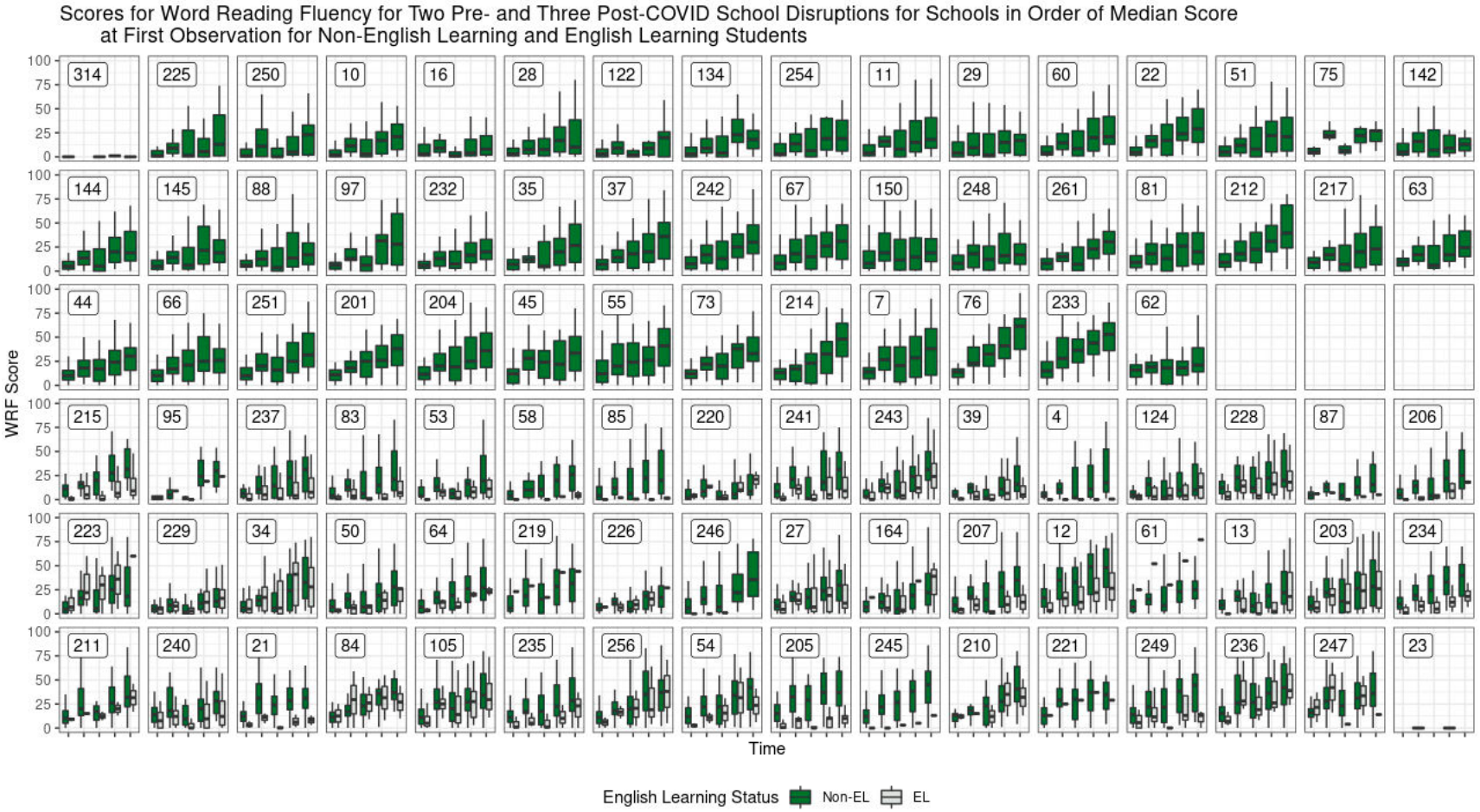

Blending Scores for Nonsense Word Fluency for Two Pre- and Three Post-COVID School Disruptions for Schools in Order of Median Score at First Observation for Non-English Learning and English Learning Students
NWF-WRC Score
Time
English Learning Status
Non-EL
EL

Figure D.3


Scores for Word Reading Fluency for Two Pre- and Three Post-COVID School Disruptions for Schools in Order of Median Score at First Observation for Non-English Learning and English Learning Students
WRF Score
Time
English Learning Status
Non-EL
EL

APPENDIX E

Residual Plots for Piecewise Hierarchical Base Models

Figure E.1

Residual Plot for NWF-CLS - All Students

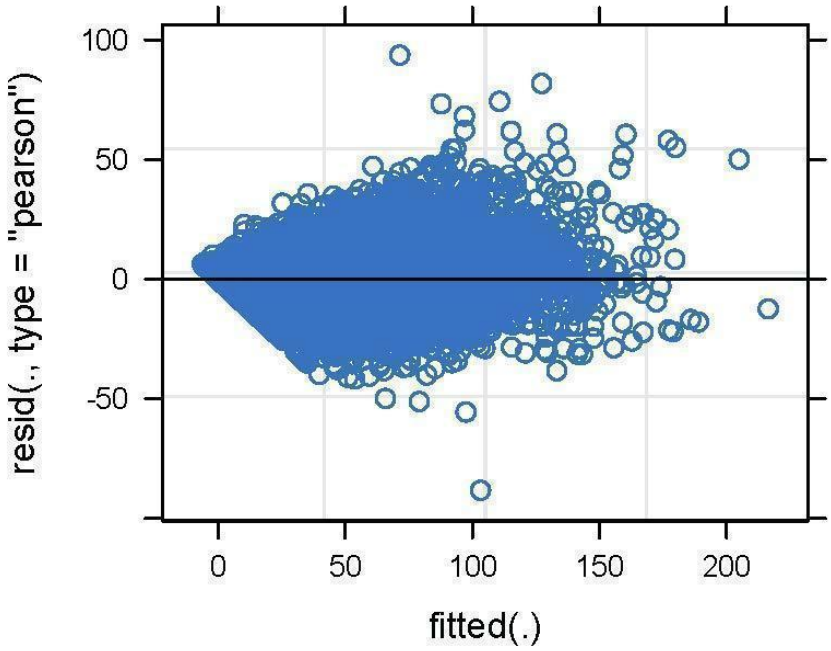


Figure E.2

Residual Plot for NWF-CLS - EL Students

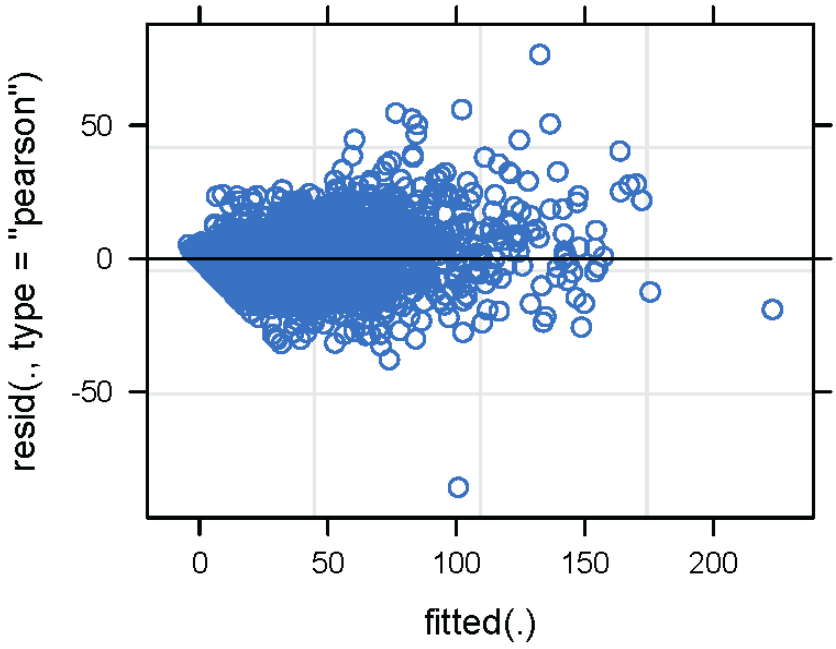


Figure E.3

Residual Plot for NWF-WRC - All Students

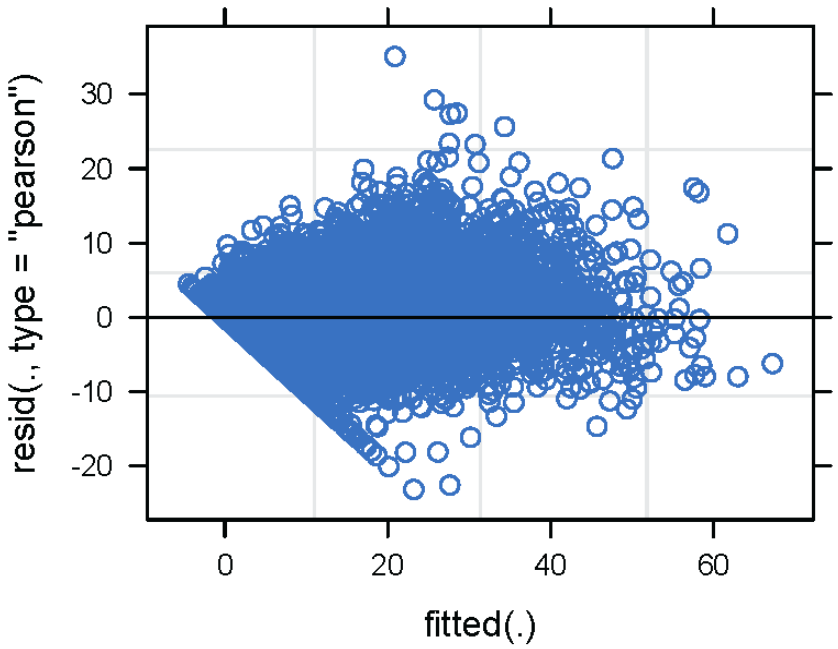


Figure E.4

Residual Plot for NWF-WRC - EL Students

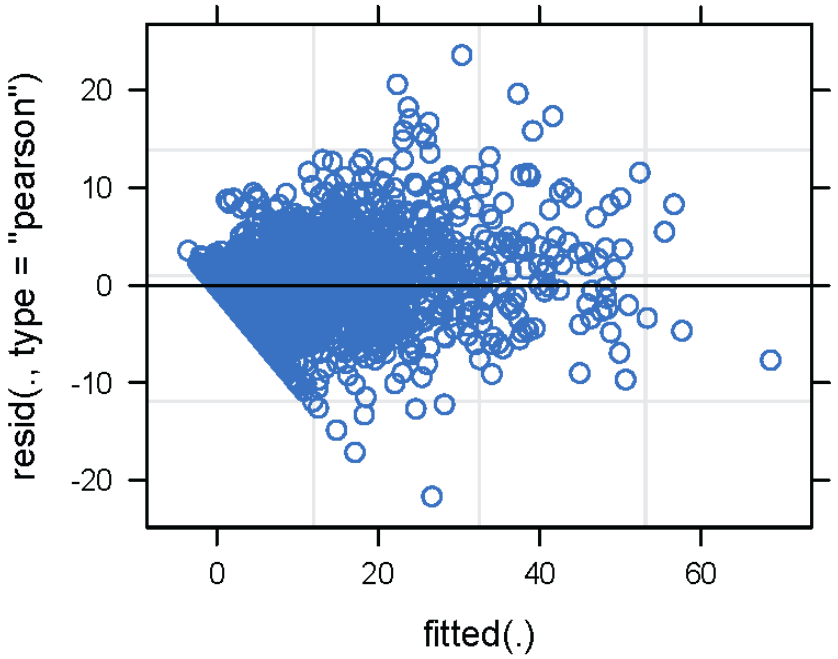


Figure E.5

Residual Plot for WRF - All Students

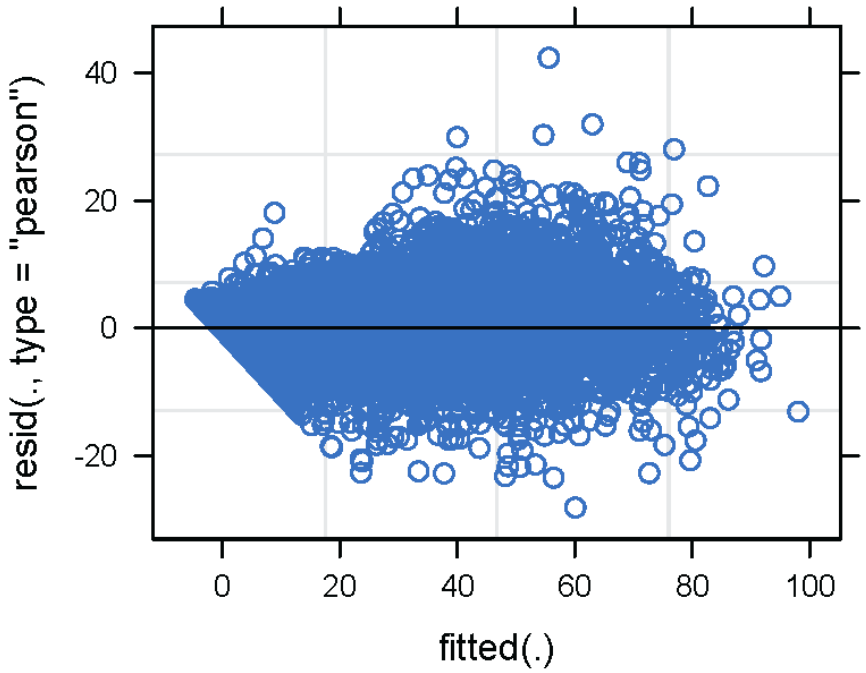
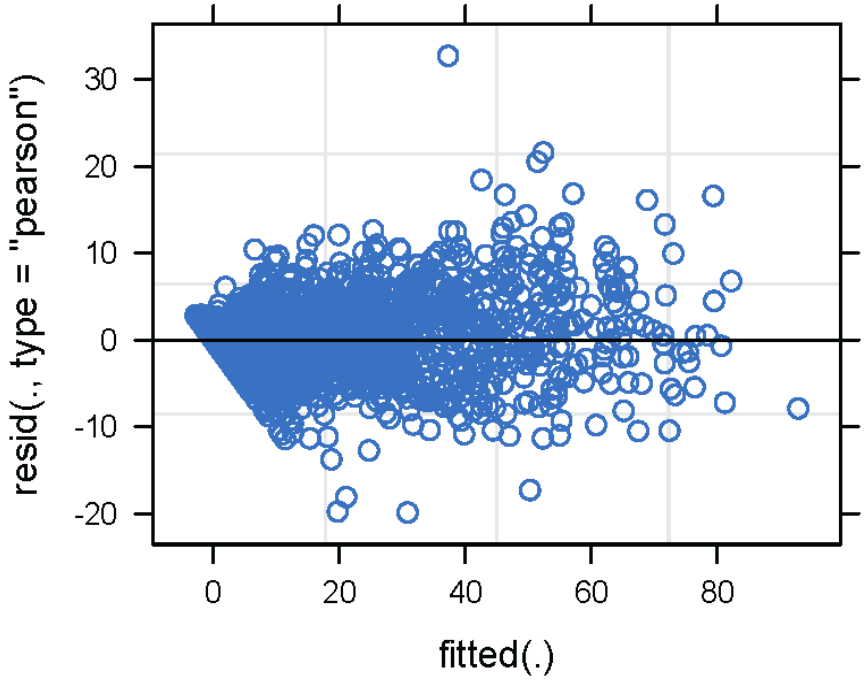


Figure E.6

Residual Plot for WRF - EL Students



APPENDIX F

EL Student Base Model Fixed Effects

	NWF-CLS			NWF-WRC			WRF		
Fixed Effects	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*	Coefficient (*SE*)	*t*	*g*
Intercept	16.01 (1.44)	11.10***	0.02	2.36 (0.46)	5.12***	0.03	5.12 (0.69)	7.47***	0.03
Time_Instruction	20.43 (1.16)	17.61***	0.77	6.31 (0.47)	13.55***	0.66	8.07 (0.52)	15.42***	0.63
COVID_Summer	-9.93 (1.07)	-9.24***	-0.37	-2.33 (0.39)	-6.02***	-0.24	-2.88 (0.46)	-6.19***	-0.23
COVID_Instruction	-4.02 (1.25)	-3.22**	-0.15	-0.82 (0.58)	-1.41	-0.09	-0.76 (0.50)	-1.52	-0.06
Spring21	-9.52 (1.05)	-9.08***	-0.36	-2.55 (0.39)	-6.51***	-0.27	-3.63 (0.41)	-8.89***	-0.28

Note. NWF-CLS = Nonsense Word Fluency - Correct Letter Sounds. NWF-WRC = Nonsense Word Fluency - Words Read Correct. WRF = Word Reading Fluency.
*** $p < .001$, ** $p < .01$

REFERENCES CITED

Amplify Education, Inc. (2021a). *Research brief: COVID-19 means more students not learning to read.* https://amplify.com/wp-content/uploads/2021/02/Amplify-mCLASS_MOY-COVID-Learning-Loss-Research-Brief_022421.pdf

Amplify Education, Inc. (2021b). *Research brief: Schools start to recover from COVID instructional loss.* https://amplify.com/wp-content/uploads/2021/07/mCLASS_EOY-COVID-Learning-Loss-Brief_Flyer_072821.pdf

Atteberry, A., & McEachin, A. (2021). School's out: The role of summers in understanding achievement disparities. *American Educational Research Journal, 58*(2), 239-282. https://doi.org/10.3102/0002831220937285

Bailey, D. H., Duncan, G. J., Murnane, R. J., & Yeung, N. A. (2021). Achievement gaps in the wake of COVID-19. *Educational Researcher,* 1-10. https://doi.org/10.3102/0013189X211011237

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *ArXiv Preprint ArXiv:1406.5823*. https://www.jstatsoft.org/index.php/jss/article/view/v067i01/946

Bielinksi, J., Brown, R., & Wagner, K. (2021). No longer a prediction: What new data tell us about the effects of 2020 learning disruptions [White Paper]. Illuminate Education. https://www.illuminateed.com/wp-content/uploads/2021/06/LearningLoss_2021_Update_FB_031221.pdf

Brenan, M. (2020, April 8). *Over 8 in 10 parents now say child is learning remotely*. Gallup. https://news.gallup.com/poll/307754/parents-say-child-learning-remotely.aspx

Catalano, A. J., Torff, B., & Anderson, K. S. (2021). *The International Journal of Information and Learning Technology, 38*(2), 258-270. https://doi.org/10.1108/IJILT-06-2020-0111

Domingue, B. W., Hough, H. J., Lang, D., & Yeatman, J. (2021). *Changing patterns of growth in oral reading fluency during the COVID-19 pandemic* (PACE Working Paper). Policy Analysis for California Education. https://edpolicyinca.org/publications/changing-patterns-growth-oral-reading-fluency-during-covid-19-pandemic

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review, 69*(5), 613-635. https://doi.org/10.1177/000312240406900501

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinsein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446. http://dx.doi.org/10.1037/0012-1649.43.6.1428

Every Student Succeeds Act, 20 U.S.C. § 6301 (2015). https://www.congress.gov/bill/114th-congress/senate-bill/1177

Fien, H., Baker, S. K., Smolkowski, K., Smith, J. L. M., Kame'enui, E. J., & Beck, C. T. (2008). Using nonsense word fluency to predict reading proficiency in kindergarten through second grade for English learners and native English speakers. *School Psychology Review, 37*(3), 391-408. https://doi.org/10.1080/02796015.2008.12087885

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3–17. https://psycnet.apa.org/doi/10.1037/0022-0663.88.1.3

Furjanic, D., Ives, C., Fainstein, D., Kennedy, P. C., & Biancarosa, G. (revision in progress). COVID-19 and Oral Reading Fluency.

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide. (NCEE 2009-4045). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://www.trumbullesc.org/Downloads/rti_reading_pg_021809.pdf

Herbers, J. E., Cutuli, J. J., Supkoff, L. M., Heistad, D., Chan, C.-K., Hinz, E., & Masten, A. S. (2012). Early Reading Skills and Academic Achievement Trajectories of Students Facing Poverty, Homelessness, and High Residential Mobility. *Educational Researcher, 41*(9), 366–374. https://doi.org/10.3102/0013189X12445320

Hernandez, J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Baltimore, MD: Annie E. Case Foundation. https://www.aecf.org/m/resourcedoc/AECF-DoubleJeopardy-2012-Full.pdf

Hopkins, M., Thompson, K. D., Linquanti, R., Hakuta, K., & August, D. (2013). Fully accounting for English learner performance: A key issue in ESEA reauthorization. *Educational Researcher, 42*(2), 101-108. https://doi.org/10.3102%2F0013189X12471426

January, S. A., & Klingbeil, D. A. (2020). Universal screening in grades K-2: A systematic review and meta-analysis of early reading curriculum-based measures. *Journal of School Psychology, 82*, 103-122. https://doi.org/10.1016/j.jsp.2020.08.007

Johnson, A., & Kuhfeld, M. (2020). Fall 2019 to fall 2020 MAP growth attrition analysis. *NWEA*. https://www.nwea.org/content/uploads/2020/11/Technical-brief-Fall-2019-to-fall-2020-MAP-Growth-attrition-analysis-NOV2020.pdf

Kaminski, R. A., & Good, R. H. III. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*(2), 215-227. https://doi.org/10.1080/02796015.1996.12085812

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, *100*(4), 851. https://psycnet.apa.org/doi/10.1037/0022-0663.100.4.851

Kieffer, M. J. (2011). Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to grade 8. *American Educational Research Journal*, *48*(5), 1187-1225. https://doi.org/10.3102%2F0002831211419490

Kieffer, M. J., & Parker, C. E. (2016). Patterns of English learner student reclassification in New York City Public Schools. REL 2017-200. *Regional Educational Laboratory Northeast & Islands*. https://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2017200.pdf

Kieffer, M. J., & Thompson, K. D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher, 47*(6), 391-398. https://doi.org/10.3102/0013189X18777740

Kogan, V., & Lavertu, S. (2022). *How the COVID-19 pandemic affected student learning in Ohio: Analysis of Spring 2021 Ohio state tests.* The Ohio State University. http://glenn.osu.edu/educational-governance/reports/reports-attributes/210828_KL_OST_Final.pdf

König, C., & Frey, A. (2022). The impact of COVID-19-related school closures on student achievement – A meta-analysis. *Educational Measurement: Issues and Practice, 41*(1), 16-22. https://doi.org/10.1111/emip.12495

Kraft, M. A., & Simon, N. S. (2020). *Teachers' experiences working from home during the COVID-19 pandemic.* Upbeat. Downloaded May 3, 2021 from https://f.hubspotusercontent20.net/hubfs/2914128/Upbeat%20Memo_Teaching_From_Home_Survey_June_24_2020.pdf

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, Jing. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher, 49*(8), 549-565. https://doi.org/10.3102/0013189x20965918

Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020, November). *Learning during COVID-19: Initial findings on students' reading and math achievement and growth.* NWEA Research. https://www.ewa.org/sites/main/files/file-attachments/learning_during_covid-19_brief_nwea_nov2020_final.pdf?1606835922

Lesnick, J., George, R., Smithgall, C., & Gwynne, J. (2010). *Reading on grade level in third grade: How it is related to high school performance and college enrollment.* Chicago, IL: Chapin Hall at the University of Chicago. https://www.chapinhall.org/wp-content/uploads/Reading_on_Grade_Level_111710.pdf

Malkus, N. (2020, June 16). School districts' remote learning plans may widen student achievement gap. *Education Next.* https://www.educationnext.org/school-districts-remote-learning-plans-may-widen-student-achievement-gap-only-20-percent-meet-standards/

Map: Coronavirus and school closures. (2021, May 12). Education Week. Retrieved May 12, 2021 from https://www.edweek.org/leadership/map-where-are-schools-closed/2020/07

Map: Coronavirus and school closures. (2020, March 6). Education Week. Retrieved May 6, 2021 from https://www.edweek.org/leadership/map-coronavirus-and-school-closures-in-2019-2020/2020/03

McCoach, D. B., & Kaniskan, B. (2020). Using time-varying covariates in multilevel growth models. *Frontiers in Psychology, 1*(17). https://doi.org/10.3389/fpsyg.2010.00017

McIntosh, K., Girvan, E. J., Horner, R. H., & Smolkowski, K. (2014). Education not incarceration: A conceptual model for reducing racial and ethnic disproportionality in school discipline. *Journal of Applied Research on Children: Informing Policy for Children at Risk, 5*(2). http://digitalcommons.library.tmc.edu/childrenatrisk/vol5/iss2/4

Morgan, L., Parkas, G., & Hibel, J. (2008). Matthew Effects for Whom? *Learning Disability Quarterly, 31*(4), 187–198. https://doi.org/10.2307%2F25474651

Muyskens, P., Betts, J., Lau, M. Y., & Marston, D. (2009). Predictive validity of curriculum-based measures in the reading assessment of students who are English language learners. *The California School Psychologist, 14*(1), 11-21. https://doi.org/10.1007/BF03340947

National Center for Education Statistics. (2020). *Table 221.12: Average National Assessment of Educational Progress (NAEP) reading scale score and percentage of students attaining NAEP reading achievement levels, by selected school and student characteristics and grade: Selected years, 1992 through 2019.* In U.S. Department of Education, Institute of Education Science (Ed.), *Digest of Education Statistics: 2020*. Retrieved from https://nces.ed.gov/programs/digest/d20/tables/dt20_221.12.asp

National Center for Education Statistics. (2019). *National Assessment of Education Progress (NAEP) Report Card: Reading.* U.S. Department of Education. https://www.nationsreportcard.gov/reading/nation/achievement/?grade=4

National Reading Panel. (2000). *Report of the National Reading Panel--Teaching Children to Read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, D.C.: National Institute of Child Health and Human Development. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Nese, J., Biancarosa, G., Cummings, K., Kennedy, P., Alonzo, J., & Tindal, G. (2013). In search of average growth: Describing within-year oral reading fluency growth across Grades 1-8. *Journal of School Psychology, 51*(5), 625-642. https://doi.org/10.1016/j.jsp.2013.05.006

NWEA. (2019). MAP Growth technical report. NWEA Research Report. https://www.nwea.org/content/uploads/2021/11/MAP-Growth-Technical-Report-2019_NWEA.pdf

Office of English Learning Acquisition. (2021). *Profile of English language learners in the United States.* U.S. Department of Education. https://ncela.ed.gov/sites/default/files/fast_facts/DEL4.4_ELProfile_508_1.4.2021_OELA.pdf

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the Black-White test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 229-272). Brookings Institution Press.

Pier, L., Christian, M., Tymeson, H., & Meyer, R. H. (2021). *COVID-19 impacts on student learning: Evidence from interim assessments in California* [Report]. Policy Analysis for California Education. https://edpolicyinca.org/publications/covid-19-impacts-student-learning

Porter, T., Cook, H. G., & Sahakyan, N. (2019). *Less than four domains: Creating an overall composite score for English learners with individualized education plans.* (WIDA Working Paper No. 2019-3). https://wcer.wisc.edu/docs/working-papers/Working_Paper_No_2019_3.pdf

R Core Team. (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing.* https://doi.org/10.1007/978-3-540-74686-7

Ragan, A., & Lesaux, N. (2006). Federal, state, and district level English language learner program entry and exit requirements: Effects on the education of language minority learners. *Education Policy Analysis Archives, 14*(20). https://www.redalyc.org/pdf/2750/275020543020.pdf

Rapaport, A., Saavedra, A., Silver, D., & Polikoff, M. (2020). *Surveys show things are better for students than they were in the spring: Or do they?* Brown Center Chalkboard, Brookings Institution. https://www.brookings.edu/blog/brown-center-chalkboard/2020/11/18/surveys-show-things-are-better-for-students-than-they-were-in-the-spring-or-do-they/

Richardson, R. D., Rocconi, L. M., & Crewdson, M. A. (2020). Evaluating English learner progress in reading: How much growth can we expect?. *School Psychology Review, 49*(4), 480-492. https://doi.org/10.1080/2372966X.2020.1787080

Sawchuk, S., & Samuels, C. A. (2020). *Where are they? Students go missing in shift to remote classes.* Education Week. https://www.edweek.org/leadership/where-are-they-students-go-missing-in-shift-to-remote-classes/2020/04

Slama, R. B. (2012). A longitudinal analysis of academic English proficiency outcomes for adolescent English language learners in the United States. *Journal of Educational Psychology, 104*(2), 265. https://doi.org/10.1037/a0025861

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology, 41*(1), 225-234. https://doi.org/10.1037/0012-1649.41.1.225

University of Oregon. (2020a). *8th Edition of Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Administration and scoring guide.* https://dibels.uoregon.edu/sites/dibels1.uoregon.edu/files/2021-05/dibels_8_admin_and_scoring_guide_05_2020.pdf

University of Oregon. (2020b). *8th Edition of Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Technical manual.* https://dibels.uoregon.edu/docs/techreports/DIBELS8-TechnicalManual_04152020.pdf

University of Oregon. (2020c). *DIBELS 8th Edition Benchmark Goals.* https://dibels.uoregon.edu/sites/dibels1.uoregon.edu/files/2021-06/DIBELS8thEditionGoals.pdf

U.S. Census Bureau. (2018). American community survey: Language spoken at home (Table ID S1601). https://data.census.gov/cedsci/table?q=Language%20Spoken%20at%20Home&t=Language%20Spoken%20at%20Home&tid=ACSST1Y2018.S1601&hidePreview=false

Vanderwood, M. L., Linklater, D., & Healy, K. (2008). Predictive accuracy of nonsense word fluency for English language learners. *School Psychology Review, 37*(1), 5-17. https://doi.org/10.1080/02796015.2008.12087904

von Hippel, P. T., Workman, J., & Downey, D. B. (2018). Inequality in reading and math skills forms mainly before kindergarten: A replication, and partial correction, of "Are Schools the Great Equalizer?". *Sociology of Education, 91*(4), 323-357. https://doi.org/10.1177/0038040718801760

von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. *Sociological Science, 6*, 43-80. https://sociologicalscience.com/download/vol-6/january/SocSci_v6_43to80.pdf

WIDA. (2020). *Annual technical report for ACCESS for ELLs paper English language proficiency test, series 403, 2018-2019 administration.* (WIDA Technical Report No. 15B). https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/ACCESS/Technical_Documents/2018-2019%20ACCESS-ATR-Series-403-Paper-Redacted.pdf

Wisconsin Center for Education Research. (n.d.). *WIDA consortium report: 2018-2019*. https://wida.wisc.edu/sites/default/files/Website/Memberships/Consortium/Annual-Reports/2019WIDAAnnualReport.pdf