



RESEARCH REPORT

Luke Waites Center for Dyslexia and Learning Disorders

Luke Waites Center Curriculum Efficacy Project

Report of Interim Findings

Summer 2022

Anna E. Middleton, Ph.D., CALT

Sheryl L. Frierson, M.D., M.Ed.

SCOTTISH RITE

FOR CHILDREN

®

Table of Contents

LWCCEP Interim Findings	1
Comparative Intervention Methods.....	2
<i>Take Flight: A Comprehensive Intervention for Students with Dyslexia</i>	2
<i>Bridges: A Dyslexia Intervention Connecting Teacher, Avatar, and Student</i>	2
Comparison Curricula.....	2
Measures.....	2
District Data.....	2
Study Specific Outcome Measures.....	3
Aggregate Student Sample Report	4
Participants.....	4
Statistical Analysis	5
Aggregate Treatment Effects	6
Comparative Growth in Reading and Related Language Skills.....	7
Comparative Growth in Fundamental Reading Skills.....	8
Comparative Growth in Word-Level Skills.....	9
Comparative Growth in Passage-Level Skills.....	10
Conclusions	11
References.....	12
Appendix	13

Tables

Table 1. Study-specific outcome measures	3
Table 2. Demographic characteristics of Aggregate Sample.	5
Table 3. Pre-test and mid-test scores on all study measures for Aggregate Sample.	7
Table A4. Demographic characteristics of Richardson ISD students	14
Table A5. Analysis of pre-test and mid-test scores using doubly multivariate ANCOVA.....	14
Table A6. Analysis of student gains using MANCOVA.....	15

Figures

Figure 1. Demographic characteristics for participating Richardson ISD students	Error! Bookmark not defined.
Figure 2. Growth in Oral and Written Language for Richardson ISD students	Error! Bookmark not defined.
Figure 3. Demographic characteristics of the Aggregate Sample.....	4
Figure 4. Growth in Oral and Written Language for the Aggregate Sample	6
Figure 5. Comparative growth in Fundamental Reading Skills	8
Figure 6. Comparative growth in Word-Level Skills.....	9
Figure 7. Comparative growth in Passage-Level Reading Skills	10

LWCCEP Interim Findings

Success and a Challenge

The Luke Waites Center for Dyslexia and Learning Disorders has previously developed and disseminated a successful dyslexia intervention called *Take Flight: A Comprehensive Intervention for Students with Dyslexia* which has been widely adopted in schools across the country since its release in 2006. *Take Flight* is designed to be taught by a certified teacher who has completed an additional two years of advanced dyslexia-specific training to become credentialed as a Certified Academic Language Therapist (CALT). The combined barriers of extensive investment of time and resources necessary for a teacher to attain CALT status and the limited number of qualified CALT training facilities place significant constraints on the number of CALTs available in schools. Yet, as a result of Texas state laws about dyslexia, an increasing number of public-school students are being identified as needing access to high quality dyslexia instruction; hence, the current demand for CALTs exceeds the number available in schools. The current project is designed to address this gap in resources.

A Promising Solution

In order to increase student access more rapidly to the beneficial *Take Flight* intervention, the Luke Waites Center developed a modified delivery model that is designed to be taught by a certified teacher. The certified teacher, without the advanced CALT training, can deliver the instruction with the help of expert curriculum support. This support includes the use of a 3-D anthropomorphic animated virtual co-teacher whose role is to introduce specific aspects of *Take Flight* content that typically require years of mentored training for a CALT to present with accuracy, consistency, and fidelity. For the technology to fulfill this role, it was necessary to design the virtual co-teacher with natural human-like mouth movement and facial expression. The accuracy of the visual aspects of speech, such as the correspondence of mouth movements to the auditory script is a critical component in language learning and may bolster auditory speech perception in children, particularly those with dyslexia (e.g., Navarra et al., 2012; Ziegler et al., 2009). Furthermore, direct instruction in articulatory phonetics can help students with dyslexia to develop the phonological processing skills needed for reading (Castiglioni-Spalten & Ehri, 2003). The primary goal of the *Bridges* program is to shorten the teacher training time, thereby making effective, evidence-based dyslexia intervention accessible to more children.

Expertise to develop a virtual co-teacher with the necessary anthropomorphic qualities comes from the University of Texas at Dallas Lab for Virtual Humans and Synthetic Societies, which has previously produced award winning 3-D anthropomorphic virtual humans for use in military and medical training settings. The dyslexia intervention using this virtual co-teaching avatar is now called *Bridges: A Dyslexia Intervention Connecting Teacher, Avatar, and Student (Bridges)*. Pilot introduction of the *Bridges* program in the controlled Luke Waites Center Dyslexia Laboratory School setting has been well received by students and teachers and also effective for student growth.

Next Steps

The next step is to determine the efficacy of this promising solution in a less controlled, more typical remedial classroom setting. To evaluate the effectiveness of the *Bridges* program, student outcomes must be compared to outcomes for students receiving both *Take Flight* and other high-quality, validated dyslexia interventions. All instruction will be provided by school district personnel. SRC is only collecting data to assess outcomes. The next sections of this document describe the study interventions and outcome measures. The following pages present preliminary analysis of student outcome data.

Comparative Intervention Methods

Take Flight: A Comprehensive Intervention for Students with Dyslexia

Take Flight, is a two-year dyslexia intervention derived from Orton Gillingham-based instructional principles. It integrates evidence-based best practices for teaching the important components of a comprehensive reading program and has demonstrated efficacy and effectiveness in both laboratory and public-school settings (Avrit et al., 2006; Ring et al., 2017). The curriculum scope and sequence is presented in 230- sessions using two alternating daily lesson types. The first lesson plan (New Learning, 132 lessons) introduces combinations of phonemic awareness, phonics concepts (e.g., grapheme-phoneme correspondences), syllable division rules, morphology, spelling rules, vocabulary, and comprehension strategies. Importantly, each new learning is integrated into each of these daily instructional components to allow for additional practice and consolidation across multiple activities. The alternate lesson plan (Application, 98 lessons) provides an opportunity to consolidate student learning by applying previously learned skills and strategies in repeated reading exercises, spelling, dictation, combined with vocabulary development and comprehension strategy use when reading continuous text.

Bridges: A Dyslexia Intervention Connecting Teacher, Avatar, and Student

The *Bridges* program follows the exact scope and sequence as *Take Flight* but is taught by a certified teacher supported by the virtual co-teacher (the avatar). The virtual co-teacher is designed to introduce new learning concepts with high fidelity and accuracy. These avatar-led lesson activities make up 10-15 minutes of daily lesson time. During the remaining 45-50 minutes of class time, the certified teacher will use scripted lesson plans to then lead the students through practice applying new learning.

Comparison Curricula

Students receiving other high-quality, validated dyslexia intervention to provide a valid comparison are also participating in the study. The exact program content and structure varies by district. Participating teachers used programs and combinations of programs including: Neuhaus *Basic Language Skills*, Neuhaus *Language Enrichment*, *Multisensory Teaching Approach (MTA)*, and *Lexia. Basic Language Skills* focuses on phonemic awareness, spelling, handwriting, and reading readiness by teaching graphemes, phonemes, cursive handwriting, and reading practice/comprehension in 192 lessons. Reading practice includes three repetitions to provide for decoding, intonation, and reading rate. Mastery checks are used to monitor progress. *Language Enrichment* is an evidence-based language arts solution for developing successful readers. It includes teaching graphemes, phonemes, and spelling of sounds and words. This is a 30-minute per day curriculum for struggling readers who need intervention with supplementary instructional activities in reading practice and comprehension. *MTA* is a comprehensive, integrated language arts program addressing phonemic awareness, phonics, reading decoding and comprehension, spelling and composition, cursive handwriting, and alphabet and dictionary skills. It is an ungraded program suitable for use with elementary and secondary level students. It consists of 7 kits and continuation to the next level of *MTA* intervention is based on continuing student mastery. These programs provide explicit, systematic, intensive literacy instruction for students with dyslexia or related language learning difficulties.

Measures

District Data

Select information was collected from the school district for each participating student. Students participating in our LWCCEP project have been identified as having the characteristics of dyslexia by the school district evaluation and referred to receiving dyslexia intervention. Results of their district dyslexia evaluation were collected by the study team to confirm eligibility for the study and to provide a baseline of ability. Demographic data received from

the district included age, gender, ethnicity, race, free/reduced lunch, English language learner status, related comorbidities, and any special services the participant may be receiving. Additional intervention-related information was also collected as available, including progress monitoring reports, dyslexia homework completion rates, and student attendance rates.

Study Specific Outcome Measures

Participating students were evaluated twice over the 2021-2022 school year by SRC diagnosticians: once at the start of the school year (Fall 2021) and again at the end of the school year (Spring 2022). The assessments administered included norm-referenced tests of language and literacy achievement and provide information regarding student ability as they progress through their intervention program. The assessments utilized during these evaluations are outlined in Table 1. Participating students will be evaluated on these measures once more at the end of the upcoming school year to measure post-intervention language ability. Additional information regarding outcome measures is presented in the Appendix of this document.

Table 1. Study Specific Outcome Measures

Assessment Name	Subtests/Domains	Reliability Metric
Comprehensive Test of Phonological Processing	Phonological Awareness, Rapid Naming	$\alpha = .92$
Gray Oral Reading Test	Oral Reading Fluency & Comprehension	$\alpha > .91$
Peabody Picture Vocabulary Test	Receptive Vocabulary	$\alpha = .89$
Woodcock Reading Mastery Tests	Word Attack, Word Recognition, Listening Comprehension, Passage Comprehension	$r > .85$
Word Identification and Spelling Test	Regular & Irregular Word Reading, Regular & Irregular Spelling, Letter-Sound Knowledge	$\alpha = .98$

Aggregate Student Sample Report

Participants

The current study aims to evaluate growth in reading ability for students receiving routine dyslexia intervention services in public-school settings. To be eligible for participation, students must have had a school-based identification of dyslexia and must have been enrolled in their first year of school-based intervention services. Families of all eligible students at 34 participating elementary campuses were provided with study information and offered the opportunity to participate. A total of 137 students are actively enrolled in the study. The final sample includes 137 students (61 female) in Grades 2 through 5 (Median: Grade 3). Demographic and baseline characteristics of the full sample are presented in Figure 3 and Table 2.

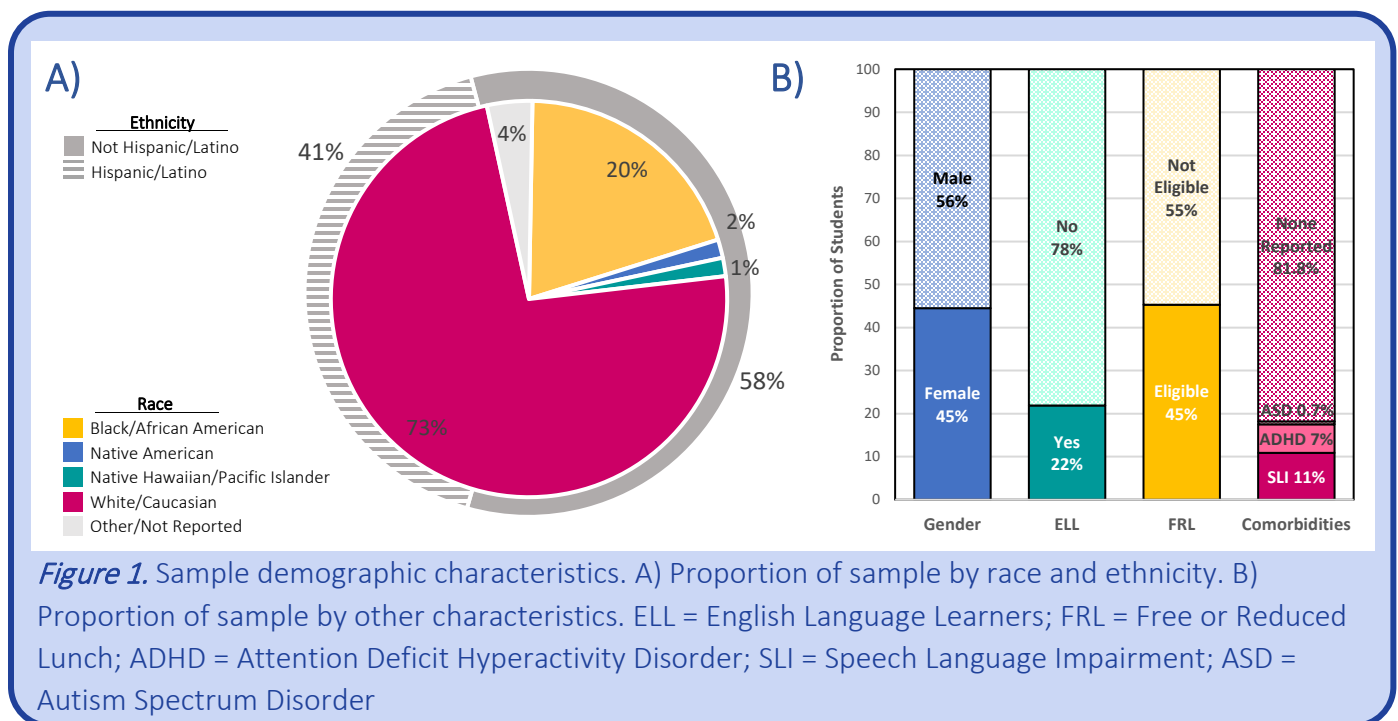


Figure 1. Sample demographic characteristics. A) Proportion of sample by race and ethnicity. B) Proportion of sample by other characteristics. ELL = English Language Learners; FRL = Free or Reduced Lunch; ADHD = Attention Deficit Hyperactivity Disorder; SLI = Speech Language Impairment; ASD = Autism Spectrum Disorder

Summary Findings: Sample Demographics

- Students in the sample were between 7-11 years old at baseline ($M=8y;8m$, $SD=1y;1m$).
- Nearly half of the sample was eligible for free/reduced lunch in the 2021-22 school year.
- Approximately 22% of students in the sample were English Language Learners.
- Nearly 1 in 5 students in the sample had a reported comorbid language- or learning-related diagnosis.
- The *Bridges* and *Other* groups had a higher proportion of Hispanic/Latino and ELL students than the *Take Flight* group. Intervention groups did not differ on any other baseline demographic variable.

Table 2. Demographic Characteristics of the Aggregate Sample.

	Aggregate Sample <i>n</i> = 137	<i>Take Flight</i> <i>n</i> = 59	Bridges <i>n</i> = 59	Other <i>n</i> = 19	Test Values
Age in Years; Months [M(SD)]	8y;8m (1y;1m)	8y;8m (1y;3m)	8y;7m (1y;0m)	8y;10m (1y;0m)	$F(2,134) = 0.35$
Grade [Median]	3	3	4	3	
Sex (% F)	44.5	42.4	44.1	52.6	$\chi^2(2)=0.62$
Race (%)					$\chi^2(8)=4.27$
Asian/Asian American	0.0	0.0	0.0	0.0	
Black/African American	19.9	17.2	23.7	15.8	
Native American	1.5	0.0	1.7	5.3	
Native Hawaiian/Pacific Islander	1.5	1.7	1.7	0.0	
White/Caucasian	73.5	77.6	69.5	73.7	
Other/ Not Reported	3.7	3.4	3.4	5.3	
Ethnicity					$\chi^2(2)=6.12^*$
Hispanic/Latino (%)	41.4	29.8	47.4	57.9	
Free/Reduced Lunch (% Eligible)	54.7	45.8	61.0	63.2	$\chi^2(2)=3.40$
English Language Learner (% Yes)	21.9	11.9	30.5	26.3	$\chi^2(2)=6.25^*$
Comorbidities (%)					$\chi^2(4)=4.92$
ADHD	6.6	8.5	3.4	10.5	
ASD	0.7	1.7	0.0	0.0	
SLI	10.9	8.5	15.3	5.3	
Other	0.0	0.0	0.0	0.0	

Note: * $p \leq .05$

Statistical Analysis

Group composition was compared across all Intervention Types using a series of t-tests and chi-squared analyses. Groups were equivalent across all demographic characteristics (age, sex, race, free/reduced lunch status, and comorbidities) except ethnicity and ELL status. Namely, the *Bridges* and *Other* intervention groups had a greater proportion of Hispanic/Latino and ELL students than the *Take Flight* group.

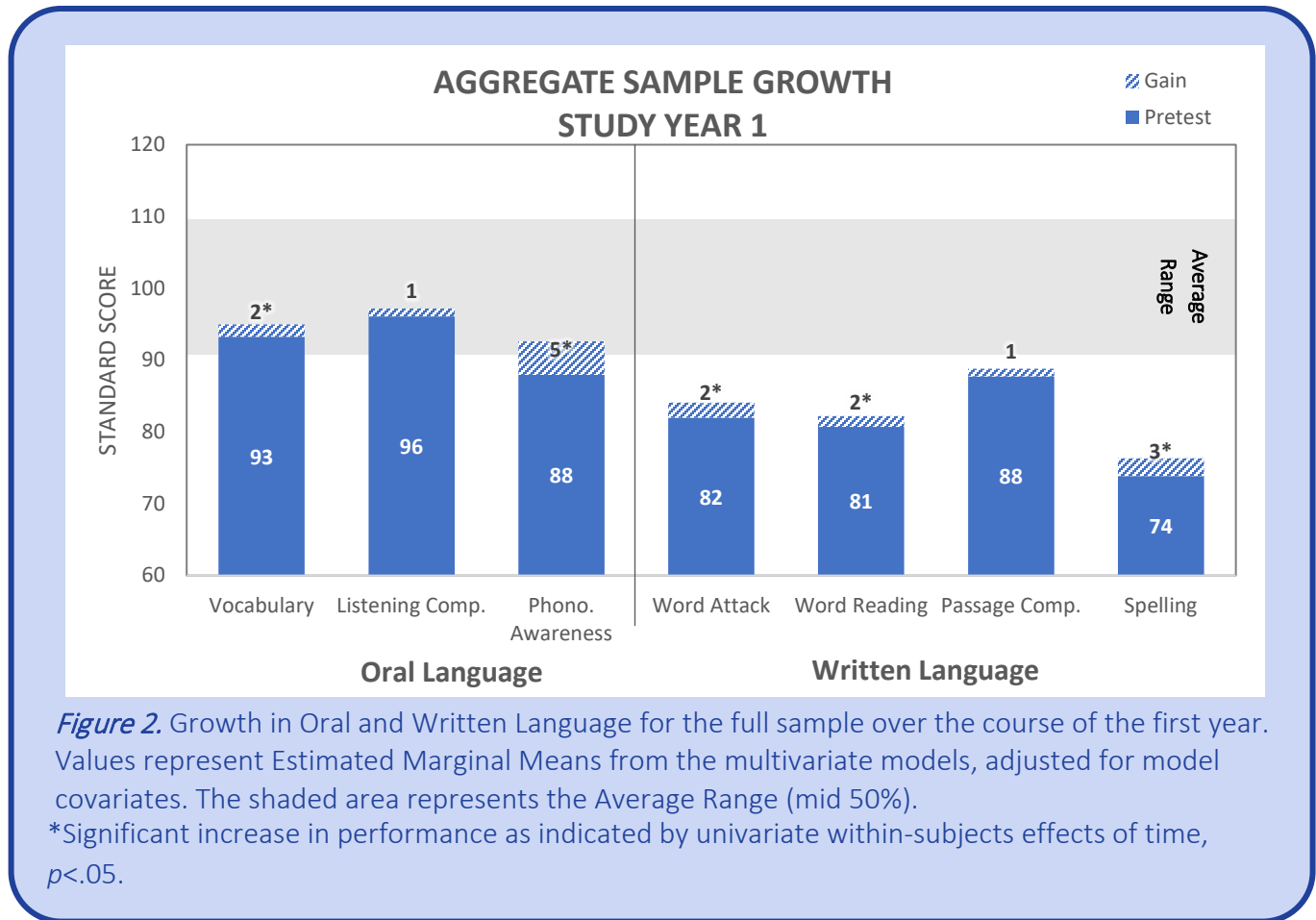
Due to unforeseen complications arising from the COVID-19 pandemic in Fall 2021, the study enrollment period was extended to allow additional time for recruitment and enrollment. Overall, students receiving *Other* instruction enrolled later in the semester compared to the *Take Flight* and *Bridges* groups so the gap of time between pre- and mid-testing was shorter. Latency (time between pre- and mid-testing) was calculated using the number of weeks and compared across groups. Pre-mid testing latency was shorter for the *Other* group ($M=26$ weeks) compared to the *Bridges* and *Take Flight* groups ($M=28$ and $M=29$, respectively).

Pre- and mid-testing data were used to evaluate growth in oral and written language skills. Results of tests of statistical assumptions were satisfactory. Instances of missing data were imputed using sample means. Due to correlations between variables, omnibus multivariate analyses were run on all dependent variables by language domain and followed by univariate analyses. Mean-centered covariates (age, test latency) were used in all analyses. The data were modeled using a series of doubly multivariate analysis of covariance on Oral and Written Language measures across Interventions at pre- and mid-testing, controlling for baseline age, and pre-mid testing latency.

Gain scores were calculated for variables clustered into three domains: Fundamental Reading Skills (Phonological Awareness, Letter-Sound Knowledge), Word-Level Skills (Word Attack, Word Identification, Spelling), and Passage-Level Skills (Fluency and Comprehension). These data were modeled using a series of multivariate analyses of covariance for each domain across Intervention Groups. Baseline variables ethnicity, ELL status, mean-centered age, and mean-centered test latency were covariates in each model to account for preexisting group differences.

Aggregate Treatment Effects

Summary statistics of oral and written language performance for the aggregate sample are depicted in Figure 4. The data show mean performance of the full sample at the beginning of treatment and observed gains over the course of the first academic year. These data indicate several important observations about the sample and interim treatment outcomes.



Summary Findings:

- Students in the sample exhibited characteristics of dyslexia at pre-test:
 - Weak phonological processing scores,
 - Word reading and spelling deficits
 - Relative strengths in oral language skills
- Significant improvements in standard score performance were found for both Oral and Written Language skills over the course of the first year of intervention. This indicates that these students' reading related skills are accelerating at a rate faster than their age-equivalent peers (i.e., narrowing the gap).
- The strongest independent growth effects were found in Phonological Awareness and Spelling, with moderate effects found for Word Attack, Word Identification, and Vocabulary. Effects for Listening and Reading Comprehension were small and statistically non-significant.

Comparative Growth in Reading and Related Language Skills

Summary statistics of all Year 1 outcome measures across intervention type are presented in Table 3. Select measures of oral and written language performance were included in follow-up analyses investigating differences in growth patterns across intervention types. Results of these analyses are depicted in Figures 5-7. These figures depict average change in standard score performance for each group over the course of the first academic year.

Overall, the *Take Flight* sample outperformed both other groups at pre-test and demonstrated the greatest growth for most measured skills. The *Bridges* group was the weakest at pre-test and demonstrated comparatively moderate growth in oral and written language skills. Students receiving instruction in Other interventions fell in between the *Take Flight* and *Bridges* groups at baseline, and demonstrated modest growth over the first year of intervention. Changes in standard score performance for individual group were examined using paired-samples *t*-tests.

Table 3. Pre-test and Mid-test Performance on all Study Evaluation Measures for the Aggregate Sample.

	<i>Take Flight</i> <i>n</i> = 59		<i>Bridges</i> <i>n</i> = 59		Other <i>n</i> = 19		<i>F</i>	η^2
	Pre	Mid	Pre	Mid	Pre	Mid		
Comprehensive Test of Phonological Processing								
Phonological Awareness	91.05 (13.72)	97.53 (14.02)	85.02 (14.43)	89.81 (13.76)	87.47 (12.70)	86.42 (9.74)	2.01	.03
Rapid Naming	86.33 (10.28)	90.02 (12.20)	86.49 (10.45)	89.81 (10.59)	86.89 (7.27)	91.79 (12.44)		
Woodcock Reading Mastery Tests								
Word Attack	83.47 (11.43)	87.56 (13.05)	80.80 (11.12)	81.64 (11.93)	81.26 (11.61)	81.26 (8.85)	1.32	.02
Word Identification	83.29 (13.05)	85.47 (13.74)	78.17 (10.95)	79.42 (11.08)	80.58 (12.67)	80.68 (11.85)	0.36	.01
Passage Comprehension	90.64 (16.15)	93.22 (14.42)	85.07 (13.03)	86.03 (12.75)	87.05 (16.85)	83.63 (10.38)	3.20*	.04
Listening Comprehension	98.33 (13.85)	100.54 (13.41)	93.37 (13.86)	94.86 (11.52)	97.26 (12.15)	94.21 (11.47)	2.14	.04
Gray Oral Reading Test								
Fluency†	6.57 (2.39)	6.80 (2.06)	5.95 (2.13)	6.34 (2.29)	6.05 (2.27)	6.26 (1.76)	0.15	.00
Comprehension†	7.72 (2.60)	7.66 (2.25)	6.42 (2.17)	7.07 (2.44)	6.79 (2.07)	6.95 (1.50)	1.62	.02
Word Identification and Spelling Test								
Spelling	75.44 (9.76)	79.53 (12.04)	71.69 (7.36)	74.44 (9.51)	75.26 (9.42)	72.26 (9.71)	6.90**	.10
Letter-Sound Knowledge††	50.05 (12.50)	65.17 (13.55)	47.00 (12.40)	57.34 (12.18)	48.47 (10.60)	51.37 (10.47)	5.73**	.08
Peabody Picture Vocabulary Test								
Receptive Vocabulary	96.03 (12.34)	99.10 (13.30)	91.36 (13.56)	91.59 (12.15)	90.63 (11.28)	92.79 (12.14)	2.29	.03

Note: All scores presented are standard scores unless otherwise noted. Standard scores are adjusted for student age at testing and fall on a distribution with an average of 100 (50th percentile) and standard deviation of 15.

† Represents scaled scores (*M* = 10, *SD* = 3).

†† Represents raw scores.

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

Comparative Growth in Fundamental Reading Skills

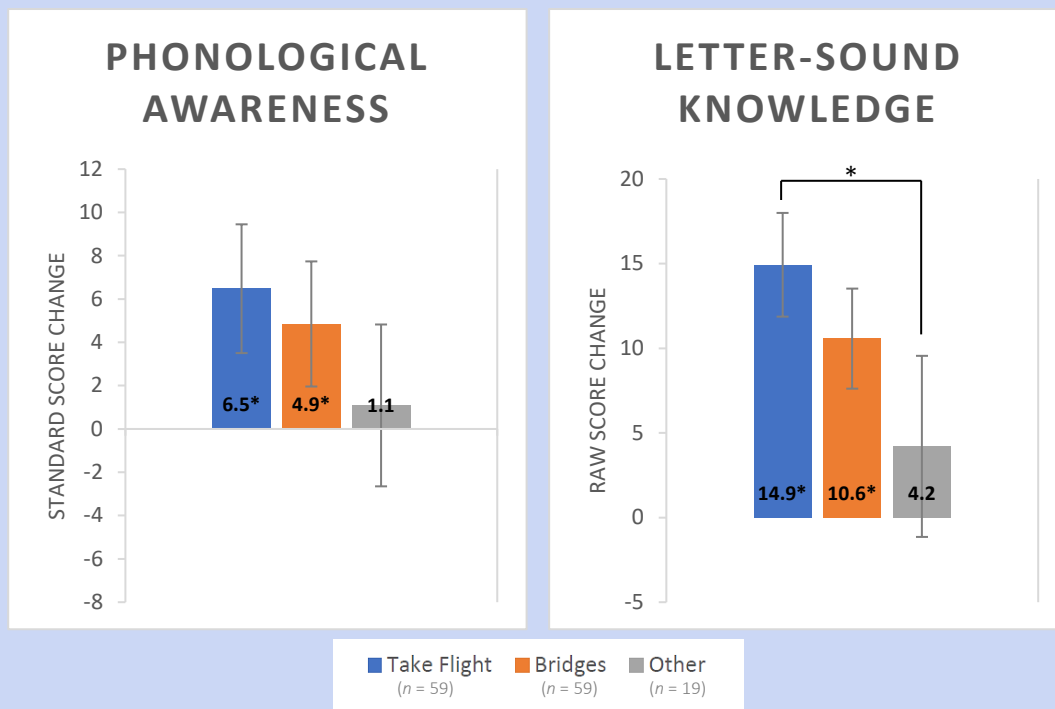


Figure 3. Comparative treatment outcomes in Fundamental Reading Skills across Intervention Type. Values represent Estimated Marginal Means adjusted for model covariates.

Note: Vertical error bars represent the 95% Confidence Interval around group means. Horizontal bars indicate significant differences in growth across groups as indicated by univariate pairwise comparisons using Bonferroni's alpha correction ($p < .05$).

*Change significant at $p < .05$ using paired-sample t -test.

†Change marginally significant ($p = .05-.10$) using paired-sample t -test.

Summary Findings:

- Fundamental reading skills support reading acquisition and are characteristically weak in students with dyslexia. These skills typically respond well to intervention with early and sustained improvements with direct systematic instruction.
- Both measured fundamental skills increased over the course of intervention Year 1. Continued growth for students in all instructional groups is expected in Fundamental Reading Skills over the course of intervention Year 2.
- Performance on Phonological Awareness and Letter-Sound Knowledge assessments revealed significant growth for the *Take Flight* and *Bridges* groups. Changes in fundamental reading skills were more variable for students receiving instruction in Other programs and these changes did not reach a level of statistical significance.
- Group differences were only significant for Letter-Sound Knowledge between the *Take Flight* and Other groups, suggesting that the groups relative performance over time is similar across interventions.

Comparative Growth in Word-Level Skills

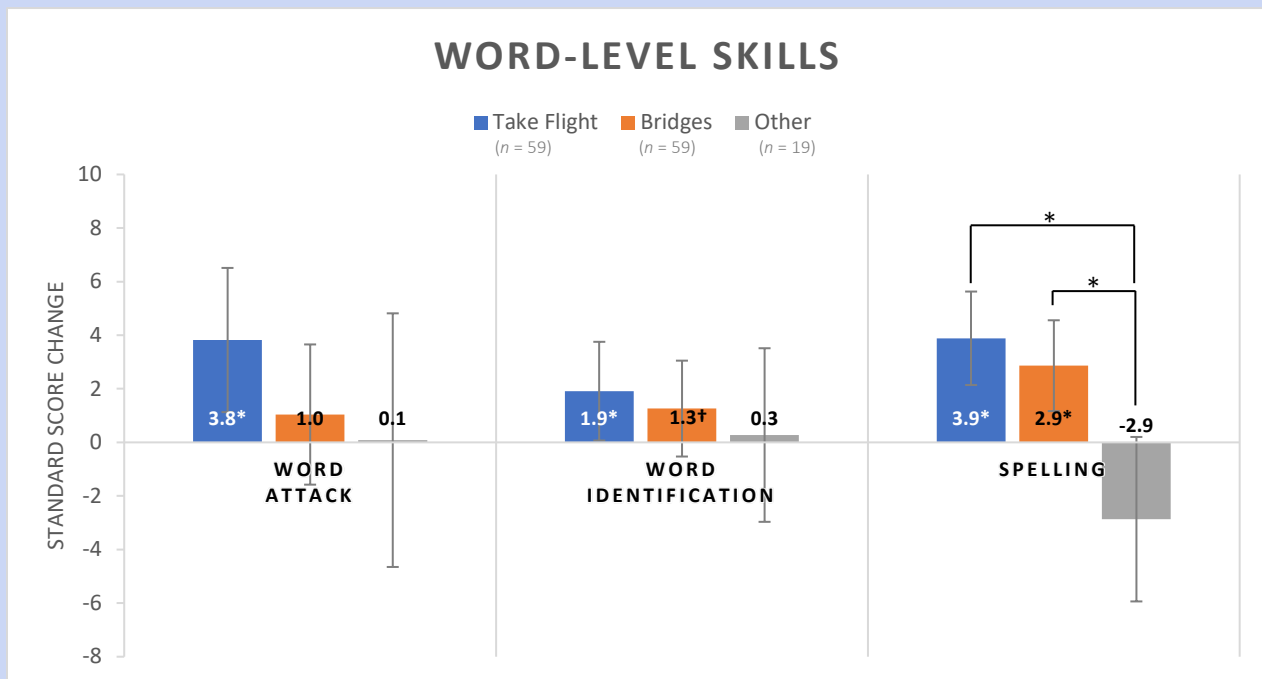


Figure 4. Comparative growth in Word-Level Skills across Intervention Type. Values represent Estimated Marginal Means adjusted for model covariates.

Note: Vertical error bars represent the 95% Confidence Interval around group means. Horizontal bars indicate significant differences in growth across groups as indicated by univariate pairwise comparisons using Bonferroni's alpha correction ($p < .05$).

*Change significant at $p < .05$ using paired-sample t -test.

†Change marginally significant ($p = .05-.10$) using paired-sample t -test.

Summary Findings:

- Student performance on standardized tests of word-level reading skills increased over the first year of intervention, though this growth was modest in comparison to fundamental reading skills. Continued growth is expected across all word-level skills in the second year of intervention for students in all instructional groups.
- The groups did not differ in amount of growth for Word Attack and Word Identification, though the *Take Flight* group demonstrated the greatest growth, followed by the *Bridges* group and then the Other group.
- Students receiving Other instruction demonstrated little change in Word Attack and Word Identification performance, as well as a decrease in Spelling performance, though this change did not reach a level of statistical significance. Variability in Spelling performance may be the result of differences across programs in their respective approaches to spelling instruction.

Comparative Growth in Passage-Level Skills

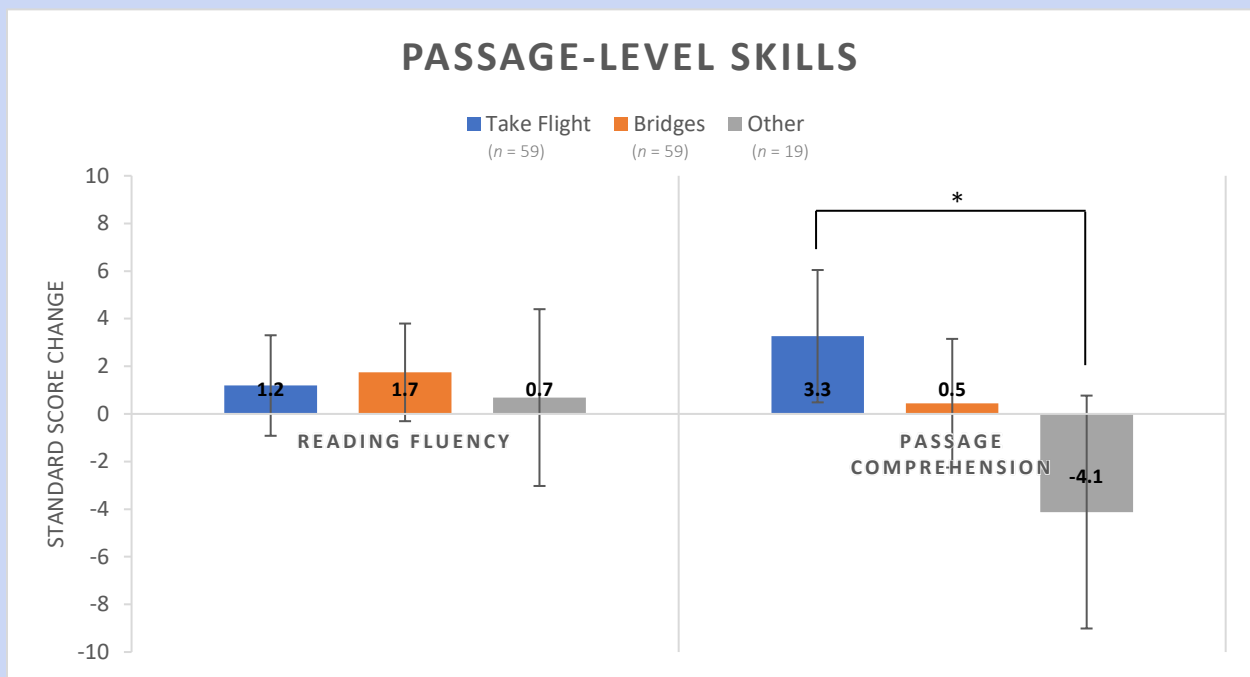


Figure 5. Comparative growth in Passage-Level Reading Skills across Intervention Type. Values represent Estimated Marginal Means adjusted for model covariates.

Note: Vertical error bars represent the 95% Confidence Interval around group means. Horizontal bars indicate significant differences in growth across groups as indicated by univariate pairwise comparisons using Bonferroni's alpha correction ($p < .05$).

*Change significant at $p < .05$ using paired-sample t -test.

†Change marginally significant ($p = .05-.10$) using paired-sample t -test.

Summary Findings:

- Passage-level skills are typically acquired later in the developmental timeline after word-level skills are developed to proficiency. Modest and variable growth in these reading skills is to be expected at this early stage in the intervention timeline.
- Student performance on standardized tests of passage-level reading skills generally increased over the first year of intervention, though there were significant differences in growth across groups for Passage Comprehension.
- Direct instruction in reading comprehension skill is incorporated into both SRC intervention programs, accounting for up to 25% of instructional time. The extent to which these skills are taught in the various Other intervention programs is variable, which may impact student growth in this area.

Conclusions

Students in this sample receiving public-school dyslexia intervention improved in Oral and Written Language skills over the first year of instruction. As expected, growth was generally greatest for Phonological Awareness, followed by word-level skills such as word reading and spelling, and then higher-level skills such as fluency and comprehension. This pattern of skill acquisition is typical for students with dyslexia who are developing characteristically weak reading abilities and suggests students in the sample are acquiring reading skills in a developmentally appropriate way. Additional growth is expected for students in all instructional programs over the course of the second year of intervention.

For most measured skills, including Phonological Awareness, Word Attack, Word Identification, and Reading Fluency, differences across groups in the amount of Year 1 growth was not significant. For these measures, changes in performance over the first year of intervention were not reliably different across programs; that is, student growth was generally similar regardless of instructional group. This suggests that the programs are equally effective in eliciting growth in these skills during the first year of intervention. However, significant group differences were found for Letter-Sound Knowledge, Spelling, and Comprehension. These differences were driven by the following patterns: for each of these three measured skills, the growth demonstrated by the *Take Flight* group was significantly greater than that of the Other intervention group. The *Bridges* group also demonstrated greater growth than the Other intervention group on a measure of Spelling. The *Take Flight* and *Bridges* groups did not differ on any measured reading skill. This pattern of findings suggests that there may be differences in the design of dyslexia intervention programs which leads to greater or lesser growth in certain skill development over time. Specifically, both SRC programs, *Take Flight* and *Bridges*, are designed such that direct instruction in both Spelling and Reading Comprehension is systematically incorporated into the lesson sequence, accounting for up to 10% and 20% of total instructional time, respectively. Relative proportions of instructional time for students in the Other intervention group vary, as this group includes students receiving instruction in various programs and implementation structures. It is important to note that these findings are aggregated at the student level and do not reflect the individual performance of any given teacher or student. However, these findings do suggest that the direct and systematic approach to instruction in these areas for SRC programs provides additional benefit over other traditional practices.

These data provide additional support for the hypothesis that the innovative *Bridges* approach to dyslexia intervention is not inferior to traditional approaches such as those used in the *Take Flight* and Other intervention instructional groups. Students receiving *Bridges* instruction demonstrated similar improvements in measured skills in comparison to those receiving traditional instruction.

Limitations

There are several important limitations to the interpretation of these findings. First, the relatively small sample and large variability for the Other intervention group may obscure significant relationships between variables and over time. Collecting data on additional students across all instructional groups will increase the statistical power of the models to uncover reliable differences in performance. Furthermore, although covariates were entered into each of the models to account for pre-existing differences across groups, these differences cannot be nullified and therefore warrant caution in the comparison of performance across groups. Additionally, these data were collected over the first year of a multi-year intervention and additional growth is expected over the course of the second academic year. These results should not be interpreted as representative of outcomes produced by a full course of dyslexia intervention that typically takes four or more semesters to complete. However, these findings do support the initial efficacy and equivalence of instructional approaches in eliciting early reading growth. Finally, and perhaps most critically, these data were collected on a group of students receiving instruction during a year which saw several surges in the COVID-19 pandemic, which caused upheaval in both educational and personal contexts for many students and their families. These limitations preclude the generalization of the current findings to broader populations and educational contexts. However, the improvements documented in this sample support the benefit of explicit, systematic, intensive dyslexia intervention even under the most challenging of circumstances.

References

- Avrit, K., Allen, C., Carlsen, K., Gross, M., Peirce, D., & Rumsey, M. (2006). *Take Flight: A Comprehensive Intervention for Students with Dyslexia*. Dallas, TX: Texas Scottish Rite Hospital for Children.
- Castiglioni-Spalten, M. L., & Ehri, L. C. (2003). Phonemic awareness instruction: Contribution of articulatory segmentation to novice beginners' reading and spelling. *Scientific Studies of Reading, 7*(1), 25-52.
- Dunn, D.M. (2019). *Peabody Picture Vocabulary Test (Fifth Edition)*. Bloomington, MN: NCS Pearson.
- Lindamood, P., & Lindamood, P. (1975). *Auditory Discrimination in Depth*. Austin: Pro-ed.
- Navarra, J., Yeung, H. H., Werker, J. F., & Soto-Faraco, S. (2012). 24 Multisensory Interactions in Speech Perception.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and instruction, 1*(2), 117-175.
- Ring, J.J., Avrit, K.J., & Black, J.L. (2017). Take Flight: the evolution of an Orton Gillingham-based curriculum. *Annals of Dyslexia, 67*, 383-400.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive test of phonological processing: CTOPP (Second Edition)*. Austin, TX: Pro-ed.
- Wiederholt, J. L., & Bryant, B. R. (2012). *GORT-5: Gray Oral Reading Tests Examiner's Manual*. Austin, TX: Pro-Ed.
- Wilson, B. A., & Felton, R. H. (2004). *Word Identification and Spelling Test: Examiner's Manual*. Pro-ed.
- Woodcock, R. (2011). *The Woodcock Reading Mastery Tests (Third Edition)*. Bloomington, MN: Pearson.
- Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental science, 12*(5), 732

Appendix

Student Study Outcome Measures

Woodcock Reading Mastery Test, 3rd Edition (WRMT-3). A standardized test of single word recognition, phonological decoding, listening comprehension, and reading comprehension. Phonological decoding ability as assessed with the Word Attack subtest from the Woodcock Reading Mastery Test 3rd Edition (Woodcock, 2011). The instrument requires participants to pronounce a list of isolated pseudowords with no time limits on participant response. Single-word reading was measured with the WRMT-3 Word Identification subtest. Participants were required to read isolated real words, again with no time constraints. These two untimed word-level reading measures have reported split-half reliabilities of .97 and .98, respectively.

Receptive language ability was measured with the Listening Comprehension subtest from the WRMT-3. Participants listen to passages and dialogues and respond to orally presented questions. This subtest has a reported split-half reliability of .88. Reading comprehension was measured using the Passage Comprehension subtest of the WRMT-3. This test utilizes a cloze-type procedure to measure comprehension of narrative and expository passages. This subtest has an average split-half reliability of .86.

Gray Oral Reading Test, 5th Edition (GORT-5). A standardized test of timed oral passage reading that measures oral reading fluency and comprehension. Oral reading fluency will be measured using the GORT-5 (Wiederholt & Bryant, 2012). The test requires participants to read aloud short narrative and expository passages and respond to orally presented comprehension probes. Fluency scores are derived from a combination of Rate and Accuracy subscale scores. Comprehension scores are derived from total student performance on the comprehension probes across all passages. The GORT-5 manual reports high internal consistencies ($\alpha > .91$) for all subscales (Rate, Accuracy, Fluency, and Comprehension).

Comprehensive Test of Phonological Processing, 2nd Edition (CTOPP-2). Additional assessments of reading-related skills included the Phonological Awareness and Rapid Symbolic Naming composite measures from the Comprehensive Test of Phonological Processing 2nd Edition (CTOPP-2; Wagner et al., 2013). The Elision subtest requires participants to elide individual phonemes from verbally presented words to form real word responses. The Blending Words subtest requires participants to combine verbally presented phonemes to form real word responses. The Phoneme Isolation subtest requires participants to provide the first, last, or middle sound from verbally presented real words. The composite measure of the three subtests has reported internal consistency of $\alpha = .92$. The Rapid Naming composite requires participants to verbally-name arrays of alpha (Rapid Letter Naming) or numeric (Rapid Digit Naming) stimuli under timed conditions. The composite measure of the two subtests has reported internal consistency of .92.

Word Identification and Spelling Test (WIST). The WIST provides a nationally standardized, diagnostic instrument designed specifically for students having difficulty with reading and/or spelling (Wilson & Felton, 2009). The Word Identification and Spelling subtests of the WIST include stimuli sets of both regular and irregular words. Regular words for both reading and spelling performance are recorded at the sub-lexical level (i.e., syllables and affixes), allowing for a more granular assessment of decoding and word recognition ability. These subtest measures were administered to all students at each evaluation period. The Word Identification and Spelling subtest measures each have reported internal consistencies of $\alpha = .98$, and strong convergent validity with other common and reliable measures of reading and spelling (i.e., WIAT-II, WRMT-R/NU, TWS; all $r_s \geq .8$).

Peabody Picture Vocabulary Test, 5th edition (PPVT). This test is an individually administered, norm referenced instrument that assesses receptive (hearing) vocabulary of examinees from 2 to 90 years of age. It measures verbal ability in standard American English vocabulary (Dunn, 2019). The Peabody Picture Vocabulary Test has reported internal consistencies of .89 – .97 and demonstrates convergent validity with other common and reliable measures of vocabulary (i.e., CELF-4, $r = .67-.75$; EVT-2; $r = .80-.84$).

Table A4. Demographic Characteristics of Participating Richardson ISD Students

	Total <i>n</i> = 65	Bridges <i>n</i> = 25	Take Flight <i>n</i> = 30	Other <i>n</i> = 10	Test Values
Age in Years; Months	8y;8m	8y;11m	8y;5m	8y;8m	$f(2, 62) = 1.3$
Grade (Median)	3	4	3	3	
Sex (% Female)	40.0	48.0	36.7	30.0	$\chi^2(2) = 1.2$
Race (%)					$\chi^2(6) = 7.7$
Asian/Asian American	0.0	0.0	0.0	0.0	
Black/African American	16.9	28.0	13.3	0.0	
Native American	0.0	0.0	0.0	0.0	
Native Hawaiian/Pacific Islander	1.5	4.0	0.0	0.0	
White/Caucasian	75.4	60.0	80.0	100.0	
Other/Not Reported	6.2	8.0	6.7	0.0	
Ethnicity (% Hispanic/Latino)	35.4	13.3	40.0	90.0	$\chi^2(2) = 18.5$
English Language Learner (% Yes)	24.6	44.0	0.0	50.0	$\chi^2(2) = 18.3$
Economically Disadvantaged (% Free/Reduced Lunch)	47.7	76.0	13.3	80.0	$\chi^2(2) = 26.4$
Comorbidities (%)					$\chi^2(2) = 2.60$
Attention Deficit/Hyperactivity Disorder	7.7	8.0	10.0	0.0	
Autism Spectrum Disorder	0.0	0.0	0.0	0.0	
Specific Language Impairment	15.4	28.0	6.7	10.0	
Other	0.0	0.0	0.0	0.0	

Note: Results of all between-groups comparisons were non-significant.

Table A5. Model results from doubly multivariate analyses of variance for pre-test and mid-test scores.

	Type of Effect	Assessment Name	<i>df1, df2</i>	<i>F</i>	η^2
<i>Oral Language Skills</i>	<i>MV</i>		<i>3, 132</i>	<i>8.04***</i>	<i>.15</i>
	UV	PPVT	1, 134	6.40*	.05
	UV	WRMT	1, 134	1.00	.01
	UV	CTOPP	1, 134	22.37***	.14
<i>Written Language Skills</i>	<i>MV</i>		<i>4, 131</i>	<i>5.41***</i>	<i>.14</i>
	UV	WRMT	1, 134	6.18*	.04
	UV	WRMT	1, 134	6.49*	.05
	UV	WRMT	1, 134	1.41	.01
	UV	WIST	1, 134	14.44***	.10

Note: Results of Multivariate analyses in italics. MV = multivariate, UV = Univariate. PPVT = Peabody Picture Vocabulary Test, WRMT = Woodcock Reading Mastery Test, CTOPP = Comprehensive Test of Phonological Processing, WIST = Word Identification and Spelling Test.

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

Table A6. Model results from multivariate analyses of covariance on student gain scores.

	Type of Effect	Assessment Name	$df1, df2$	F	η^2
<i>Fundamental Skill Growth</i>	<i>MV</i>		<i>4, 260</i>	<i>2.86*</i>	<i>.04</i>
	Phonological Awareness	CTOPP	2, 130	2.01	.03
	Letter-Sound Knowledge	WIST	2, 130	5.73**	.08
<i>Word-Level Reading Growth</i>	<i>MV</i>		<i>6, 258</i>	<i>2.50*</i>	<i>.06</i>
	Word Attack	WRMT	2, 130	1.32	.02
	Word Identification	WRMT	2, 130	0.36	.01
	Spelling	WIST	2, 130	6.90***	.10
<i>Passage-Level Reading Growth</i>	<i>MV</i>		<i>4, 260</i>	<i>1.71</i>	<i>.03</i>
	Reading Fluency	GORT	2, 130	0.15	.00
	Passage Comprehension	WRMT	2, 130	3.20*	.05

Note: Results of Multivariate analyses in italics. CTOPP = Comprehensive Test of Phonological Processing, WIST = Word Identification and Spelling Test, WRMT = Woodcock Reading Mastery Test, GORT = Gray Oral Reading Test.

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