RESEARCH REPORT

Luke Waites Center for Dyslexia and Learning Disorders

Luke Waites Center Curriculum Efficacy Project

Non-Inferiority of Tech-Assisted Dyslexia Instruction in Remediating Literacy Skills

Anna E. Middleton, Ph.D., CALT

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

SCOTTISH RITE FOR CHILDREN®

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 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.

To evaluate the effectiveness of the *Bridges* program, student outcomes were compared to outcomes for students receiving *Take Flight* in a typical classroom setting. All instruction was 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 findings from a preliminary analysis of student outcome data.

Comparative Intervention Methods

Take Flight: A Comprehensive Intervention for Students with Dyslexia

Take Flight is an extensive, multi-componential 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 is presented in 230 sessions using 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 concept 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 students with opportunities to consolidate learning by applying previously learned skills and strategies through oral 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 content of the *Bridges* intervention 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, a trained certified teacher uses scripted lesson plans to then lead the students through practice applying new learning.

Measures

District Data

Select information was collected from participating school districts for each participating student. Students participating in this project were identified as having the characteristics of dyslexia by standard school district procedures and referred for dyslexia intervention. Results of each participant's district dyslexia evaluation were collected by the study team to confirm eligibility for the study and to provide a baseline of ability level across various literacy skills. Demographic data received from the district included age, gender, ethnicity, race, free/reduced lunch status, English learner status, related comorbidities, and any special services the participant may be receiving.

Study Specific Outcome Measures

Participating students were evaluated three times over the course of two academic years by SRC diagnosticians: once at the start of the school year (pre-test), at the end of the first school year (mid-test), and again at the end of the second school year (post-test). 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 (see Table 1).

Table 1. *Study Specific Outcome Measures*

Assessment Name	Subtests/Domains	Reliability Metric
Comprehensive Test of Phonological Processing, 2nd Ed.	Phonological Awareness	α = .92
Gray Oral Reading Test, 5th Ed.	Accuracy, Rate, Oral Reading Fluency	α > .91
Woodcock Reading Mastery Tests, 3rd Ed.	Word Attack, Word Identification, Passage Comprehension	*r* > .85
Word Identification and Spelling Test	Spelling	α = .98

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 27 participating elementary campuses were provided with study information and offered the opportunity to participate. Those who provided consent were enrolled and scheduled for study evaluations and data collection. The final sample includes 162 students (86 female) in Grades 2 through 5 (Median: Grade 3). Demographic and baseline characteristics of the full sample are presented in Figure 1 and Table 2.

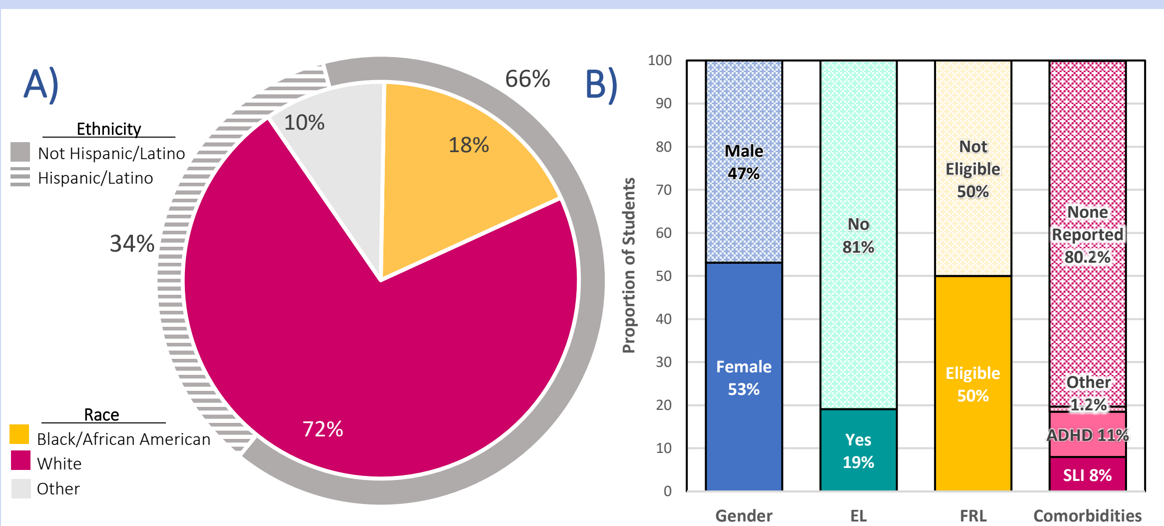

Figure 1. Sample demographic characteristics. A) Proportion of sample by race and ethnicity. B) Proportion of sample by other characteristics. EL = English Language Learners; FRL = Free or Reduced Lunch; ADHD = Attention Deficit Hyperactivity Disorder; SLI = Speech Language Impairment.

Table 2. *Demographic characteristics of the sample*

	Aggregate Sample n = 162	*Bridges* n = 82	*Take Flight* n = 80	Test Values
Age in Years; Months [M(SD)]	8y;7m (1y;1m)	8y;6m(1y;0m)	8y;8m(1y;2m)	$t(160) = 0.81$
Grade [Median]	3	3	3	
Sex (% F)	53.1	53.7	52.5	$\chi^2(1)=0.02$
Race (%)				$\chi^2(2)=1.99$
Black/African American	17.9	22.0	13.8	
White/Caucasian	72.2	69.5	75.0	
Other	9.9	8.5	11.3	
Ethnicity (% Hispanic/Latino)	35.0	45.0	25.0	$\chi^2(1)=7.03$**
Free/Reduced Lunch (% Eligible)	50.0	58.5	41.3	$\chi^2(1)=4.84$*
English Learner (% Yes)	19.1	29.3	8.8	$\chi^2(1)=11.02$***
Comorbidities (%)				$\chi^2(3)=7.49$
ADHD	8.0	3.7	12.5	
SLI	10.5	13.4	7.5	
Multiple/Other	1.2	0.0	2.5	
None Reported	80.2	82.9	77.5	

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

Statistical Analysis

Group composition was compared across Intervention Type 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 EL status. Namely, the *Bridges* group had a greater proportion of Hispanic/Latino, FRL eligible, and EL students than the *Take Flight* group. To account for these differences, EL and FRL entered as covariates in all models.

Pre-, mid-, and post-test standard scores were used to evaluate growth in encoding, decoding, and text reading skills. To address the question of comparative growth across groups in reading scores over time, linear mixed effects modeling was used to account for dependencies within the data (i.e., time nested within student, student crossed with examiner). Models were fit using maximum likelihood estimation in the lme4 package in R (Bates et al., 2015). Each model included the following terms: time, group, EL status, FRL status, and time*group interaction. Dichotomous variables were sum coded to aid in the interpretation of fixed effects (i.e., EL status: No = -1, Yes = 1; FRL status: No = -1, Yes = 1). Random intercepts for both Student and Examiner were included in each model. For all models, normality of residuals and random effects were evaluated using histograms and Q-Q plots. Residuals for all models were normally distributed with a mean of approximately zero.

Results of these analyses are presented in the Appendix and depicted in Figures 2-4. These figures depict estimated change in standard score performance for each group over the course of the two-year intervention period, adjusting for covariates.

Comparative Growth in Reading and Related Language Skills

Summary statistics of outcome measures across intervention types 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. Figures 2-4 depict average change in standard score performance for the intervention groups over the course of two academic years. Overall, the *Take Flight* group outperformed the *Bridges* group prior to and over the course of the intervention period. However, the two groups demonstrated comparatively similar growth in measured skills over time, with significant improvements across all measured skills.

Table 3. *Pre-, Mid, and Post-test performance on all study evaluation measures across groups.*

	Bridges (*n* = 82)			*Take Flight* (*n* = 80)		
	Pre	Mid	Post	Pre	Mid	Post
Encoding						
Phonological Awareness [a]	86.01 (14.23)	90.96 (13.34)	92.13 (14.06)	91.51 (14.79)	97.28 (14.48)	99.34 (12.99)
Spelling [b]	71.51 (7.52)	74.54 (10.44)	80.1 (15.38)	76.14 (10.52)	79.8 (12.71)	83.48 (15.00)
Decoding						
Word Attack [c]	79.95 (11.55)	79.6 (12.38)	86.32 (14.27)	84.21 (13.51)	85.91 (14.86)	89.68 (13.39)
Word Identification [c]	80.79 (11.46)	82.54 (12.53)	85.91 (13.57)	85.69 (13.09)	87.98 (13.60)	90.39 (14.51)
Accuracy [d†]	5.43 (2.33)	6.09 (2.39)	6.43 (2.61)	5.43 (2.33)	6.89 (2.53)	6.72 (2.57)
Text Reading						
Rate [d†]	5.63 (2.33)	6.49 (2.16)	6.75 (2.06)	5.63 (2.33)	7.21 (2.49)	7.71 (2.30)
Fluency [d†]	5.74 (2.14)	6.17 (2.19)	6.44 (2.30)	6.81 (2.53)	6.99 (2.41)	7.07 (2.20)
Passage Comprehension [c]	84.77 (13.64)	85.99 (12.16)	89.06 (13.21)	90.83 (16.05)	92.51 (14.17)	95.06 (14.83)

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 (50^{th} percentile) and standard deviation of 15. Standard error in parentheses. a: CTOPP-2, b: WIST, c: WRMT-3, d: GORT-5. †Scaled scores with a mean of 10 and standard deviation of 3.

Comparative Growth in Encoding Skills

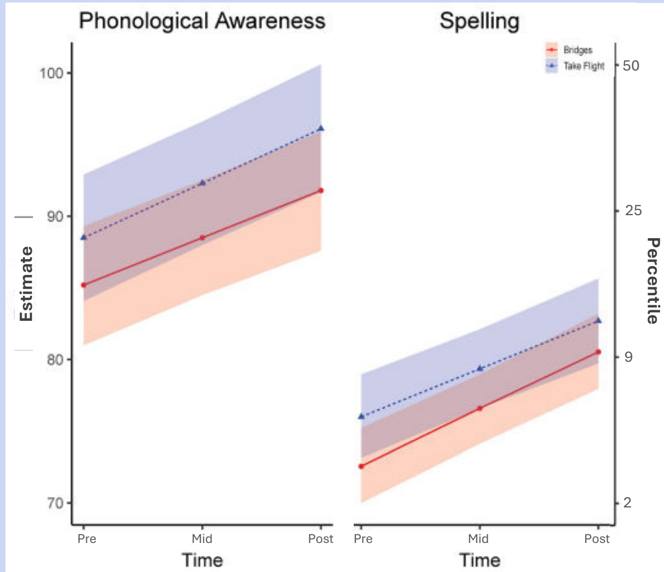

Figure 2. Comparative treatment outcomes in Encoding Skills across Intervention Type. Values represent Estimated Marginal Means adjusted for model covariates. Shaded regions represent the 95% Confidence Interval around group means. Greater overlap in these regions represents greater similarity across groups.

The encoding of oral language into its written counterpart is a critical and foundational component of reading acquisition. As an early indicator of reading ability, phonological awareness is often a key component of core literacy instruction in early grades and is monitored through universal screening measures and targeted interventions. When deficient, phonological awareness may indicate an impaired ability to encode speech sounds as their respective written representations (letters). Students at risk for reading disabilities are often weak in these skills but typically respond well to intervention with early and sustained improvements in response to direct systematic instruction. Spelling involves a more complex encoding task which requires the integration of both phonological awareness and phonics knowledge to correctly produce a written word. In samples of students with dyslexia, growth in spelling is often modest in relation to reading and phonological awareness abilities. However, spelling is a key indicator of an individual's knowledge of a written language and is informative as an outcome of interest in the evaluation of growth.

Encoding skills improved over time for the sample as a whole. Mean pre-test performance for both groups was within or just below the Average Range on a measure of Phonological Awareness, suggesting this skill to be a relative strength for our sample. This may reflect previous intervention efforts targeting phonological awareness and/or other literacy intervention. Despite relatively strong performance on PA at the beginning of the intervention, both groups demonstrated significant growth over time.

As expected, much greater deficits were exhibited for the Spelling measure. This is in line with previous findings suggesting that spelling skills are less malleable compared to reading skills. Nevertheless, the increase in spelling performance over time was statistically and clinically significant for both groups, and the rate of growth did not differ across groups. Given the severity of deficits in Spelling, students will likely need continued support to further develop spelling skills.

Comparative Growth in Decoding Skills

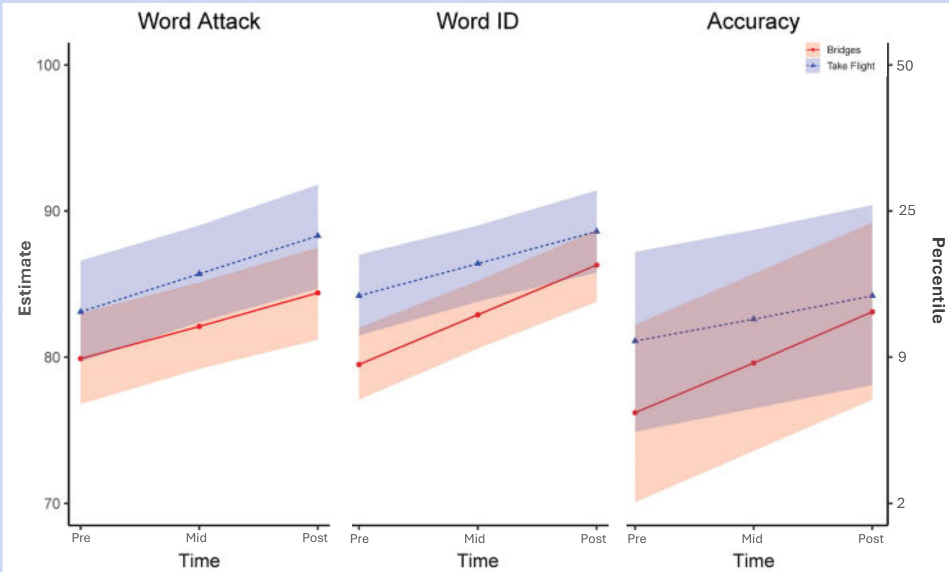

Figure 3. Comparative growth in Word-Level Skills across Intervention Type. Values represent Estimated Marginal Means adjusted for model covariates. Shaded regions represent the 95% Confidence Interval around group means. Greater overlap in these regions represents greater similarity across groups.

Dyslexia is a reading disorder characterized by deficits which often manifest as inaccuracy and/or dysfluency at the level of the individual word. Thus, improving word-level skills is the focus of interventions designed to improve literacy skills for students with dyslexia. Furthermore, improving word-level skills can have the added benefit of providing additional support for higher-order reading skills such as passage-level fluency and comprehension. Decoding activities are a key component of instruction in both the *Take Flight* and *Bridges* curricula. Each phonics concept introduced as part of these programs is practiced through word, sentence, and passage-level exercises.

In the current sample, performance on standardized tests of decoding skills increased over the course of intervention. As shown in Figure 3, both groups were deficient across a variety of decoding skills prior to the intervention. Despite the stronger overall performance for the *Take Flight* group, decoding improved for students in both groups over the course of intervention. Growth rates were similar across groups on measures of Pseudoword Decoding and Word Identification, which are both untimed measures of isolated word reading. Reading Accuracy, a timed measure of decoding in the context of passage reading, revealed a difference in growth rates over the course of intervention: students receiving *Bridges* instruction demonstrated significantly greater growth over time, achieving similar levels of performance by the end of the intervention period. The accelerated growth in Reading Accuracy for the *Bridges* group suggests that the intervention supported growth in these skills even for a relatively more impaired group of learners.

Comparative Growth in Passage-Level Skills

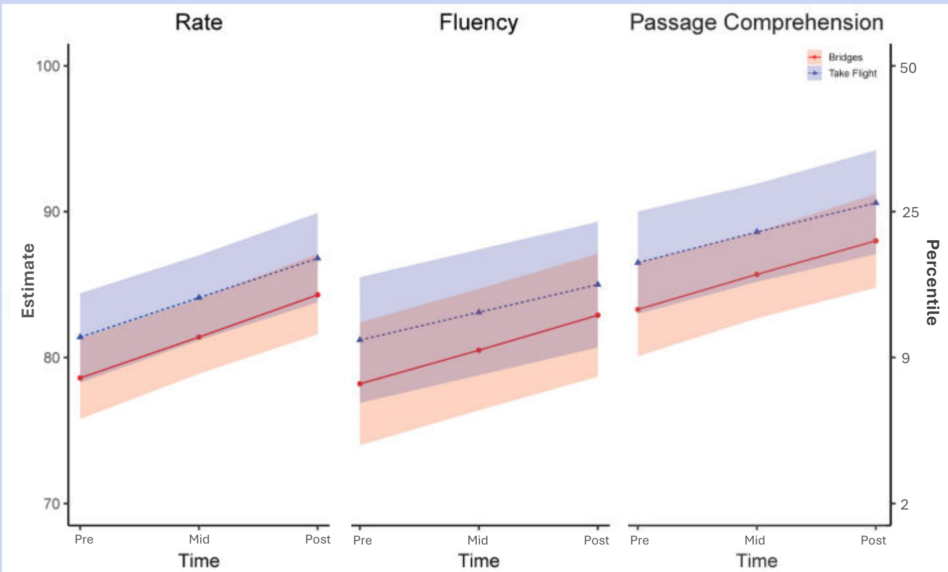

Figure 4. Comparative growth in Passage-Level Reading Skills across Intervention Type. Values represent Estimated Marginal Means adjusted for model covariates. Shaded regions represent the 95% Confidence Interval around group means. Greater overlap in these regions represents greater similarity across groups.

Passage-level reading skills, such as oral reading fluency and comprehension, are often secondary deficits for students with dyslexia who experience weaknesses at the word-level. These abilities are synonymous with skilled reading and are typically acquired later in the developmental sequence as word-level skills are developed to proficiency and with additional reading experience. Best practices for supporting these complex skills rely on increasing exposure to print, including repeated reading, connected text reading, and strategy instruction. These components, among others, are incorporated into both the *Take Flight* and *Bridges* curricula to support the development of higher-order reading ability.

Student performance on standardized tests of passage-level reading skills significantly improved over the course of intervention, bringing these students closer to age-appropriate skill levels on complex reading tasks. Students in both groups improved in Reading Rate, indicating that students were not only more accurate when reading passages aloud (i.e., improved Reading Accuracy), but they completed passages in a shorter amount of time. These two aspects passage reading (Rate and Accuracy) are combined to create a measure of Reading Fluency. Therefore, improvements in passage-level fluency scores require accelerated growth in both the accuracy and speed with which students read increasingly challenging texts. Students in both groups significantly improved Reading Fluency over the course of the intervention.

Both programs systematically include instructional activities targeting reading comprehension. Average standard score performance on a measure of passage comprehension increased over time, but this change did not reach statistical significance. Students in both groups exhibited strengths in passage comprehension relative to word-level reading skills prior to intervention and reached the average range by the end of the intervention.

Conclusions

Students in this sample receiving public-school dyslexia intervention improved in oral and written language skills over the course of instruction. Growth in standard score performance was evident for all measured skills, with slightly greater growth exhibited on measures of encoding and decoding in comparison to passage-level reading skills. 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 in all skills may be evident with additional practice and exposure to print.

Rate of student growth was generally similar across the two methods of instruction, suggesting that the traditional and tech-assisted instructional programs are similarly effective in eliciting growth in literacy skills over the course of intervention. Groups did differ in standard score growth on a measure of Oral Reading Accuracy. Students in the *Bridges* group demonstrated weaker baseline skills on this measure and improved standard score performance at a faster rate compared to those in the *Take Flight* group. Further research is needed to understand the extent to which these differences are attributable to student characteristics or curriculum delivery. 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 utilized by both the *Bridges* and *Take Flight* programs elicits similar growth across traditional and tech-assisted program approaches across various reading skills.

These data provide additional support for the hypothesis that the innovative *Bridges* approach to dyslexia intervention is not inferior to traditional approach of *Take Flight*. Students receiving *Bridges* instruction demonstrated similar improvements in measured skills in comparison to those receiving traditional instruction both in the amount and the rate of growth observed over two years of dyslexia intervention.

There are several important limitations to the interpretation of these findings. First, although covariates were entered into each of the models to account for compositional differences across groups, these differences cannot be nullified and therefore warrant caution in the comparison of performance across groups. However, these findings do support the initial efficacy and equivalence of instructional approaches in eliciting reading growth. Finally, and perhaps most critically, these data were collected on a group of students receiving instruction during a range of school years 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.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

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.

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

Table A1. *Parameter estimates for crossed-random effects models.*

	Phonological Awareness	Spelling	Word Attack	Word ID	Accuracy	Rate	Fluency	Comprehension
Parameters	*B (SE)*	*B (SE)*	*B (SE)*	*B (SE)*	*B (SE)*	*B (SE)*	*B (SE)*	*B (SE)*
Fixed Effects								
Intercept	85.17 (1.99)***	71.55 (1.31)***	79.92 (1.54)***	79.52 (1.22)***	76.16 (2.62)***	78.55 (1.36)***	78.18 (1.91)***	83.28 (1.57)***
EL Status	-1.31 (1.34)	-0.77 (1.09)	-1.37 (1.23)	-0.69 (1.04)	-1.69 (1.06)	-1.78 (1.06).	-1.92 (0.99).	-2.92 (1.17)*
FRL Status	-2.69 (1.03)**	-3.43 (0.84)***	-2.47 (0.95)*	-2.97 (0.8)***	-2.17 (0.82)**	-3.15 (0.82)***	-2.6 (0.77)***	-3.86 (0.9)***
Group	3.35 (2.02)	3.46 (1.67)*	3.2 (1.88)	4.72 (1.57)**	4.91 (1.62)**	2.82 (1.61)	3.04 (1.5)*	3.22 (1.84)
Time	3.29 (0.55)***	4.05 (0.48)***	2.22 (0.55)***	3.38 (0.43)***	3.48 (0.48)***	2.89 (0.45)***	2.36 (0.41)***	2.38 (0.62)***
Group*Time	0.51 (0.76)	-0.64 (0.67)	0.35 (0.76)	-1.23 (0.6)*	-1.9 (0.66)**	-0.16 (0.62)	-0.48 (0.57)	-0.32 (0.86)
Random Effects (var)								
Student intercept	119.87	78.73	99.07	71.85	73.67	75.27	65.86	83.9
Examiner intercept	9.34	0.62	1.95	0.42	31.8	1.95	13.78	2.89
Residual	39.77	30.65	40.71	25.53	30.21	26.67	22.59	52.18
Model Fit								
AIC	3356.7	3175.36	3338.3	3144.84	3210.25	3162.59	3101.3	3395.78
BIC	3393.82	3212.38	3375.44	3181.99	3247.37	3199.71	3138.43	3432.93
Conditional R^2	0.8	0.78	0.75	0.79	0.81	0.79	0.82	0.71
Marginal R^2	0.13	0.2	0.12	0.18	0.12	0.2	0.16	0.21
ICC	0.77	0.72	0.71	0.74	0.78	0.74	0.78	0.62

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; ICC = Intraclass Correlation Coefficient.
*$p < .05$, **$p < .01$, ***$p < .001$