Design and implementation of digital aids to empower struggling readers

Patrick M. Donnelly

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Patricia K. Kuhl, Chair

Jason D. Yeatman

Sara T. Kover

Program Authorized to Offer Degree:

Speech & Hearing Sciences

University of Washington

Abstract

Design and implementation of digital aids to empower struggling readers

Patrick M. Donnelly

Chair of the Supervisory Committee:
Professor Patricia K. Kuhl
Speech & Hearing Sciences

Decades of research have established effective curricula for struggling readers. In the digital age, educational technologies expand access to apps and resources to supplement classroom instruction, but most lack scientific backing. *Sound it Out*, a collaborative effort between reading researchers and technology developers, aims to address this disconnect. This dissertation describes the design and development of the *Sound it Out* tool using three intervention studies of children ages 7-15 (n = 175). Each study answers questions of efficacy and the context for which this tool is ideally suited. We find that, with practice, emerging readers can effectively leverage the tool's symbolic vowel annotation to decode novel words more accurately. We also find generalizable gains with extended practice and enhanced benefits with caregiver-supervised usage. Results demonstrate that *Sound it Out* is an effective scaffold for vowel recognition in emerging readers.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This project would not have been possible with the constant support of a number of people.

Firstly, I would like to thank my many supportive mentors in pursuit of this degree: Jason Yeatman, Patricia Kuhl, and Sara Kover. Also, I would like to thank my fellow doctoral students, now friends, who provided support and community: Maggie Clarke, Reva Zimmerman, Ruofan Cai, Kelly Jahn, Liesbeth Gijbels, and Elle O'Brien. And my lab mates, Emily Kubota, Michelle Nue, Piper Doering, and Alex White who have vastly influenced my work and commitment to the thrilling world of research.

For their collaboration with *Sound it Out*, I am immensely grateful for Kevin Larson, Greg Hitchcock, Paul Linnerud, and Tanya Matskewich at Microsoft for giving me support and freedom to pursue this project for my graduate work.

For her constant support, both in life and in research, I am especially thankful to Maryanne Wolf, who has taught me time and time again how truly transformative a teacher can be.

Finally, I thank my parents, my grandparents, Aunt Sue, my siblings, and my husband Evan for inspiring my commitment to learning and having confidence in me every step of the way.

DEDICATION

This dissertation is dedicated to my research participants and their families.

Chapter 1. INTRODUCTION

One of the major feats of humans has been the development and mastery of literacy. Not only a font for imagination, literacy provides a tool for learning, expression, and innovation. Although we are born to use spoken language, the ability to read is not hard-wired (Dehaene & Cohen, 2011; Wolf, 2007) and many children struggle acquiring this skill that is non-trivial to their future success. Due to disability and widespread inequities, there are prominent barriers facing a multitude of children. The most commonly identified learning disability in school-aged youth (Fletcher, Reid Lyon, Fuchs, & Barnes, 2018), developmental dyslexia, affects between 5 and 17% of the population (Shaywitz et al., 1998; Snowling, 2004). Moreover, it is estimated that 80-90% of children in special education programs have impairments in word recognition (Lerner, 1989). More broadly, literacy rates nationwide are declining (National Assessment of Educational Progress, 2019), highlighting a widespread deficiency in how we teach our children to read.

With reading performance being one of the most significant predictors of long-term outcomes (Lyon, Shaywitz, & Shaywitz, 2003; Perie, Grigg, & Donahue, 2005; Raskind, Goldberg, Higgins, & Herman, 1999; Sum, Kirsch, & Yamamoto, 2004), health and longevity (Gilbert, Teravainen, Clark, & Shaw, 2018), policymakers and researchers alike have directed attention and resources to develop solutions. For struggling readers, the act of learning to read is more than challenging, requiring hours of instruction outside of school, with the goal of catching up with their quickly advancing peers. Stressed by the growing cost of dyslexia assessment, parents and caregivers seeking effective solutions for their child are faced with the financially burdensome task of private tutoring. Due to this burden, many families turn to technology-based tools as less costly alternatives. In fact, many apps and technologies are now widely available to augment, or even replace, the teacher in delivering intervention to remediate reading disability. These tools are advertised as educational, presented as games, and are fun to use. Although promising, many of these tools ignore the evidence-base on effective interventions, and few studies have tested their efficacy. Moreover, the decades of research on effective interventions for the struggling reader are not being incorporated into this technological boom.

We begin with a brief review of the landscape of educational technologies (EdTech) for literacy. With this foundation laid, we introduce a novel program – *Sound it Out* – and the questions that this dissertation will address.

1.1 A REVIEW OF THE EDTECH LANDSCAPE

Since the late 1960s, the technological revolution has spurred innovation in education and literacy. Researchers and educators, motivated by the novelty and potential of information technology, have looked for ways to leverage computation in the classroom and, in theory, improve the way we teach our children. After decades of innovation, the results are mixed; there exists a juxtaposition of a plethora of options with a disappointing lack of evidence base, especially in effectiveness for struggling readers.

1.1.1 *Computer-assisted instruction*

The most prolific area of research in the world of educational technologies for literacy has been in computer-assisted instruction (CAI). CAI consists of three levels of instruction: drill and practice, tutorial, and dialogue. Through drill and practice, students are given reinforcement of existing concepts and skills. Tutorials provide clarification of important concepts, and dialogue consists of active feedback mechanisms to scaffold learning. Combining all three with extensive professional development so that the program supports classroom instruction was the ideal.

Research, however, tells a different story. In their 2013 review, Alan Cheung and Richard Slavin reviewed 20 studies of CAI tools for struggling readers in grades 1-6. They found the largest effect sizes (0.32) for small-group, tutorial integrated applications, moderate effects for supplementary models (0.18), and minimal effects for comprehensive models that aimed to provide classroom management in addition to instruction (0.04) (Cheung & Slavin, 2013).

In theory, CAI aims to aid the delivery of learning content in the classroom, but research has consistently found limited effects (Becker, 1992; Blok, Oostdam, Otter, & Overmaat, 2002; Cheung & Slavin, 2013; Kulik & Kulik, 1991; Slavin, Lake, Davis, & Madden, 2010). These results are not encouraging, but researchers emphasize that the predominant issues facing CAI are a direct result of a lack of rigorous research, including small sample sizes, lack of control groups, and limited research in general (Blok et al., 2002; Cheung & Slavin, 2011, 2013; Guernsey & Levine, 2015; Soe, Koki, & Chang, 2000; Stetter & Hughes, 2015).

1.1.2 *Tablets, mobile computers, and the era of screens*

Although research has mainly focused on publisher/researcher-developed CAI for classrooms, the most prolific medium for EdTech is with consumer technologies (i.e. tablets, smartphones, etc.). The marketplace for educational apps and services is expansive and unregulated, leading to widespread concern about the legitimacy of EdTech for reading education (Alexander, 2020; Greenfield, 2009; Wolf, 2018; Zuckerman, 2017).

In their book, “Tap, Click, Read: Growing Readers in a World of Screens”, Lisa Guernsey and Michael Levine layout the challenging environment that exists for evidence-based practice in the app marketplace (Guernsey & Levine, 2015). They contend that the standards for being labelled as educational are lax, and the relationship between popularity and adherence to the evidence-base is disappointing. In a 2012 analysis of commercially available reading software programs, Amy

Grant and her team echoed these sentiments with the finding that 30 of the most popular apps targeting early literacy skills were very low quality and provided inconsistent scaffolding to produce potential gains (Grant et al., 2012). Grant looked at several key content areas: print knowledge, alphabetic knowledge, phonological awareness, grapheme-phoneme correspondence, phonics, syntactic awareness, decoding, fluency and text comprehension, and their representation across popular apps. In a thorough coding endeavor, she found that the apps were immensely entertaining for children, but that they fell short of providing sufficient scaffolding to transfer or extend to higher-level skills (Grant et al., 2012). Namely, alphabetic knowledge was the most represented content feature across apps, but none of the apps provided the next step to develop other skills important for literacy (i.e. phonemic awareness).

Despite these results, Grant and others understand there is much promise in the digital world for highly accessible tools for struggling readers, and that there are some comprehensive, high-quality apps that are available and gaining momentum (Barzillai, Thomson, & Mangen, 2017; Hutchison, Beschorner, & Schmidt-Crawford, 2012; Stetter & Hughes, 2015; Takacs, Swart, & Bus, 2015; Wolf, Gottwald, Breazeal, Galyean, & Morris, 2017).

1.1.3 *EdTech affordances for emerging readers*

From the rich history of reading research with and without technology, we have learned a lot about what works, what does not work as well, and what to avoid when we design programs for emerging readers. The true magnitude in the number of programs that exist is non-trivial (Guernsey & Levine, 2015), and there persists a disconnect between what research shows is effective and what is introduced into schools (Hurford & Hurford, 2016; Seidenberg, 2017), but research has also revealed essential considerations for a successful application of EdTech for struggling readers:

1. **We cannot replace the teacher.** Generally, and from what we have learned in the last century: a good teacher with a bad program is often better than a bad teacher with a good program (Chall, 1983). We can, however, help teachers to focus their attention in the most effective ways. Teachers are interested in incorporating technology into the classroom (Hutchison et al., 2012), and must be considered an essential stakeholder in program success (Kim, McKenna, & Park, 2017).
2. **We cannot do everything in one program.** There is no one solution for every child, and the prospect is impractical. The most expansive programs appear to achieve this end, but their success is in their composition of many, interweaving elements, not an attempt at a one-size-fits-all approach (Cheung & Slavin, 2013). Instead, the focus should be on individualization and optimization to create a program that grows and adapts (Ok & Rao, 2019).
3. **We must motivate and ensure use in the home with parents/family.** With practice, repetition, and reinforcement, a consistent affordance of technology is its ability to routinize exposure and provide limitless practice (Gee, 2003). Moreover, families of struggling readers are always looking for ways to help their child. Instead of being told time after time to read with their child, technology should instead provide them with the tools to support their child's learning. Even with the ubiquity of technology, however, there are still many barriers to access (Bowles, 2018; Macgilchrist, 2018) for low income, and minority children. A major premise, then, should be for educational technologies to provide affordable and accessible solutions.
4. **We must focus on what we know works.** Literacy and reading disabilities are the most studied of the learning disabilities. Technology must take advantage of the pedagogical

techniques that have proven effective (Grant et al., 2012). An individual program doesn't have to do them all, but an effective technology should incorporate one or a combination of the following: phonics, sight words, morphology, strategies, syntax, semantics, orthography, grapheme-phoneme correspondence, and letter knowledge.

5. **We must account for development.** Research sustains early intervention as the gold-standard approach, but also that older children are not a lost cause (Suggate, 2010; Vaughn et al., 2008). Effective methods for instruction are constant as children age, but the proper mix of foundational skills and generalization strategies are not. For older, struggling readers in 3rd and 4th grade classrooms, this doesn't mean that the whole class should return to the building blocks; instead, the technology should provide the reader with the foundational scaffolding needed to keep up with their peers and constantly adapt to change.
6. **We must make whatever we do fun and engaging.** When we enroll a struggling reader in an intensive intervention, we are asking them to perform hours and hours a day of what they find to be most difficult. Technology is unique in its ability to gamify learning and make even the most menial task entertaining (Kahne, Middaugh, & Evans, 2009). If we incorporate essential elements of video game design, such as interactive feedback and adaptive control (Benton, Vasalou, Berkling, Barendregt, & Mavrikis, 2018), with content focused on the needs of challenged populations (Benton, Vasalou, Khaled, Johnson, & Gooch, 2014), we can make highly motivating, effective tools.
7. **We must rigorously test the tools we develop.** To keep pace with the commercialization of educational programs and provide tools for parents and educators, research must use carefully designed experiments to examine efficacy (National Reading Panel, 2000). No

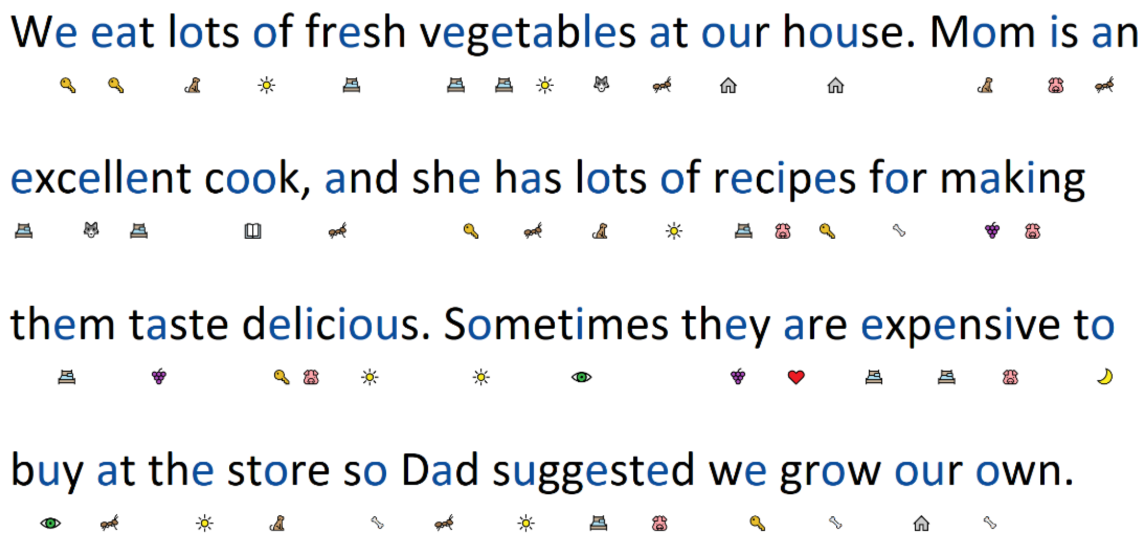
solution will work for everyone, but there is a pressing need for more evidence about what works for what children under what conditions (Morris et al., 2012).

8. **We must design technologies based on the needs of struggling readers.** Just as literacy instruction has been informed by best-practices for struggling readers, educational technologies for literacy must be informed by research on challenged populations. The theoretical premise of many comprehensive and supplemental CAI programs is sound, but their effects are limited because they are designed for the typical reader. Improving upon existing frameworks with a 'Diversity for All' paradigm is a necessary element of the future (Benton et al., 2014).

1.2 *SOUND IT OUT*: A COLLABORATIVE VENTURE

In partnership with a team of engineers and scientists in the Advanced Reading Technologies Team at Microsoft, we developed *Sound it Out*, a symbolic annotation tool to support vowel recognition. Depicted in Figure 1.1, *Sound it Out* provides a phonemic image cue below each vowel; for example, in the word 'lab', the 'a' will appear in blue font with the image of an ant just below. If a child struggles, they can click on the ant to hear the auditory cue, "ant, /æ/" and learn that that the 'a' in 'lab' makes the sound, /æ/. Compared to digital text-to-speech alternatives, *Sound it Out* incorporates a strategy that reinforces foundations of literacy.

Figure 1.1. Example of the *Sound it Out* annotation



Example text with Sound it Out phonemic image cues. The text is an excerpt of "Planting a Garden" from the DIBELS 6th Edition assessment battery (Good & Kaminski, 2007).

Motivating the development of *Sound it Out* is previous research which has highlighted digital tools' flexibility to manipulate the visual-spatial presentation of text. Research in our lab and elsewhere shows the efficacy of increased spacing, both between lines and between letters in words for improving reading speed for struggling readers (Joo, White, Strodtman, & Yeatman, 2018; Legge, Mansfield, & Chung, 2001). These simple manipulations, for a subset of individuals, can immediately impact reading performance, producing an adaptation with instant benefit. As an extension of this work, *Sound it Out* utilizes a simple manipulation, but in doing so embeds an element of pedagogy to target a skill that is particularly difficult for emerging readers: vowel decoding.

As will be discussed in the chapters that follow, *Sound it Out* is designed to: (1) harness evidence-based techniques for teaching sound-letter correspondence, (2) encourage autonomy and motivated use for independent learning, (3) empower children and families in shared reading experiences, (4) supplement classroom instruction, and (5) provide a flexible tool that will adapt to and empower a reader's sociolinguistic identity.

1.3 Outline of the Dissertation

This dissertation details a series of studies that were performed to test the efficacy of *Sound it Out* as a meaningful tool for emerging readers. These studies are of significance to our further understanding of the role of technology in reading education but are also relevant to a future of similar partnerships between developers and researchers in EdTech.

Three experiments are presented in this dissertation, involving a total of 151 participants between the ages of 7 and 15:

- **Determining whether *Sound it Out* can effectively aid struggling readers more accurately decode novel words.** This study represents proof-of-concept work for a phonemic image cue to scaffold vowel decoding for both isolated word decoding and connected-text reading.
- **Characterizing *Sound it Out* as a spontaneous affordance or a practiced intervention for reading performance.** This work aims to determine if *Sound it Out* provides spontaneous benefit or requires extended exposure for realized gains in decoding accuracy.
- **Investigating the role of extended exposure and supervised practice on generalizable improvement in decoding skill.** This study extends the intervention period and provides prescriptive practice to ascertain the context for implementation.

Chapter 2. PROOF OF CONCEPT

This chapter is published as Donnelly, et al. (2020) with co-authors Kevin Larson, Tanya Matskewich, & Jason Yeatman in PloS ONE.

An advantage of digital media is the flexibility to personalize the presentation of text to an individual's needs and embed tools that support pedagogy. The goal of this study was to develop a tablet-based reading tool, grounded in the principles of phonics-based instruction, and determine whether struggling readers could leverage this technology to decode challenging words. The tool presents a small icon below each vowel to represent its sound. Forty struggling child readers were randomly assigned to an intervention or control group to test the efficacy of the phonemic cues. We found that struggling readers could leverage the cues to improve pseudoword decoding: after two weeks of practice, the intervention group showed greater improvement than controls. This study demonstrates the potential of a text annotation, grounded in intervention research, to help children decode novel words. These results highlight the opportunity for educational technologies to support and supplement classroom instruction.

2.1 INTRODUCTION

The discrepancy between the value that society places on literacy and reading achievement levels in American youth (National Assessment of Educational Progress, 2006) is a source of concern both among policy makers and scientists (Moats, 1999; Snow, Burns, & Griffin, 1999). Developmental dyslexia, a learning disability that impacts reading, is widespread, affecting between 5 - 17% of the population (Lyon et al., 2003; Wolf, 2007). Beyond dyslexia, poor literacy rates are a nationwide issue, with 34% of fourth graders performing below the Basic level on national achievement tests (National Assessment of Educational Progress, 2019). Together, these results paint a troubling landscape of literacy achievement and illuminate a non-trivial need for expanding access to evidence-based instruction and intervention.

For many families of struggling readers, access to high quality, evidence-based interventions outside of school are not only limited, but also represent a significant financial burden (Delany, 2017). Even with a diagnosis, children with reading disabilities struggle to find the support they need in their typical classrooms, necessitating supplemental, after-school programs (Kristen D. et al., 2018). For this reason, and with the ever-growing landscape of educational technologies, families are turning to digital alternatives. Many apps and technologies are now widely marketed to augment, or even replace, the teacher in delivering intervention for struggling readers. These new tools are advertised as educational, presented as games, and, in general, are fun to use. Although promising, many of these tools ignore the evidence base on effective instruction and intervention techniques and, of the many educational technologies that are available today, very few have scientific studies testing their efficacy (Cheung & Slavin, 2011; Guernsey & Levine, 2015; Shaheen & Lohnes Watulak, 2019).

In 2012, it was estimated that hundreds of thousands of educational apps had been released on the Apple iOS app store (Guernsey & Levine, 2015). According to a RAND report, children between the ages of three and five years-old spend, on average, four hours per day interacting with communications technology (i.e., smart phones, tablets, etc.) (Daugherty, Dossani, Johnson, & Wright, 2014). The Joan Ganz Cooney Center reported, based on a survey of parents, that 35% of children aged two to ten years-old use educational apps at least once per week. Furthermore, federal funds are being devoted to bringing Internet and devices to our nation's schools, providing infrastructure for even greater involvement with digital media in education. However, despite the exciting potential offered by educational technology, parents and educators alike feel overwhelmed by the plethora of options and the lack of guidelines surrounding technologies advertised as educational (Guernsey & Levine, 2015).

Decades of scientific research into the behavioral and neural mechanisms of literacy learning has led to the development and testing of effective intervention programs for struggling readers, and established comprehensive guidelines and best-practices for implementation of an effective curriculum (Castles, Rastle, & Nation, 2018; Maureen W. Lovett et al., 2017; National Reading Panel, 2000; Snow et al., 1999). Unfortunately, these evidence-based practices (e.g., phonemic awareness, phonics) are largely not being incorporated into the current technological boom. For example, a consistent finding in the intervention literature is that children with dyslexia benefit from direct instruction in phonological awareness and curricula that make clear links between orthography and phonology whereas children with stronger reading skills can often infer grapheme-phoneme correspondences without direct instruction (for review of the extensive literature on the importance of phonics/phonemic awareness see (Bradley L & Bryant P E, 1983; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Frijters et al., 2017; Scanlon & Vellutino, 1997; Storch & Whitehurst, 2002; J. K. Torgesen et al., 2001; Vaughn et al., 2008; Wanzek et al., 2018)). From a sample of 184 apps compiled from online lists of award-winning or highly-rated apps, researchers discovered that although they were entertaining, they lacked scientific backing and "their content, design, production, and distribution are [...] an incomplete response to children's literacy needs, especially for struggling readers"(Guernsey & Levine, 2015). Thus, there is great need for researchers studying literacy development and reading disabilities to work with tech developers on the design of tools that are grounded in the extensive scientific literature on what works for struggling readers, systematically test their effectiveness, and contribute to the development of standards of practice for educational apps targeted at literacy.

Recent metanalyses demonstrate much promise for digital solutions in the context of literacy, yet also describe the multitude of ways that technology is an inappropriate substitute for many aspects of pedagogy (Cheung & Slavin, 2013; Slavin et al., 2010). Namely, these meta analyses demonstrate that technologies focused on supplementing what is provided in the evidence-based classroom (i.e. explicit phonics), rather than restructuring at the classroom level, have demonstrated the most promise in the digital landscape. These findings, however, should be interpreted with caution as the authors further contend that the preponderance of studies in this area are characterized by small samples and poor study design (Cheung & Slavin, 2013). In parallel with research outside of technology, the onus for researchers is to rigorously test digital solutions to discover "what works best in which programs for what students under which conditions" (Morris et al., 2012).

As a means of scaffolding learning in the classroom, technological tools provide limitless practice (Clark, Tanner-Smith, & Killingsworth, 2016; de Souza et al., 2018; Laurillard, 2016) that can be individualized (Cheung & Slavin, 2013; Dowker, 2005; Rose & Strangman, 2007) to optimize for an individual reader's strengths and weaknesses. Additionally, modern computational tools utilizing speech recognition and synthesis application programming interfaces (APIs), allow for embedded tools to be provided in real-time for any given text. One promising avenue for such technology has been the use of embedded support features such as visual images to facilitate learning sound-symbol correspondence. For example, Trainertext, a program in the United Kingdom which provides visual mnemonics above each phoneme in a given text, saw significant improvement after 10 months of exposure and at-home practice (Messer & Nash, 2018). By providing a visual scaffold, the authors concluded that readers could leverage a phonics-based text adaptation to improve their decoding skill. Moreover, SeeWord Reading app, a digital tool which uses a picture-embedded font to demonstrate grapheme-phoneme correspondence (Seward, O'Brien, Breit-Smith, & Meyer, 2014), has demonstrated efficacy in studies performed both in Singapore and the United States (O'Brien, Habib, & Onnis, 2019). Both Trainertext and SeeWord Reading, demonstrate the utility of an embedded support to provide concrete visual relationships between written text and spoken language: a foundational skill for literacy. Together these studies demonstrate that technologies focusing at the level of the phoneme and the syllable (or rime) not only benefit reading performance, but also adhere to known learning and pedagogical principles (Kyle, Kujala, Richardson, Lyytinen, & Goswami, 2013; O'Brien et al., 2019).

This paper outlines a collaboration between the Brain Development & Education Lab at the University of Washington and the Learning Tools team at Microsoft to develop *Sound it Out*, a web-based app that annotates text with phonemic image cues to assist in decoding. This tool is the product of collaborative goals to: (1) create an app informed by the literature – in this case, explicit phonics instruction (National Reading Panel, 2000), (2) focus on an adaptation over an intervention – a supplemental tool that would assist, not replace the teacher, and allow children to bring skills from the classroom to at-home practice with reading (Chall, 1983; Hutchison et al., 2012; Kim et al., 2017), (3) design a fun and whimsical interface that children would want to use (Morris et al., 2012; Wolf et al., 2009), and (4) enable children to confront their challenges and build the skills to decode more complex words. Our focus on explicit phonics instruction is based on decades of research detailing its importance in literacy education (Maureen W. Lovett et al., 2000; Mandel Morrow & Asbury, 1999; Morris et al., 2012; Oakland, Black, Stanford, Nussbaum, & Balise, 1998; Scanlon & Vellutino, 1997; Wanzek et al., 2018), and the unique role that technology affords to provide limitless exposure and practice inside and outside the classroom. The tool was designed through collaboration between the researchers at the University of

Washington (P.M.D. and J.D.Y) and Microsoft (K.L., T.M.), and then tested (independently) in a laboratory study using a pre-registered randomized control trial (RCT) design to determine whether a conceptually simple digital tool can lead to improved word reading outcomes for struggling readers.

As reflected in the preregistered report (available at https://osf.io/q8tpz), the study aimed to answer the following questions: Can struggling readers use phonemic cues to improve reading fluency?, Do struggling readers benefit from phonemic cues when decoding difficult words without time constraints?, and Can we predict which individuals will benefit from the tool based on a standardized battery of reading-related assessments? We hypothesized that the phonemic cue would aid struggling readers in more accurately reading short passages, and with repeated reading, will increase their reading rate. Further, we hypothesized that the phonemic cue would aid struggling readers in more accurately decoding individual real and pseudo-words and that this benefit will relate to the amount of practice they have had with the tool. Finally, regarding individual differences, we hypothesized that those struggling readers that have a specific impairment in phonological processing would benefit the most from the support provided by this tool.

2.2 Methods

2.2.1 *Pre-registration*

The methods, including study design, hypotheses, and analysis plan, were pre-registered using the Open Science Framework (OSF) open-access, pre-registration pipeline. We obtained initial reviews and feedback on this pre-registration from an independent OSF reviewer, revised and re-submitted our methodology, and then adhered (with some minor deviations) to this pre-registered plan throughout the duration of the study. Deviations are noted and explained within this manuscript and are compiled in a ‘Transparent Changes” document in the project repository. Documentation is available at https://osf.io/q8tpz.

2.2.2 *Participants*

Forty children between the ages of 8 and 12 were recruited from the University of Washington Reading & Dyslexia Research Program database, an online repository of families interested in dyslexia research in the Puget Sound region. The participants (19 females; 21 males) were classified as struggling readers based on a battery of behavioral measurements administered within one year prior to participation in the present study. Here, we use the term “struggling reader” rather than “dyslexia” because there is substantial variability in diagnostic criteria for dyslexia and our goal was to design a tool that would support literacy development for anyone that was struggling, regardless of a dyslexia diagnosis. To be considered a struggling reader, participants needed to have reading skills that were more than one standard deviation (SD) below the mean on either the Woodcock-Johnson IV Tests of Achievement Basic Reading Skills composite (WJ BRS) or the Test of Word Reading Efficiency - 2 Index (TOWRE Index), and scores above 1 SD below the mean on the Wechsler Abbreviated Scale of Intelligence Full Scale-2 composite (WASI FS-2). A threshold of 1 SD, rather than the 1.5 SD threshold defined in our preregistration, was adopted to better account for the heterogeneity in the struggling reading population and to expedite participant recruitment. Further, phonological processing abilities were measured for each of the participants using the Comprehensive Test of Phonological Processing – 2 (CTOPP 2) but was not used as an enrollment criterion in the study. Together with age and IQ (WASI-II), CTOPP-2 scores were collected for the purpose of analyzing individual differences in intervention effects to determine if the tool is particularly effective for subjects with certain characteristics. Participants were previously screened for potential speech/language/hearing disorders, neurological impairments, and psychiatric disorders and had none. ADHD was not a disqualifying factor as there is a high co-occurrence with reading disability (Pennington, 2006). In our sample 12 children had a diagnosis of ADHD (6 Control, 6 Intervention). Demographic information on the sample can be found in Table 2.1.

Table 2.1: Summary statistics for the study sample

	Intervention (n=20) *Mean (SD)*	**Control** (n=20) *Mean (SD)*
Age (years)	10.34 (1.3)	9.79 (1.1)
Female (proportion)	0.5	0.45
WJ Basic Reading Skills Composite	78.5 (12.92)	81.15 (8.43)
TOWRE-2 Index Composite	70.25 (7.48)	70.65 (7.10)
WASI Full-Scale 2 Composite	97.6 (8.75)	100.15 (17.20)
CTOPP Phonological Awareness Composite	87.2 (9.7)	83.4 (11.94)
CTOPP Rapid Naming Composite	79.05 (8.57)	77.75 (6.7)

See Methods for descriptions of the individual characteristics. For each characteristic, the mean is provided with the standard deviation within parentheses. Independent t-tests - and Wilcoxon signed-rank tests for Age/Gender - demonstrated no significant differences across all characteristics.

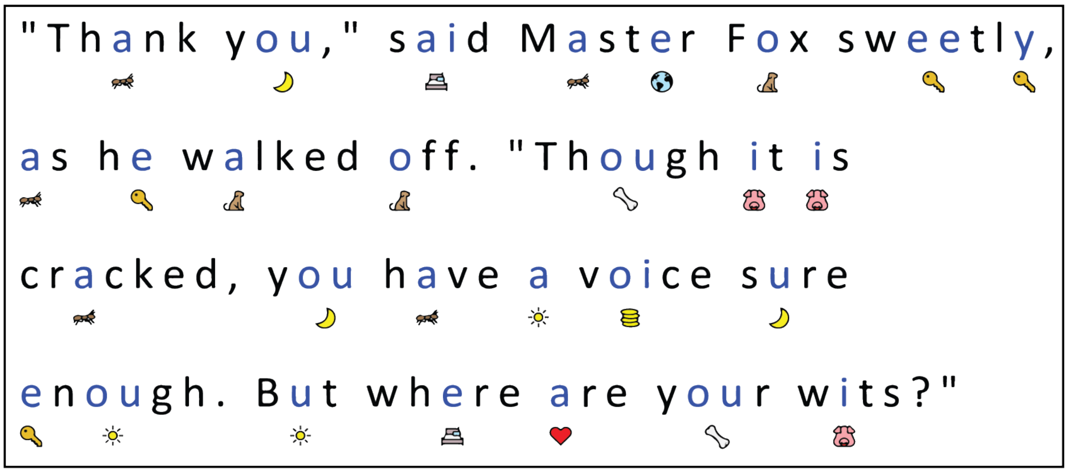
The parents of all participants in the study provided written and informed consent under a protocol that was approved by the University of Washington Institutional Review Board and all procedures, including recruitment, child assent, and testing, were carried out under the stipulations of the University of Washington Human Subjects Division.

2.2.3 *App design*

Sound it Out is a web-based application (web app) that annotates text passages with visual phonemic cues to assist decoding. When a passage is viewed using the web app, the vowels appear in blue font with the image cues (located just underneath) indicating the associated phoneme. Each image cue is a highly recognizable symbol whose name contains the target vowel sound for the vowel above. For example, in the word “cow”, “o” would appear in blue, with the symbol of a house below. The house cues the child that the letter “o” in “cow” makes the same sound as the /aʊ/ in word “house”. Fig 1 shows three sentences taken from a grade level passage and annotated with *Sound it Out*.

Figure 2.1. *Sound it Out* example text with legend.

A



B

IPA	Cue Word	Cue Symbol	IPA	Cue Word	Cue Symbol	IPA	Cue Word	Cue Symbol	IPA	Cue Word	Cue Symbol
/ɑ/	Heart		/ɝ/	Earth		/aʊ/	House		/oʊ/	Bone	
/æ/	Ant		/eɪ/	Grape		/əl/	Wolf		/ɔɪ/	Coin	
/ʌ/	Sun		/ɪ/	Pig		/aɪ/	Eye		/ʊ/	Book	
/ɔ/	Dog		/i/	Keys		/ɛ/	Bed		/u/	Moon	

The vowels in this excerpt from Aesop’s ‘The Fox and the Crow” fable appear blue with phonemic image cues provided below. The name of the symbols cues the reader to the sound of the target vowel.

To aid in symbol recognition and retention, the app also integrates a voice cue; when a child presses the phonemic cue symbol, a voice narrates the symbol name followed by an isolated presentation of the target vowel sound. For example, in the “cow” example, when a child presses the image of the house, below “o”, a voice will say, “house, /aʊ/”. The vowel sounds were recorded by a native English speaker with training in phonetics. The recordings were judged by the three native English-speaking authors to be typical examples of the given vowel sounds and, during the training period, participants were exposed to all the vowel sounds and were able to correctly identify each vowel.

The goal of this app is for the cues to provide helpful hints that aid in decoding and provide children the support they need to attempt to decode difficult words and, eventually, learn the highly inconsistent grapheme-phoneme correspondences of vowels in English. Instead of simply reading challenging words, as is typical in a speech-to-text tool (Olson & Wise, 1992; Takacs et al., 2015), *Sound it Out* focuses on a particularly difficult task for struggling readers: vowel decoding. Because vowels in English represent the majority of 'mutually interfering discriminations' - or multiple sound associations - that young readers must master (Bryant, 1965), learning the rules for vowels represents a significant portion of phonics curricula (Castles et al., 2018; Hornung, Martin, & Fayol, 2017; Moats, 1999). With the understanding that vowels may not be the only challenge for the reader, *Sound it Out* was designed to provide a tool that scaffolds learning and empowers struggling readers to read more complex passages independently. *Sound it Out* was designed as a feature that can be turned on or off when a child is reading. In this study, both intervention and control participants were given a tablet and taught how to use it for reading. The primary manipulation was whether the *Sound it Out* feature was turned on. Beyond this manipulation, the text in the table-based reader was identical.

2.2.4 *Procedure*

2.2.4.1 Study design

In a randomized pre-post design, participants were randomized to a control or intervention condition. Randomization was unconstrained with group assignment determined at time of consent; however, sibling participants were assigned to the same group to better control participant adherence. Both intervention and control participants completed an initial, baseline session that collected all outcome measures using the normal text condition presented on a Kindle fire tablet without the *Sound It Out* cue (see **Outcome Measures**). Control participants completed a brief training in use of the tablet then a two-week, at-home practice program without the *Sound It Out* tool. Intervention participants completed the full training period that included both the use of the tablet as well as a formalized introduction to and practice with the *Sound It Out* cue prior to an identical at-home practice program with the tool turned on. After two weeks, all participants returned for a post-intervention session where the intervention participants were tested with the cue, while control participants experienced the normal text condition for all study stimuli. As the goal of this study was to test proof-of-concept for a digitally embedded phonetic cue in a small scale RCT, generalization to un-cued reading in the intervention group was not tested.

2.2.4.2 Training program

For the intervention group, each child was first oriented to the presence of the cue in an example passage. The researcher would show a passage with the cue and say the following: “Now we are going to use a cool tool that we made to help you read the tricky words. Underneath each word we have symbols that help you figure out the sound that the blue letters make.” The researcher would walk through a word and demonstrate how the cues could be used to help decode words. Then the researcher explained when to use the cues: “When you come to a word you don’t know, just look at the symbols and that should help you figure out the sounds that the blue letters make.” The child was then instructed to read through the passage. When they came to tricky words the researcher alternated between clicking on the symbols and naming the symbols to help sound out the words that cause difficulty. If the child read through the first example passage with ease, a more challenging passage was added to ensure that the child had demonstrated efficient and correct usage. Then, with the use of flashcards, the researcher reviewed each symbol explicitly with the child.

For the control group, participants were introduced to the tablet and instructed on how to navigate to the various passages for at home practice. They were not shown the phonemic cues, but otherwise followed an identical procedure.

2.2.4.3 At-home practice

After training with the application, and when the baseline testing session was completed, participants were provided tablets to take home for reading practice. Participants were asked to read at least one story per day over the course of two weeks using the app (at home). For the intervention group, the *Sound it Out* feature was enabled so that phonemic cues showed up in the passages. For the control group, text was rendered in the same font but without the cues. Tablets were pre-loaded with 36 first, second, and third grade supplementary passages for children to read with or without their parents. To encourage meaningful practice, parents were provided with a brief introduction to the app prior to taking home the tablet and were given instructions to allow the child to work through difficult words using the image cues prior to providing any additional guidance. For those that adhered to these instructions, and assuming each passage would take approximately ten minutes to complete, each child experienced at least 100 minutes of exposure over the two-week practice period. Adherence to the practice schedule was measured via short, three-question comprehension quizzes completed after reading a story (through a web interface). A participant was only credited for a practice passage if they received a comprehension score of at least two.

Supplementary passages for at-home reading were from ReadWorks.org, an online library of grade-level passages (used with permission of ReadWorks). Passages were phonetically coded manually by the research team based on the most common pronunciation in the Pacific Northwest dialect of English. The code was verified by both P.M.D & K.L., with inconsistencies discussed and decided via consensus. All passages were displayed on an Amazon Kindle Fire 8 tablet at a set font size and resolution. Comprehension questions, to gauge practice adherence, were created by a trained, certified teacher at a local school that specializes in working with children with developmental dyslexia and were grade-level matched to the individual passages.

2.2.5 *Outcome measures*

Measures of reading performance were collected at baseline and after the two-week period of practice. Real and pseudo word decoding accuracy was measured by having participants read lists of words that were loaded into the web app. Four unique lists of 30 real and 30 pseudo words were created with two lists being delivered at each session. All lists were developed using the orthographic wordform database MCWord (Medler & Binder, 2005). Real word lists consisted of the most frequent words in the English language with five instances each of three-letter to eight-letter words. Pseudo word lists consisted of the most frequent bigrams in English with an identical progression from three-letter to eight-letter pseudo words. (See Table A.2 and Table A.3 for detailed word statistics). All lists were unique and were rated for consistent difficulty using timed reading in ten typically reading adults. Lists were administered in a counterbalanced order for participants in each group. Lists were phonetically coded manually by the first author based on the most common pronunciation in the Pacific Northwest dialect of English. Accuracy in pronunciation was not limited to the code used for the phonemic cues but extended to acceptable pronunciations in English. At the start of administration, all participants were reminded that they were not being timed and encouraged to read as accurately as possible. For intervention participants exposed to the image cue, participants were additionally reminded that the symbols were there to help them should they come to a challenging word. Post-hoc analysis revealed that performance was highly reliable across the different word lists: performance was highly correlated for the two lists of real words ($r = 0.86, p < 0.001$) and pseudo words ($r = 0.79, p < 0.001$) in each session. Accuracy of real and pseudo word decoding was our primary outcome measure (number of words read correctly on each list akin to the Woodcock Johnson Word ID and Word Attack (both untimed measures)).

Passage reading rate and accuracy was measured by having participants read grade-level passages that were loaded into the web app. All testing passages used were from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) 6th Edition library, used with permission of the University of Oregon Center on Teaching & Learning. These passages are commonly used as benchmark assessments in schools and have been extensively used in reading research. Only passages rated at second, third, and fourth grade were used. For each testing session, every passage was presented by a research assistant in the Brain Development & Education Lab and audio recorded for accurate scoring and coding. As with the decoding measures, participants were reminded at the start of administration that they were not being timed, encouraged to read as accurately as possible, and (for intervention participants) that the symbols were there should they come to a challenging word. Instead of constraining the oral reading to the one-minute limit of the DIBELS protocol, all passages were read to completion twice in a repeated reading design. Each passage reading yielded four measures: accuracy (number of words pronounced correctly) on the first and second read; and rate of reading (number of accurate words per minute) on the first and second read. To test the influence of *Sound it Out* on connected text reading, analyses focused on the word-reading accuracy of the first read and the word-reading rate of the second read. Second reading rate was used to control for the added time that might be associated with using the phonemic cues to sound out difficult words on the first read. Testing passages were coded and rated in the same manner as the at-home practice passages.

2.2.6 *Statistics*

Due to the presence of missing data, data were analyzed with linear mixed effects (LME) models, as specified in our pre-registration. Missing data consisted of word reading accuracy and rate information for twelve passages from seven participants due to testing fatigue or inability to complete the passages. For each outcome measure, we fit an LME model with fixed effects of: (1) time (pre-intervention / post-intervention as a categorical variable); (2) group (intervention / control groups as a categorical variable); (3) the group by time interaction. The models included a random intercept for participant, to account for individual variation in baseline performance. To account for differences between the individual, lab-created word lists, we added a random intercept for word list to those models. Practice data were used to ensure that all participants engaged with the tool at home and were also used in correlational analyses to examine the impact of at-home exposure on improvement.

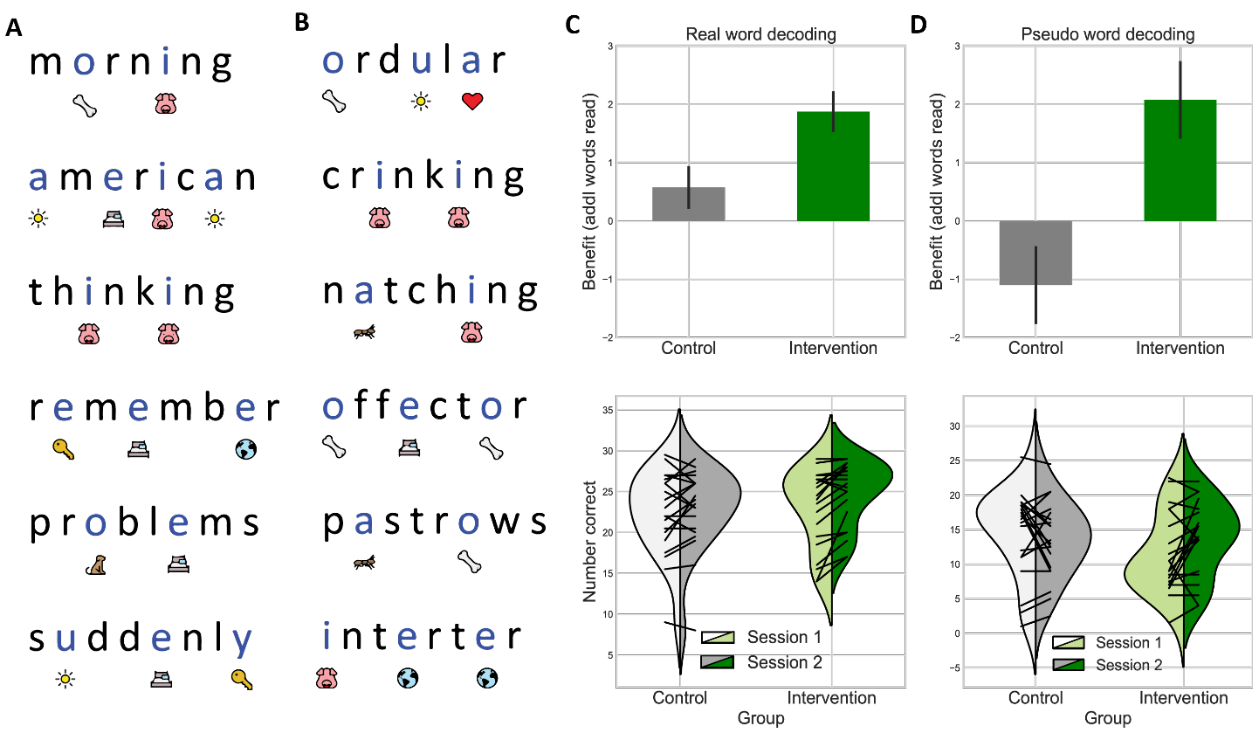
Due to issues collecting reliable usage statistics for the at home reading practice, prediction analyses were not appropriate. Instead, post-hoc correlation analyses were performed using the Pearson correlation coefficient between post-pre difference scores and the three subject characteristics collected at baseline: age, WASI-II, and the CTOPP-2. This analysis differs from that described in the preregistration due to the small number of reading variables collected and inability to collect robust measures of exposure, making methods of dimensionality reduction not appropriate. Analyses were carried out using the NumPy and SciPy libraries of Python and the MATLAB Statistics Toolbox (2019a) (The MathWorks, 2017). All data and analysis code associated with this manuscript are publicly accessible at the following link: [https://github.com/patdonnelly/Donnelly_2019_PLOSONE]

2.3 Results

2.3.1 *Phonemic cues improve decoding accuracy for real and pseudo words*

For our primary outcome measure (as specified in our pre-registration [link]), children were assessed on their ability to decode lists of increasingly more complex real and pseudo words prior to, and immediately following, the two week intervention period. For this measure of decoding accuracy, words were displayed in a list (see Figure 2.2A/B). Figure 2.2 also includes bar plots of difference scores as well as violin plots of the full score distribution.

Figure 2.2. Untimed decoding performance on real and pseudo words.



Example stimuli from the real-word (A) and the pseudo-word (B) lists with Sound it Out phonemic cues below the highlighted vowels. Bar plots show difference scores (number of words read correctly) from the first and second sessions for both the control and intervention groups on real word (C) and pseudo word (D) lists. Bar heights represent the additional words read on the *second session. Error bars reflect +/- 1 SEM. Below the bar plots, violin density plots show group performance on these measures for both sessions with superimposed line plots of individual performance.*

For real-word decoding accuracy, although both groups did show some growth, the group by time interaction was not significant (β=1.3, t(156) =1.923, p=0.056) indicating that the growth in the intervention group was not statistically different from the control group.

For pseudo-word decoding accuracy the group by time interaction was significant (β=3.175, t(156)=2.99, p=0.003) with the intervention group showing significantly greater improvement than the control group (a threshold of 0.0125 was defined in the preregistered report to adjust for multiple comparisons). At pretest, despite randomization, the intervention group by chance had lower scores than the control group: this is evidenced by the significant main effect of group in the mixed effects model.

To determine the influence of practice on individual outcomes, we examined the correlation between at-home practice and individual growth. All participants completed at-home practice (intervention Mean = 13 stories, SD = 6; control Mean = 12, SD = 6), but the correlation between amount of practice and growth was not significant for the intervention group (real words r = -0.14, p = 0.55; pseudo words r = 0.23, p = 0.35) or the control group (real words r = 0.01, p = 0.97; pseudo word r = -0.37, p = 0.10).

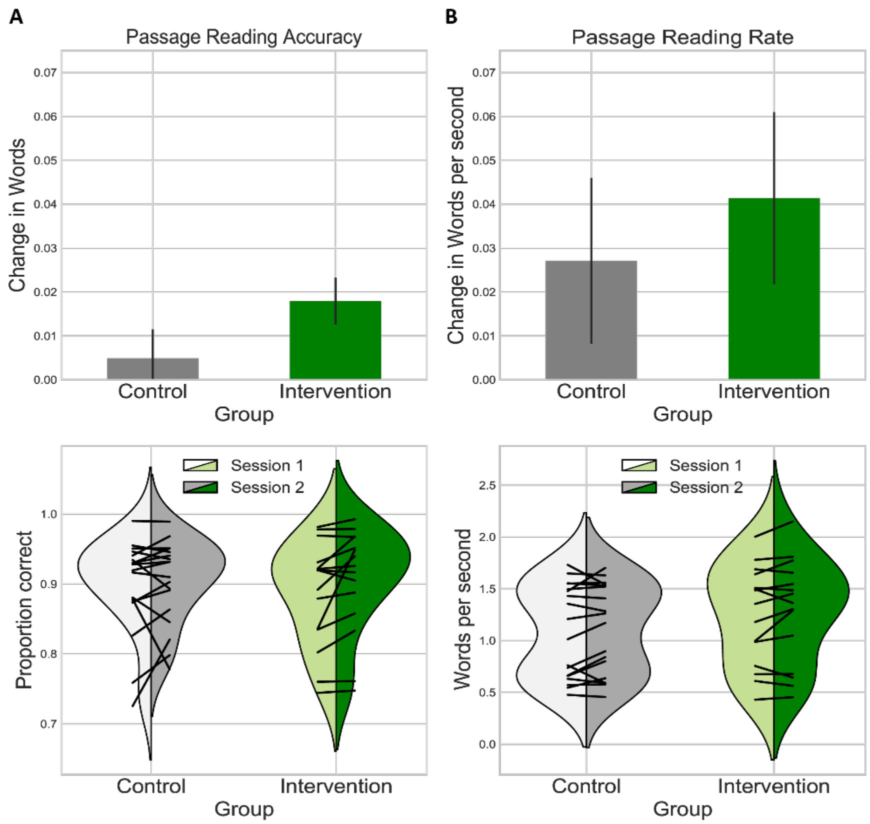
Regarding our final question on subject characteristics that predict individual response, a correlation analysis revealed no significant relationships between our baseline predictors (age, WASI-II Full Scale-2, and CTOPP-2) and improvement in real-word decoding. For pseudo-word decoding, there were negative correlations with the CTOPP-2 Phonological Awareness (PA) (r = -0.52, p = 0.018) and Phonological Memory (PM) (r = -0.48, p = 0.034) composite measures as well as the WASI-II Full Scale – 2 score (r = -0.49, p = 0.027), but these effects were not significant after correcting for multiple comparisons. Due to the heterogeneity of our sample, we tested a model with added covariates for age and initial phonological awareness ability: model fit comparison revealed no benefit to the more complex model and no significant main effects for the added covariates (see A.1.1).

These findings show that without the constraints of time during testing, there was a beneficial effect of access to the phonemic cue for single word decoding. This benefit was observed in the case of pseudoword decoding where children were asked to pronounce novel words in isolation. Moreover, correlation analyses suggest that this effect is more pronounced for those participants with more significant impairments in phonological processing and lower IQ (though these effects did not surpass our adjusted significance threshold of $p < 0.0125$). Although the effect sizes were moderate (Cohen's $d = 0.74$ for pseudoword decoding, $d = 0.57$ for real word decoding), these results suggest that children, with practice, can incorporate a novel cue to scaffold independent decoding.

2.3.2 *Phonemic cues for connected text reading*

To determine whether the phonemic cue confers a benefit for reading connected text, we assessed word reading accuracy and rate on grade level passages before and after intervention. Figure 2.3 depicts bar plots of difference scores and violin density plots for the control and intervention participants in terms of (a) reading accuracy: number of words read correctly in the passage on the first read; and (b) reading rate: number of correct words per minute in the second read.

Figure 2.3. Accuracy and rate for connected text reading.



Bar graphs depict difference scores from the first and second sessions for both the control and intervention groups for word reading accuracy (A) and rate (B). Bar heights represent the additional number of words read correctly and additional accurate words per second on the second session. Error bars reflect +/- 1 SEM. Below the bar plots, violin density plots show group performance on these measures for both sessions with superimposed line plots of individual performance.

For word reading accuracy the group by time interaction was not significant (β=0.014, t(65)=1.1, p=0.275). For word reading rate there was a non-significant group by time interaction (β=0.014, t(64)=0.368, p=0.714). Effect sizes were d = 0.36 for accuracy and 0.13 for rate. To examine the effect of heterogeneity in our sample, we tested a model with added covariates for age and initial phonological awareness ability: model fit comparison and analysis of added fixed effects demonstrated no significant effects of these covariates (see A.1.1).

Correlation analyses revealed only a near-significant negative relationship between age and word reading accuracy ($r = -0.55$, $p = 0.028$) suggesting that younger children may benefit more from *Sound it Out*.

2.4 DISCUSSION

Using a RCT design, we tested the hypotheses that struggling readers could leverage a phonemic image cue placed below the vowels in digitally presented text to improve reading accuracy for isolated words and connected text, and that this benefit would be more pronounced for those readers with lower performance on measures on phonological processing. Data collected after a two-week period of unsupervised (but digitally monitored) practice demonstrated that struggling readers could read more complex words using the tool: compared to the control group, the intervention group showed a significantly larger improvement in decoding accuracy specifically for pseudo-words. As depicted in the results, this benefit did not extend to either measure of connected-text reading (accuracy and rate) and the improvement in real-word decoding did not differ significantly between intervention and control groups.

Although there was no benefit, stable performance on measures of connected text reading was observed for all participants with no significant difference between groups. The lack of benefits for connected text reading might reflect the limited training period or the increased cognitive demands of a novel approach to reading. These are important questions for future studies as generalization to connected text is of key importance.

Correlation analyses, after multiple comparison correction, revealed no significant relationships between our variables of interest and benefit of the cue. Due to unreliable practice data (see Statistics), analyses cannot support any conclusions regarding the relationship between subject characteristics and benefits conferred by phonemic cues. However, results suggest that the tool may benefit those participants who are younger and/or have lower phonological processing scores (see Results). Together, although most analyses failed to meet our adjusted significance threshold, data suggests that participants were able to effectively use the cues in isolated situations (i.e. pseudoword reading), but the tool did not become sufficiently automatic to produce significant gains in passage reading fluency.

Decades of dyslexia research has been devoted to developing and systematically testing interventions designed for struggling readers. In the digital age, devices provide ever-expanding access to a plethora of educational apps and resources advertised as educational. Many families seek out these resources to supplement their child's education. Unfortunately, most of this educational technology lacks scientific backing (Slavin, Cheung, Groff, & Lake, 2008). A goal of this project was to embed a core feature of evidence-based practice in literacy instruction into a digital tool to scaffold learning to read. Inspired by the key tenants of phonics instruction (e.g., explicit and clear instruction in letter-sound correspondence, repeated exposure, and systematic practice), the phonemic cue was designed to provide struggling readers with a hint to aid the decoding of novel words. We focused on vowels because, in English, the highly inconsistent grapheme-phoneme mapping is a major hurdle for struggling readers. In a landscape of digital tools that provide instant corrections at the whole word level, the phonemic cue annotation in this study is one of a only a few learning aids that provides an element of instruction (at the phoneme level) to support generalizable skill (Messer & Nash, 2018; O'Brien et al., 2019; Olson & Wise, 1992; Wise & Olson, 1995). Instead of being given the answer at the first sign of struggle, the child can utilize the phonemic cue to learn part of the word, yet still needs to exercise the building blocks of literacy to get the answer.

In a similar vein to this work, researchers in the UK developed Trainertext, a scaffolding program that provides whimsical, visual mnemonics for scaffolding letter-sound correspondence. Instead of focusing on vowels, Trainertext provides a visual cue above every phoneme in connected text that is associated with short rhyme. For example, above the word “gas”, Trainertext has images related to the phrases “Goat in a Boat”, “The Ant in Pink Pants”, and “The Snake with a Shake” to represent the grapheme-phoneme correspondence of each letter. A RCT with individualized instruction and 10 months of exposure demonstrated significant improvements (Cohen’s d = ~0.80) with the largest effects seen for decoding and phonological awareness (Messer & Nash, 2018). This work demonstrated how phonics-based text annotation could be leveraged by struggling readers to bootstrap their decoding skills, and that, with extended exposure, that benefit could generalize to decoding without the cues. The present study built on this work by (a) employing a simplified symbol set focused on vowel sounds, with the hope of building a tool that would be quicker to learn and less cognitively demanding and (b) could be used immediately without requiring months of a resource-intensive intervention program. Taken together, these two studies emphasize that text annotation is a promising approach, either in combination with an in-person intervention (as in (Messer & Nash, 2018)), or as a tool to support at-home practice (as in the present study).

As was the case with Trainertext, *Sound it Out* requires children to learn a new, albeit intuitively designed, symbol system and practice sufficiently for the associations to become automatic. The symbols chosen were optimized for recognizability, but the challenge remains in teaching the child to associate a portion of the symbol name with a discrete sound segment in an often-unrelated word. Moreover, the sound-symbol association must be fast enough to not impede short term memory with increased cognitive load (Norton & Wolf, 2012; Wolf & Bowers, 1997;

Wolf & Katzir-Cohen, 2001). These two dimensions, effective use and automaticity, are captured by the two areas of measurement: the untimed word lists and the connected text reading.

The finding of improved pseudoword decoding performance indicates that without temporal constraints or the cognitive demands of connected text reading, children may be able to use the phonemic cue to improve decoding performance. This suggests that a brief, two-week practice period was enough and the tool sufficiently intuitive to have an impact. As the limited effects in passage reading accuracy and rate reveal, however, the tool did not extend to situations when time constraints were re-introduced. Either due to the limited practice period, limited supervised practice, or conflict with existing strategies children use when approaching challenging words, children did not similarly benefit from the phonemic cues in connected text. Future studies should incorporate qualitative and metacognitive methods to identify factors and circumstances that encourage struggling readers to adopt a novel strategy.

Albeit promising, these results should be interpreted cautiously: Our power analysis indicated that we only had sufficient power to detect relatively large effects and many of the analyses (e.g., individual differences) were likely underpowered. Also, as there was a significant difference at pre-test for our sole finding with pseudo word decoding, future studies are needed to rule out the role of regression toward the mean and possible ceiling effects. Thus, future studies are needed with larger sample sizes to provide more conclusive results. Moreover, two additional points merit further investigation. First, future studies should more efficiently, and quantitatively and qualitatively, monitor practice adherence and cadence at home to better explore the relationship between exposure and reading-related measures. Second, given the short intervention period, we did not examine generalization to reading improvements without the cue and across different aspects of skilled reading. We only investigated whether the cue could be effectively used to decode more complex words. Thus, examining long-term learning effects and generalization to a variety of different contexts, as well as the role of parental involvement/participation is an important future direction.

We had anticipated that the cue would require limited exposure, but our results are in stark contrast with previous studies that have instituted a more comprehensive, extended training program and observed significant benefit to fluent reading (Messer & Nash, 2018; Joseph K. Torgesen, Wagner, Rashotte, Herron, & Lindamood, 2010). Thus, although the training and practice periods were enough to encourage effective use, they were not enough to ensure automatic and fluent use in a natural setting, and by extension, were insufficient to make general claims on efficacy of *Sound it Out* for supporting long-term growth in reading skills.

On the other hand, the limited effects in accuracy and rate performance suggest that either the cue did not adversely impact reading performance or that it was underutilized given the increased cognitive demands of real-time reading. Many participants in the study have received supplemental instruction previously and have learned strategies for approaching new words. As a novel strategy, the phonemic cue may have been overridden when children were asked to read continuously and for comprehension. In line with the corpus of research on strategy instruction for literacy (Berninger, Lee, Abbott, & Breznitz, 2013; Farkas & Jang, 2019; Maureen W. Lovett et al., 2000; Wolf & Bowers, 1997; Wolf, Bowers, & Biddle, 2000), although the children in our sample demonstrated competency in the use of the cue in isolated, single-word decoding, strategy adoption would require more sustained exposure.

A strength of educational technologies for literacy is their ability to empower parents, teachers and other advocates to support and supplement their child's learning outside the classroom (Dexter & Richardson, 2019; Francom, 2019; Lindeblad, Nilsson, Gustafson, & Svensson, 2017;

Messer & Nash, 2018; Ronimus & Lyytinen, 2015). *Sound it Out* is unique in that it provides a tool that gives parents a strategy for reinforcing phonics principles with their child. Many parents, when confronted with the stress and challenge of raising a child who struggles with learning to read, are told to read more to their child, but not given the knowledge base needed to provide meaningful support (Goodall, 2017). Post-study feedback from parents in the study were overwhelmingly positive, with a majority of parents noting interest in using the tool into the future (see Table A.1). Relatedly, a study by Ronimus and colleagues demonstrated increased efficacy of GraphoGame, a digital literacy program, for reading performance when children engaged with the tool with parental involvement (Ronimus & Lyytinen, 2015). Thus, with parental support and potential alignment with the teacher and in-school curriculum (Auerbach, 1989; Cheung & Slavin, 2013; Hannon & James, 1990; Kraft & Monti-Nussbaum, 2017; Lynch, Anderson, Anderson, & Shapiro, 2006; Slavin et al., 2008; Tichnor-Wagner, Garwood, Bratsch-Hines, & Vernon-Feagans, 2016), *Sound it Out* represents a promising venture bridging research and practice.

In aggregate, these findings represent proof-of-concept for this novel approach to assisting struggling readers by merging the extensive evidence base on effective literacy instruction and the affordances that technology lends to the educational arena. Not only did it prove promising in improving decoding performance with very limited practice, but it also was observed to be non-detrimental to passage reading, meaning that it was not too cognitively demanding. Future research should focus on optimizing training and practice to produce gains that will extend beyond isolated single word decoding and lead to more confident, fluent readers.

Chapter 3. SPONTANEOUS EFFECTS

This work represents Study 1 in a manuscript currently submitted to the journal Reading Research Quarterly with co-authors Liesbeth Gijbels, Kevin Larson, Tanya Matskewich, Paul Linnerud, Patricia Kuhl, and Jason Yeatman

Research on educational technologies for reading instruction is disproportionate to the myriad applications in the marketplace. Here we assess a web-based reading tool, *Sound it Out,* that assists struggling readers in vowel decoding. Created as a collaboration between researchers and technology developers, the tool provides a phonemic-image cue below the vowels to scaffold letter-sound correspondence. This study examined whether *Sound it Out* provides an immediate benefit to reading performance in thirty struggling readers (ages 8-10) randomly assigned to counterbalanced groups in a single visit. Results showed that children were not able to spontaneously capitalize on the cues for improved text reading.

3.1 INTRODUCTION

Initial, proof-of-concept research into *Sound it Out* revealed limited, yet promising effects of the cue on reading performance (Donnelly, Larson, Matskewich, & Yeatman, 2020). In a small-scale randomized controlled trial (RCT), with a two-week practice period, our research group found that the cue had a significant impact on pseudo word decoding but no significant impact on connected-text reading or real word decoding. The pseudoword decoding benefit was small - about two additional items read, on average - however, combined with qualitative survey data which demonstrated substantial interest and engagement, overall findings suggested that, with some modifications, this technology could represent a promising approach to supplement phonics based classroom instruction.

The present work builds on these initial findings to hone the implementation of the technology. Namely, this work details a study that aim to describe the conditions that are necessary for the cue to impact reading performance beyond single word decoding. Extending from our previous research that was ambiguous with regards to the effect of practice on cue benefit, the present study utilizes a single-visit crossover design to determine if we can replicate our previous findings without the practice period. In other words, does the phonemic image provide an immediate benefit to reading? Because the *Sound it Out* cue was so easy to learn, we hypothesized that children could immediately use the cue to sound out difficult words that were above their reading level.

3.2 Methods

3.2.1 *Participants*

Thirty-two children between 8 and 10 years old were recruited from the University of Washington (UW) Reading & Dyslexia Research Program database, an online repository of families interested in reading research in Western Washington. Three participants withdrew from participation during the study, leaving a final dataset of twenty-eight (14 Female, 14 Male) participants. These participants were struggling readers, as characterized by performance on a standardized battery of reading assessments. In order to capture a more heterogeneous and matched sample to the previous study (Donnelly et al., 2020), the classification of ‘struggling readers’ was based on performance 1 SD below the standard norm on either the Woodcock-Johnson IV Test of Achievement Basic Reading Skills (WJ BRS) or the Test of Word Reading Efficiency-2 (TOWRE Index) composite measures, and above 1 SD below the mean on the Wechsler Abbreviated Scales of Intelligence Full Scale-2 (WASI FS2) composite measure. Twenty participants had a diagnosis of developmental dyslexia. Participants were also assessed for phonological processing using the Comprehensive Test of Phonological Processing - 2 (CTOPP), but that test was not used as a screening tool. Participants were screened previously for potential speech/language/hearing disorders, neurological impairments, and psychiatric disorders. Due to high co-occurrence with ADHD, a diagnosis was not a disqualifying factor: nine participants in the sample had a diagnosis. Full demographic information for this sample can be found in Table 3.1. Parents/guardians of all participants in the study were provided with informed consent that was approved by the UW Institutional Review Board. Further, all procedures, including recruitment, assent, and behavioral testing, were performed in accordance with the UW Human Subjects division (HSD).

Table 3.1: Demographic characteristics for study sample

	Participants (n=28)	
	Mean	***SD***
Age (Y)	10.18	1.37
WJ BRS	82.62	11.3
TOWRE Index	73.65	7.8
WASI FS2	101.1	14.5
CTOPP PA	91.77	11.1
CTOPP RSN	85.96	12.8

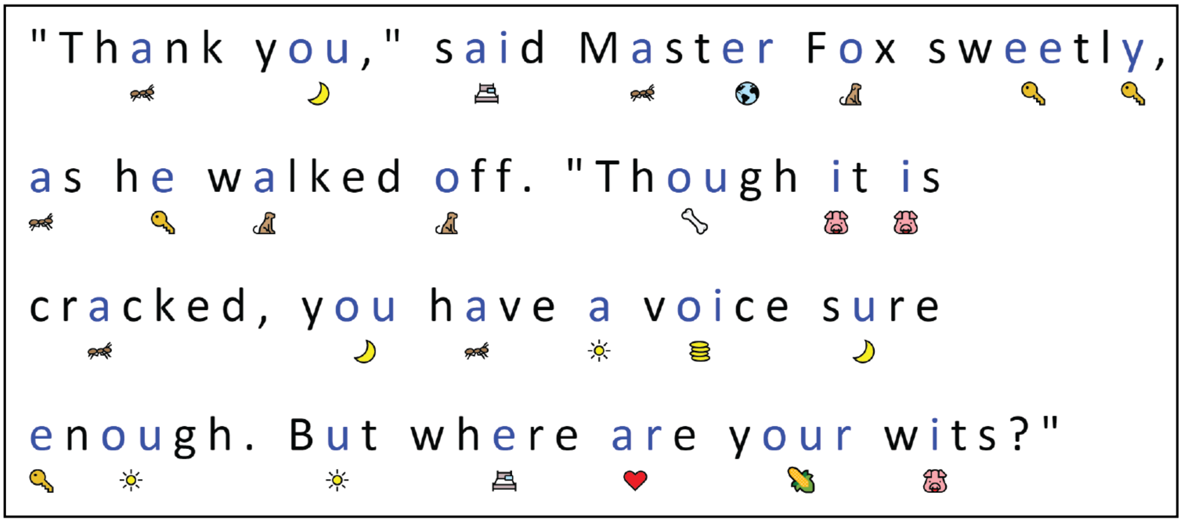
Means and standard deviations are given for each demographic variable of interest for participants. For all variables, other than age, standardized scores are provided.

3.2.2 *App Design*

Sound It Out, designed in collaboration with the Advanced Reading Technologies Team at Microsoft, is a web-based application that provides visual annotations to aid in vowel decoding of live text. When a reader is viewing a text sample in the Immersive Reader tool - in Microsoft Word or OneNote, Sound It Out is provided as an option to turn on/off at the level of the entire sample or an individual word. Selecting Sound it Out will color each of the voiced vowels blue in the text with corresponding phonemic image cues placed below each vowel. Each image cue is a highly recognizable image whose name contains the associated 'target sound' for the vowel directly above the image (an example is provided in **Fig 1**). Moreover, clicking on the image cues will provide the reader with an auditory cue for the symbol including the name of the image and the associated target vowel sound. For example, in the word 'sound', the 'ou' would appear in blue font with the symbol of a house below. The house cues the child that the 'ou' in 'sound' has the same vowel sound as /aʊ/ in word 'house'. If a child clicks on the house, the web-app will announce, "house, ou". A more detailed description of the app used can be found in our previous work (Donnelly et al., 2020).

Figure 3.1: Example of the Sound it Out annotation with legend.

A



B

IPA	Cue Word	Cue Symbol	IPA	Cue Word	Cue Symbol	IPA	Cue Word	Cue Symbol	IPA	Cue Word	Cue Symbol
/ɑ/	Heart		/ɝ/	Earth		/aʊ/	House		/oʊ/	Bone	
/æ/	Ant		/eɪ/	Grape		/əl/	Wolf		/ɔɪ/	Coin	
/ʌ/	Sun		/ɪ/	Pig		/aɪ/	Eye		/ʊ/	Book	
/ɔ/	Dog		/i/	Keys		/ɛ/	Bed		/u/	Moon	
/ɔr/	Corn										

Panel (A) shows a sample of an Aesop's fable passage with the Sound it Out symbolic vowel annotations. Panel (B) provides a legend of the image cues used, their corresponding names and target vowel sounds.

3.2.3 *Study Design*

Using a single-visit, cross-over design, this study aimed to build off previous work to determine if a similar cue benefit can be observed without a period of exposure and practice. In this study, participants were randomly assigned to one of four sets of stimuli presentations with and without the cue. Test stimuli were counterbalanced in presentation order as well as cue condition (normal or cued). Prior to testing, all participants completed a training protocol that was identical to the protocol in the previous work. The protocol involved demonstration of proficiency in identifying the image names, associated target vowel sounds, and accessing the voice cue within the app. A more detailed description of the training protocol can be found in the previous work (Donnelly et al., 2020).

3.2.4 *Measures*

The following measures were collected during the study: real word decoding, pseudo word decoding, passage reading accuracy, and passage reading rate. These measures were collected using identical stimuli to our previous work. Real and pseudo word decoding accuracy were measured using lab-created lists: four, unique lists of 30 real and 30 pseudo words were administered and were developed using the orthographic word form database MCWord (Medler & Binder, 2005). Real words, with high frequency in the English database, were selected with five instances each of three-letter to eight-letter words. Pseudo words were chosen based on high bigram frequency with an identical progression of three to eight-letter words. All lists were unique, rated as similarly difficult and were administered in a counterbalanced order and cue condition during the study. Further information about validity and reliability can be found in (Donnelly et al., 2020).

Passage reading accuracy and rate were measured using grade-level passages from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) 6th Edition database. With permission from the University of Oregon Center on Teaching and Learning, two third grade passages were administered to determine if the phonemic cue would assist readers in connected-text reading. Identical to the prior study, the passages were read to completion and scored for accuracy (number of words correct) and rate (correct words per second) in oral reading. Each passage was read once in a counterbalanced order of presentation and cue condition. All measures were administered by a research assistant in the Brain Development & Education Lab and audio recorded for accurate scoring and coding.

3.2.5 *Statistics*

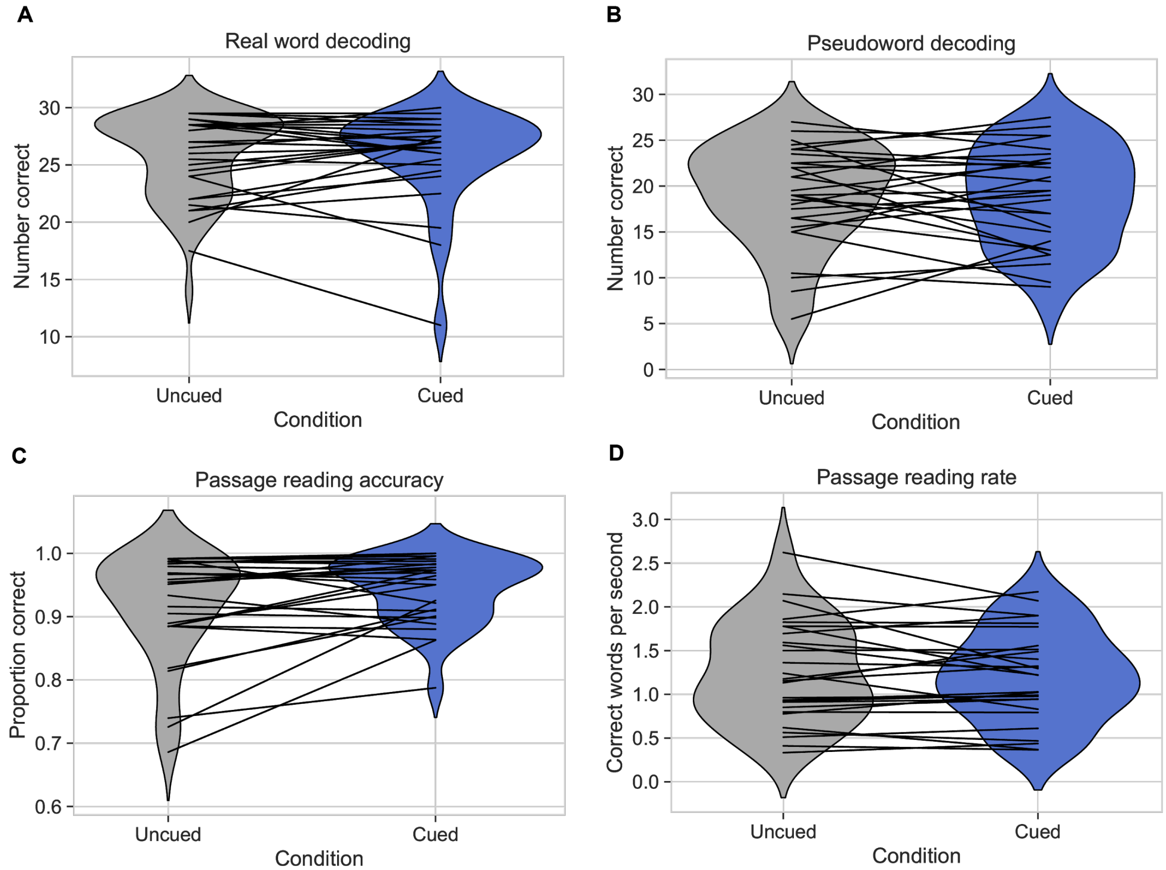
Data were analyzed using linear mixed effects (LME) modelling from the MATLAB Statistics and Machine Learning Toolbox (The MathWorks, 2017). Performance on each outcome measure (real word decoding, pseudo word decoding, passage reading accuracy, and passage reading rate) were fit with an LME model with a fixed effect for cue condition (cued/uncued). Each model also included two random effects: (1) individual subject and (2) stimulus order group. Based on model comparisons using AIC/BIC fit statistics, an additional random effect of word list was added to the real and pseudoword reading models. In line with our previous work, an adjusted alpha of 0.0125 was used for significance thresholds. All data and analysis code is publicly available at https://github.com/patdonnelly/SoundItOut.git.

3.3 Results

3.3.1 *A phonemic cue does not provide immediate benefit for struggling readers*

Using grade level passages and lab-created word and pseudoword lists with and without the *Sound it Out* image cues, we measured four facets of reading performance: (1) real word decoding, (2) pseudo word decoding, (3) passage reading accuracy, and (4) passage reading rate. **Fig 2** shows performance across each of these measures with violin plots of the distributions in each condition and line plots depicting within-subject performance across conditions.

Figure 3.2: Within-subject performance with and without the cue.



Violin density plots show distributions of performance on uncued (gray) and cued (blue) conditions with lines showing individual comparisons across conditions. Plots are shown for real word decoding, pseudoword decoding, passage reading accuracy, and passage reading rate.

Across all outcome measures, we found no significant effect of condition. (real word decoding: $\beta = 0.22$, $t(108) = 0.59$, $p = 0.557$; passage reading accuracy: $\beta = 0.03$, $t(52) = 2.48$, $p = 0.016$) (pseudoword decoding: $\beta = -0.11$, $t(110) = -0.17$, $p = 0.865$; passage reading rate: $\beta = -0.05$, $t(52) = -0.91$, $p = 0.367$). The effect for passage reading accuracy was near-significant but did not meet our four-comparison α correction of 0.0125. (See Table B.1 for full model output for each outcome measure). Thus, we conclude that a period of practice is necessary for participants to leverage the information in the phonemic cue to improve their reading.

3.4 Discussion

Study 1, which utilized a single-visit, crossover design, sought to determine if the app should be considered an adaptive affordance with spontaneous effect or as a supplemental intervention requiring a period of exposure or practice. Contrary to our hypothesis, there were no significant differences between performance in the cued and uncued conditions indicating that the cue is unable to provide immediate benefit to reading performance. Even though children were able to remember the vowel that the cue represented, they did not use the cue to derive a benefit in sounding out complex words.

The results of this study are in line with other findings of the important role of iterative practice in educational technologies for literacy (Grant et al., 2012; Messer & Nash, 2018; Ronimus & Lyytinen, 2015; Taylor, Handler, FitzPatrick, & Whittingham, 2020). Albeit intuitive, these findings demonstrate that a period of exposure is necessary for early readers to gain familiarity with the image cues in order to utilize them in an assessment context. This is supported, in part, by the suggestive benefit observed for passage reading accuracy. During passage reading, a child is additionally motivated by a desire to understand what they are reading, which may provide the necessary encouragement to incorporate even a novel strategy. This is juxtaposed with the non-significant results for measures of single word decoding, where the motivation may be just to get through the list as quickly as possible. As supported by related work, technologies such as Sound it Out are more impactful when implemented in a highly engaging, motivating context (Hutchison et al., 2012; Larson, 2010; Tare, Shell, & Jackson, 2020; Wolf, Gottwald, Galyean, Morris, & Breazeal, 2014).

Based on these findings, our focus shifted to (a) the duration of and (b) type of practice/exposure which would lead to benefits for measure of decoding and fluency. In a second, follow-up study, we implemented *Sound it Out* for an extended period (one month) with a dual intervention group; against an independent reading control group, an independent intervention group was trained and asked to practice on their own, and a dyadic reading group was trained and asked to practice only with a caregiver present.

Chapter 4. PROLONGED EXPOSURE & SUPERVISED PRACTICE

This work represents Study 2 in a manuscript currently submitted to the journal Reading Research Quarterly with co-authors Liesbeth Gijbels, Kevin Larson, Tanya Matskewich, Paul Linnerud, Patricia Kuhl, and Jason Yeatman

Research on educational technologies for reading instruction is disproportionate to the myriad applications in the marketplace. Here we assess a web-based reading tool, Sound it Out, that assists struggling readers in vowel decoding. Created as a collaboration between researchers and technology developers, the tool provides a phonemic-image cue below the vowels to scaffold letter-sound correspondence. This study utilized a repeated measures randomized controlled trial design to determine if an extended practice period (1 month) produced gains and whether a caregiver supervising practice enhanced benefits. Seventy-six struggling readers (ages 7-13) were randomly assigned to two intervention groups (independent and dyadic reading) and one control group. Results showed significant, dose-response benefits to decoding accuracy and passage reading accuracy for the combined intervention groups in comparison to controls. Moreover, supervised dyadic reading enhanced the effect of practice. These results highlight the potential for an evidenced-based supplemental learning technology to support both independent and shared reading for struggling readers.

4.1 Introduction

Even before the COVID-19 pandemic, the ubiquity of technology in education was evident. As schools went virtual, and education became inextricably linked to screens, there has been increased interest in educational technologies for reading instruction (Cassidy, Ortlieb, & Grote-Garcia, 2020; Greenfield, 2009; Wolf, 2018; Zuckerman, 2017). In their pivotal report, the

National Reading Panel suggested that computers represent a promising opportunity for teaching reading (National Reading Panel, 2000). Twenty years later, the marketplace for educational technologies for literacy is expansive (Guernsey & Levine, 2015), and 'multimodal/digital literacy' is a priority for educators and researchers nationwide (Cassidy et al., 2020). There have been many promising and successful technologies, but the vast majority of educational technologies for literacy either lack a connection to evidence (Christ, Arya, & Liu, 2018; Guernsey & Levine, 2015), have small effect sizes (Cheung & Slavin, 2011) or have not been studied empirically (Cheung & Slavin, 2013; Stetter & Hughes, 2015).

The problem, according to a recent meta-analysis, is the prevailing ideology that there is inherent 'magic' when machines enter the world of education (Cheung & Slavin, 2013). Although these technologies are implemented at-scale, analyses reveal limited short- and long-term gains, small sample sizes and less rigorous designs (Blok et al., 2002; Grant et al., 2012; Stetter & Hughes, 2015). However, juxtaposed with these findings are consistent reports of the opportunity that exists for digital solutions for struggling readers (Benton et al., 2014; Cheung & Slavin, 2013; Hutchison et al., 2012; Soe et al., 2000; Wolf et al., 2017). Namely, researchers point to various aspects of technology that make it ideally suited to the needs of emerging readers: technology provides seemingly limitless opportunities for practice and scaffolding (Campuzano, Dynarski, Agodini, & Rall, 2009; Ronimus & Lyytinen, 2015; Soe et al., 2000), persistent feedback (Benton et al., 2018; Richardson & Lyytinen, 2014), and motivated, child-driven autonomy in the learning process (Larson, 2010; McTigue, Solheim, Zimmer, & Uppstad, 2020; Wolf et al., 2014). Most poignantly, technology provides an opportunity to supplement classroom learning with individually-optimized material to support the development of foundational skills that a struggling reader requires (Atkinson & Hansen, 1966; Barker & Torgesen, 1995; Eagleton & Dobler, 2007;

Gee, 2003; Grant et al., 2012; Harris & Hofer, 2009; Hutchison et al., 2012; Ok & Rao, 2019; Soe et al., 2000; Solheim, Frijters, Lundetræ, & Uppstad, 2018; Taylor et al., 2020).

A successful application of technology targets the foundational skills of phonological awareness and letter-sound correspondence (Abrami, Lysenko, & Borokhovski, 2020; Barker & Torgesen, 1995; Joseph K. Torgesen et al., 2010). Phonics, or code-based, curricula are an essential component of evidence-based instruction (Castles et al., 2018; Moats, 1999; Petscher et al., 2020; Seidenberg, Cooper Borkenhagen, & Kearns, 2020), and are particularly important for challenged readers (Ehri et al., 2001; Maureen W. Lovett et al., 2017; Vaughn et al., 2008; Wanzek et al., 2013). Due to a multitude of factors ranging from disability and limited adoption of science-based reading practice in schools, struggling readers are unable to access high-quality, iterative and targeted instruction in traditional classrooms (Hindman, Morrison, Connor, & Connor, 2020; Petscher et al., 2020; Seidenberg, 2017). For this reason, many practitioners, advocates, and families have turned to technology as means to tailor education to a reader's unique needs.

Sound it Out, the subject of the current work, was designed to address this need in providing a foundational tool to empower emerging readers both at school and at home. Created with the Advanced Reading Technologies team at Microsoft, Sound it Out is a web-based application that is nested within the domain of educational technologies that focus on phonological awareness and sound-letter correspondence skill. For each vowel in a given text, the app utilizes a speech application programming interface (API) to access the corresponding image cue that relates to common pronunciation in Standard American English. The image cue is placed underneath the printed vowel, as an annotation, and contains an auditory cue when triggered with a click. So, when a reader comes to a challenging word, they can use both visual and auditory cues to determine the target vowel sound, or, with mastery, the visual cue alone. Sound it Out was designed in a collaborative effort between researchers at the University of Washington (UW) and Microsoft, and all testing was conducted independently by the UW laboratory to determine the legitimacy of the web-app for emerging readers.

The theoretical basis for the phonemic cue used in Sound it Out in reading instruction is well documented. For elementary educators, many evidence-based curricula use image cards and symbols to help scaffold letter/sound knowledge (Moats, 1999). For example, Fundations, a part of the Wilson Language Training program, uses pictures to represent both consonants and vowel sounds (i.e. Apple for ‘A’, Bat for ‘B’, etc) (“Fundations®,” n.d.). Moreover, Vowel Alert is a component strategy of PHAST (for Phonological and Strategy Training), a well-studied phonics program, and similarly cues readers to try out vowel sounds to find the one that ‘yields the correct word’ (M. W. Lovett, Lacerenza, & Borden, 2000). Thus, Sound it Out provides an on-demand synthetic phonics-based scaffold (Mandel Morrow & Asbury, 1999) to aid readers in learning the varied orthographic representations of English vowels.

The research on digital applications of phonics strategies is reassuring: compared to more comprehensive classroom oriented technologies, supplemental programs that focus on alphabetic skills, explicit phonological awareness instruction, and phoneme sequencing are among the most effective programs for improving outcomes (Cheung & Slavin, 2013; Joseph K. Torgesen et al., 2010). Of the supplemental programs that have shown promise, a few have stood out in their commitment to merge research and practice: Graphogame, Daisy Quest, Lexia and ROSS. Graphogame, a gamified, literacy app that has been implemented in many languages, works at the level of the sound to focus on synthetic phonics skills, rhyme, morphology, and orthography (Richardson & Lyytinen, 2014). In using a gamified platform, Graphogame is able to produce high levels of motivation and engagement, both at school and at home (McTigue et al., 2020; Ronimus

& Lyytinen, 2015). Daisy Quest, another gamified project, provides a whimsical narrative to teach rhyme, onset-rime, and sound segmentation and has been associated with measurable gains in phonological awareness mastery (Barker & Torgesen, 1995; Erickson, Foster, Foster, Torgesen, & Packer, 1992). Lexia, a technology designed for school implementation, focuses on phonetic strategies as well, but utilizes telemetry and tutorial-style learning to provide personalized and adaptive support (Macaruso, Hook, & McCabe, 2006). Reading with Orthographic and Segmented Speech (ROSS), in contrast, provides various levels of segmentation - from phonemes to syllables to words - to help readers when they encounter challenging words in connected-text. In combination with an explicit phoneme manipulation program, ROSS not only provided measurable gains but also saw high levels of engagement and motivation in a population that is prone to feeling overwhelmed by passage-based learning (Wise & Olson, 1995, 1998). Together, these and other programs provide unique methods for technology to enhance learning, but all adhere to the principle of implementation that a program's impact is dependent on its ability to help both students and teachers achieve their instructional goals (Lee, Waxman, Wu, Michko, & Lin, 2013).

In the same vein as Sound it Out, are educational technologies such as Trainertext and SeeWord Reading. Trainertext, developed by researchers and developers in the United Kingdom, provides visual mnemonics above each phoneme (vowels and consonants) to whimsically support shared reading with an adult. Results from a ten-month, longitudinal study purported significant benefits to measures of decoding and fluency (Messer & Nash, 2018). SeeWord Reading, instead integrates a picture-embedded font to provide a symbolic scaffold for learning the alphabetic principle. Studies demonstrate that the application produces measurable gain in grapheme-phoneme consolidation (Seward et al., 2014). As digital siblings of Sound it Out, both Trainertext and SeeWord Reading reveal the validity of a tool that provides a symbolic image cue to assist emerging readers. Unique to Sound it Out, is the versatility and flexibility it has to adapt to the variability in curricula for individual learners.

Additionally, providing phonemic annotations has been shown to be useful for emerging readers across different orthographies. In both Arabic and Hebrew, diacritic marks (pointed-scripts) are used to vowelize text for beginning readers (Mahfoudhi, Everatt, & Elbeheri, 2011). This vowelization transforms the deep orthography of formal Arabic and Hebrew into a shallow orthography and allows readers to use a phonological code to gain familiarity with the more complex visual script (Abu-Rabia, 1999, 2001; Schiff & Saiegh-Haddad, 2018). As a result of using this phonological cue, phonological processing skills are highly predictive of reading performance for Arabic readers even as they transition to the non-vowelized formal orthography (Elbeheri & Everatt, 2007). Akin to English, dyslexic Arabic children demonstrate a similar phonological deficit and pronounced difficulty with the vowelized script (Elbeheri & Everatt, 2007). In terms of efficacy, studies have shown that early readers who have access to these pointed-scripts in Hebrew demonstrate greater accuracy in word reading (Haddad, Weiss, Katzir, & Bitan, 2018; Shany, Bar-On, & Katzir, 2012), but that benefit decreases with age and reading proficiency (Schiff & Saiegh-Haddad, 2018; Weiss, Katzir, & Bitan, 2015a, 2015b). Benefits are conflicting for children with dyslexia (Weiss et al., 2015b), but, to knowledge, few studies have explored auditory cues alongside pointed-scripts as an intervention tool. Related to the goals of Sound it Out, the use of diacritics in Proto Semitic languages provides a phonological bootstrap to both orthography and morphology (Elbeheri & Everatt, 2007; Haddad et al., 2018).

A digital cue to scaffold vowel decoding during passage reading is also rooted in the concept of 'word learning in context' (Ehri, 2014), a major tenet of blended theories of reading development. In line with the Interactive Strategies Approach, Sound it Out is an application of explicit phonics instruction in combination with word solving skills to incorporate 'interactive and confirmatory' strategies to approach novel words and build sight vocabulary (Vellutino & Scanlon, 2002). As this theory suggests, Sound it Out uses a meaningful, engaging context to provide a code-based cue that will ensure long-term consolidation of sub-lexical units (i.e. vowels, syllables, morphemes) (Mesmer & Williams, 2015). This is duly supported by the Phase Theory of Reading (Ehri, 1995), which stipulates that "reading words in meaningful contexts ensures that syntactic and semantic identities of words become bonded to spellings and pronunciations to form amalgamated units in memory" (Ehri, 2020). Fusing phonics and comprehension extends to the dynamic, reciprocal relations of the Lattice Model of reading development (Hindman et al., 2020; McDonald Connor et al., 2009), and the coordination of implicit and explicit statistical learning mechanisms in computational models (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Seidenberg, 2005; Seidenberg et al., 2020; Share, 1995). Sound it Out, then, provides readers a context to compare sources of information and a tool to focus attention on the information that will encourage long-term consolidation and learning: the alphabetic code (Ehri, 2014; Scanlon & Anderson, 2020).

Initial, proof-of-concept research into Sound it Out revealed limited, yet promising effects of the cue on reading performance (Donnelly et al., 2020). In a small-scale randomized controlled trial (RCT), with a two-week practice period, our research group found that the cue had a significant impact on pseudo word decoding but no significant impact on connected-text reading or real word decoding. The pseudoword decoding benefit was small - about two additional items read, on average - however, combined with qualitative survey data which demonstrated substantial interest and engagement, overall findings suggested that, with some modifications, this technology could represent a promising approach to supplement phonics based classroom instruction.

The present work builds on these initial findings to (a) hone the implementation of the technology and (b) test its efficacy under various use cases. Namely, this paper details two studies that aim to describe the conditions that are necessary for the cue to impact reading performance beyond single word decoding. Extending from our previous research that was ambiguous with regards to the effect of practice on cue benefit, Study 1 utilizes a single-visit crossover design to determine if we can replicate our previous findings without the practice period. In other words, does the phonemic image provide an immediate benefit to reading? Because the Sound it Out cue was so easy to learn, we hypothesized that children could immediately use the cue to sound out difficult words that were above their reading level. Study 2, an RCT conducted one year later, employs a version of the app now powered by a speech API to answer two main questions: (1) Can an extended period of exposure produce proportional gains in reading skill? And (2) Can the benefit of the cue be enhanced by involving a parent/caregiver in the intervention (dyadic reading)? We hypothesized that an extended period of exposure would produce measurable gains and that the gain would be significantly pronounced in the dyadic reading condition.

4.2 METHODS

4.2.1 *Participants*

Eighty-one participants between 8 and 13 years old were recruited from the same database based on the same characterization criteria used in Study 1 for struggling readers. Due to participant attrition, seventy-six participants completed at least two sessions to be included in the sample. Study groups were evenly represented in terms of participant sex (control = 11 female; independent = 11 female; dyadic = 11 female) as well as the other demographic variable collected (see Table 4.1). Forty-three had a diagnosis of dyslexia (control = 16; independent = 15; dyadic = 13) and twenty-two had a diagnosis of ADHD (control = 9; independent = 6; dyadic = 7). As with Study 1, all procedures, including remote assessment due to the COVID-19 pandemic, were approved by the UW HSD.

Table 4.1: Summary statistics and group differences for study sample

	Control (n = 25)		**Independent (n = 26)**		**Dyadic (n = 25)**		**Kruskal-Wallis Test**	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	X^2	*p*
Age (Y)	10.50	1.21	10.68	1.56	10.50	1.29	0.18	0.912
WJ BRS	84.08	9.49	80.62	10.08	81.92	10.47	1.92	0.383
TOWRE Index	73.44	7.77	73.77	6.48	73.80	8.35	0.15	0.927
WASI FS2	100.68	10.55	106.46	13.29	101.80	11.34	2.89	0.236
CTOPP PA	91.12	9.33	92.31	12.88	87.72	11.58	2.27	0.321
CTOPP RSN	78.92	20.54	79.92	12.54	83.17	9.46	2.12	0.346

Means and standard deviations are given for each demographic variable of interest for participants in Study 2 for each of the study groups: control, independent readers, and dyadic readers. Chi-square statistics and p values from Kruskal-Wallis tests reveal that there is no significant difference across groups.

4.2.2 *App Design*

As described in 3.2.2 (above), *Sound it Out* is a web-app that provides a phonics-grounded, annotation scaffold for decoding vowels in connected-test reading. Study 2 utilizes the same app, but with some minor improvements and modifications. Rather than the manual coding that was performed in previous iterations, the vowel coding is standardized using a speech API developed by Microsoft: the Speech Application Programming Interface (SAPI). Additionally, the voice cues were updated to provide professionally recorded narration with extended vowel sounds.

In Study 1 and previously, there was concern (both participant-reported and researcher-observed) that the image cues were not optimal for the target population: struggling readers. The image cues were chosen as highly recognizable monosyllabic images that would require little instruction, were already used in vowel teaching in classrooms, and focused on word-middle sounds: "bed" for the vowel /ɛ/ and "pig" for the sound /ɪ/. As phonological elision is difficult for dyslexic and struggling readers, especially for word-middle sounds (Snowling, 2004), we created and tested an alternative set that optimized for word-initial sounds (an easier type of elision): "elephant" for the vowel /ɛ/ and "igloo" for the sound /ɪ/. (see Table B.1 for side-by-side image set comparison).

The two image cue sets, termed "Pigs & Suns" and "Igloos & Umbrellas", were tested in a small-scale validation study with twenty-two typically reading (WJ BRS; Mean = 105.95, Sd = 14.24) children ages 7 to 15 years old (female = 10, male = 12). Participants were randomized to either image set and given a brief training to learn the images and their associated vowel sounds. Training utilized flashcards and was complete when each participant demonstrated mastery of each image and its associated target sound. Then the participants were tested using 35 flashcards displaying single words with the vowels replaced by images in their trained image set. For example: H *dog* N ("Pigs & Suns" set) or H *octopus* N ("Igloos & Umbrellas" set), with the correct response of /han/. Accuracy was measured as the number of correctly pronounced items in the set.

We compared performance using an independent samples t-test and found that there was no significant difference in performance between the two sets ($t(20) = -0.25$, $p = 0.806$) (see Figure B.1). Although these results do not discount issues of individual preference, these findings demonstrate that there is no compelling reason to change the cue set for the current study. The Pigs & Suns set was implemented once again in the present study.

4.2.3 *Study Design*

In a randomized, repeated-measures design 76 participants were randomly assigned to one of three participant groups: a single control group and two intervention groups - independent and dyadic. Participants were randomly assigned based on time of consent and sibling participants were assigned to the same group for protocol adherence. All participants were asked to complete three in-person or virtual testing sessions and two, two-week practice periods with the aid of a touch-screen device larger than a smartphone. To reduce issues of access, participants were given the option of a lab-provided Kindle Fire HD 8 for use during the study.

During visit one, all participants completed baseline testing of all measures using the uncued condition (test without Sound it Out) (see 4.2.4). The two intervention groups then underwent a short training program that introduced them to the image cues, the target vowel sounds and voice cues (see (Donnelly et al., 2020) for a more detailed description of the paradigm). For the dyadic group, a parent/caregiver was asked to participate actively in the training and demonstrate similar proficiency in using the web-app. After training, both intervention groups were instructed to complete at least five passages per week with Sound it Out enabled for the next two weeks prior to visit two. Those in the dyadic reading group were instructed to read with a parent/caregiver while independent readers were asked to read alone. Practice was tracked using brief comprehension quizzes (see 4.2.4). In visit two, intervention participants completed testing in all outcome measures, but using the cued condition (with Sound it Out). Participants then completed an abbreviated ‘refresher’ training program (with a caregiver for the dyadic group) and asked to adhere to the same practice regimen for the final two-week period. In their final visit, participants once again completed all outcome measures in the cued condition before completing a researcher-administered questionnaire.

Control participants completed identical, two-week practice periods and testing sessions, but using only the uncued condition. Training, moreover, was limited to a tour of the app and instructions for completion of the comprehension quizzes. Control participants were asked to read independently.

4.2.4 *Measures*

One primary and four secondary outcome measures were collected at all three time points: decoding accuracy, real word decoding, pseudo word decoding, passage reading accuracy and passage reading rate. Real and pseudoword decoding were measured using lab-created word lists generated using the orthographic word form database MCWord (Medler & Binder, 2005). Each of the six lists consisted of 30 unique real or pseudo words, with two lists administered at each visit. Real words lists contained five items each of three- to eight-letter words, selected for high English-language occurrence. Pseudoword lists were similarly structured but were generated based on English bigram frequency identically scaled from three- to eight-letter words. All lists consisted of unique words/pseudowords and were administered in a counterbalanced order across participants. (See Table B.3 and Table B.4 for detailed word statistics). As with the previous study (Donnelly et al., 2020), the lists were calibrated for difficulty using timed reading in ten typically reading adults and were phonetically coded manually by the first author based on common pronunciation in the regional dialect. As there are multiple pronunciations for many of the words, accuracy was scored using any acceptable pronunciation of the word items. For pseudowords, accuracy was scored using most common pronunciations of any given bigram, as defined prior to the start of the study. List reliability was calculated using the correlation between the first and second list administered at each visit. Results show high reliability in performance for both real word ($r = 0.89$, $p < 0.001$) and pseudoword ($r = 0.89$, $p < 0.001$) lists. Decoding accuracy (the primary outcome measure) was calculated as a composite of real- and pseudo-word reading accuracy. Real and pseudoword decoding were pre-defined as secondary outcome measures.

Passage reading accuracy and rate were measured using levelled test passages from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) 8th Edition library. These passages consisted of six, fourth-grade level passages from the Oral Reading Fluency subtest and were used with permission from the University of Oregon Center on Teaching & Learning. Designed to be administered for progress monitoring in the classroom, these passages are commonly used in both research and practice. For testing, passages were viewed within the Immersive Reader tool to both standardize font properties across participant groups and reduce distractions typical of a digital reading environment. For the cued condition, test passages were phonetically coded by SAPI. At each visit, two passages were read twice in a repeated reading design and were audio recorded for accurate scoring and coding. Analyses focused on the accuracy of the first read (number of words pronounced correctly) and the rate of the second read (accurate words per second). With these two measures, we aimed to uncover the effect of the cue independent of the expected consequences of implementing a novel strategy for fluency. First read accuracy and second read rate were pre-defined as secondary outcomes.

For practice, participants were provided with a bank of seventy-five, second- to fifth-grade passages for daily reading. The passages were accessed (with permission) from both the DIBELS 6th Edition library and ReadWorks.org. Passages were chosen for diverse interests/topics and, for the intervention groups, coded using the Microsoft speech API. For each passage, a short, 3-question comprehension quiz followed to log practice adherence. Comprehension questions were created by a Master's level teacher certified to work with struggling readers and were grade-level matched to the specific passage. Practice was measured as the number of quizzes completed.

4.2.5 *Statistics*

Using the results and effect sizes from previous work (including Study 1), we pre-defined our analysis to test (A) the difference between the combined intervention groups and the control (intervention effect) and (B) the difference between the two independent intervention groups (group effect) using one primary (decoding accuracy) and four secondary outcome measure (real word decoding, pseudoword decoding, passage reading accuracy, and second read rate). An a priori power analysis was conducted using G*Power3 (Faul, Erdfelder, Lang, & Buchner, 2007) to test (A) the difference between the combined intervention groups and the control group and (B) the difference between the two independent intervention groups. A medium-to-large effect size of .80 standard deviation (SD) units was hypothesized for both comparisons A and B. In this case, a total sample size of 75 participants across the three groups (control = 25, independent = 25, dyadic = 25) would provide approximately 90% power for comparison A and 80% power for comparison B using a two-sided alpha = .05. To accommodate possible dropout over the course of the study, 81 participants were recruited to ensure that a sufficient number would complete all study visits.

Thus, analyses examined the impact of *Sound it Out* on the primary and secondary outcome measures at two levels: the intervention effect and the group effect. For both levels, we implemented LME modelling to explore both within- and between-subject effects. Models were fit with fixed-effects of (1) time (session as a categorical variable), (2) group (intervention/control or independent/dyadic as a categorical variable), and (3) group by time interaction. All models contained random effects for time and participant to account for individual variation in baseline and slope. For decoding measures, an additional random effect of specific list was added based on model fit comparison analysis using AIC/BIC statistics.

Post-hoc analyses were performed to explore the impact of four predictor variables: age, WJ BRS, TOWRE Index, CTOPP phonological awareness and practice adherence (number of quizzes completed by each visit, demeaned for each participant). These analyses utilized an LME model with fixed effects of (1) predictor and (2) time, with identical random effects structures to the primary analyses. All data and analysis code is publicly available at https://github.com/patdonnelly/SoundItOut.git.

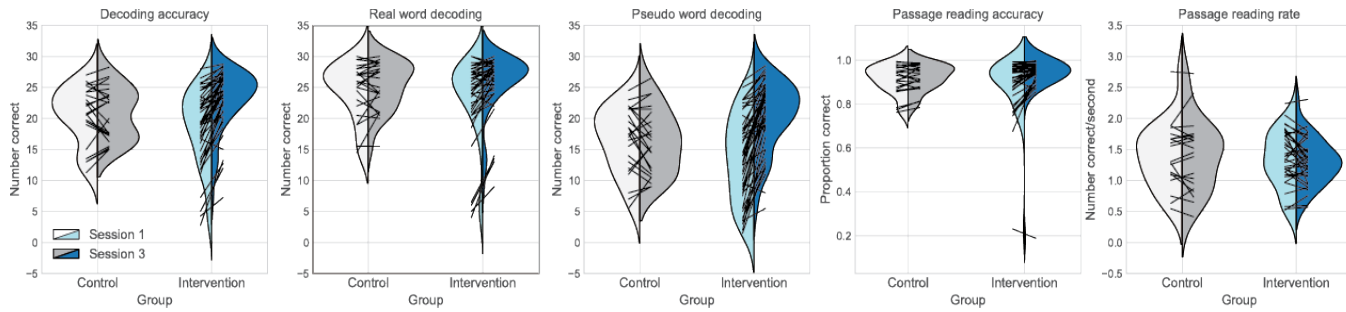
4.3 RESULTS

4.3.1 *With prolonged exposure, benefits increase and extend to connected-text reading*

We first compared the combined intervention groups to the control group (Comparison A) to determine if we could, first, replicate previous findings of an effect on isolated, pseudoword decoding and, second, whether this reading benefit would extend to higher-order skills with a doubling to the practice period.

The primary outcome measure, Decoding accuracy, was calculated as a combination of performance on real word and pseudoword decoding measures, which were individually analyzed as secondary outcome measures. Words were presented in a vertical list and administered free of time constraints. Moreover, secondary outcome measures of passage reading accuracy and rate were assessed using digitally presented passages. Figure 4.1 depicts group level distributions and individual trends at the start and end of the intervention period.

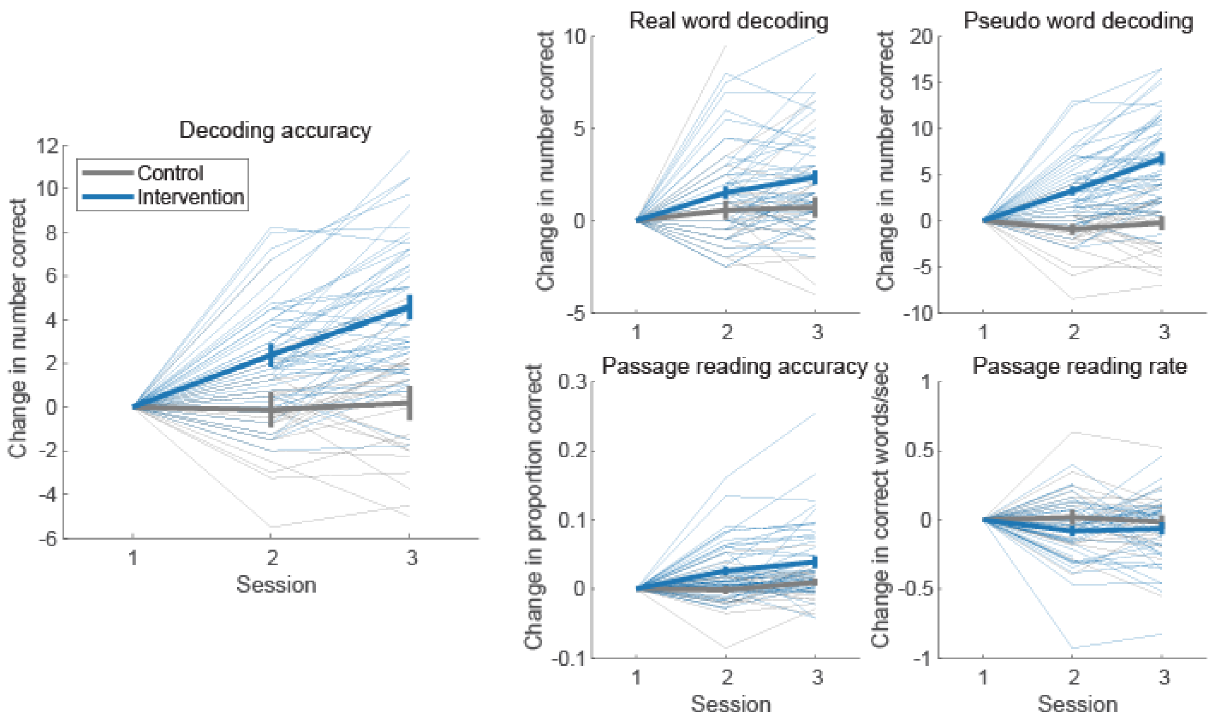
Figure 4.1: Within-subject pre-post intervention performance.



Violin density plots show distributions of performance for the control (gray) and the combined intervention (blue) conditions. Individual line plots depict participant-level performance at sessions 1 and 3. Plots are shown for the decoding accuracy composite, real word decoding, pseudoword decoding, passage reading accuracy, and passage reading rate.

For comparison A of the decoding accuracy primary outcome measure, the group by time effect was highly significant at both midpoint ($\beta = 2.58$, $t(895) = 4.26$, $p = 2.28*10^{-5}$) and post-test ($\beta = 4.28$, $t(895) = 5.73$, $p = 1.34*10^{-8}$). The amount of growth was roughly equivalent between each timepoint. These results indicate that growth in the combined intervention groups was significantly different than the control group, and that the observed growth follows a roughly linear, dose-response trajectory (see Figure 4.2).

Figure 4.2: Rate of change plots for intervention effect comparison.



For each outcome measure line plots depict change at the group-level from the LME model for the combined intervention group (blue) and the control group (gray). Although individual intercepts are removed for visualization, model fits are calculated with the full fixed and random effects structure defined in the Methods. Individual slopes are plotted to demonstrate variability between-subjects. Error bars represent +/- 1 SEM.

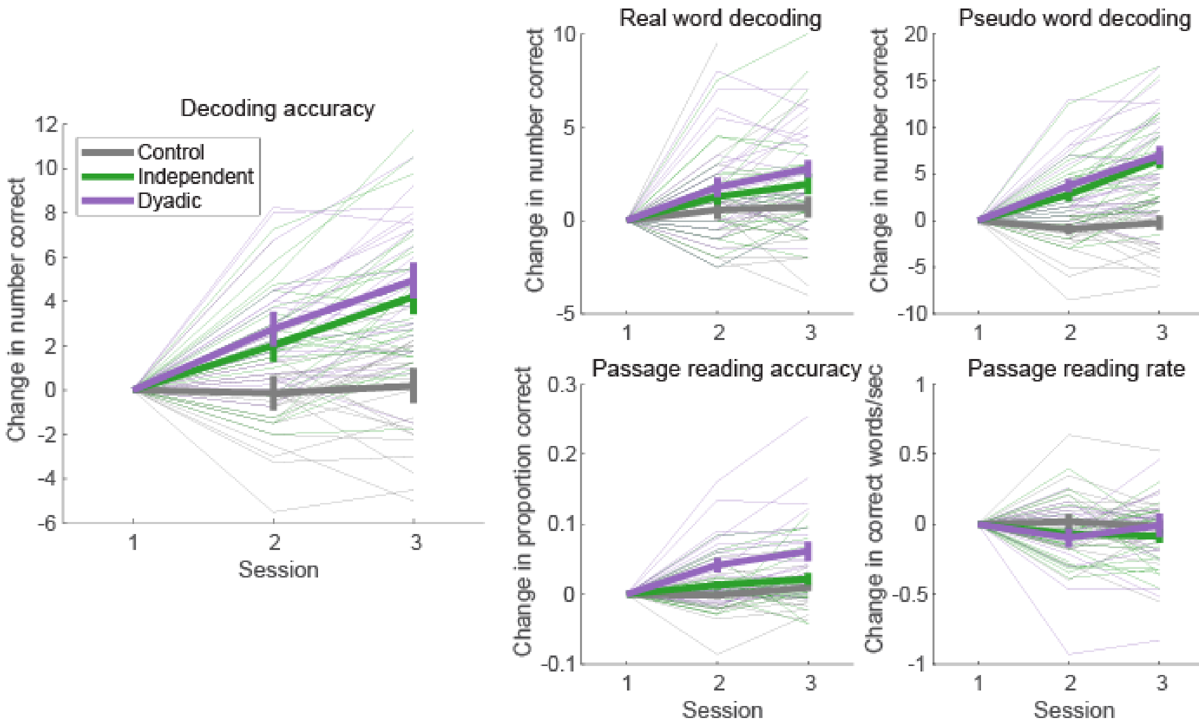
As decoding accuracy was a composite of real and pseudoword decoding accuracy, we examined the component measures as secondary outcome measures. For pseudo word decoding, the group by time effect was significant at both time points (midpoint: β = 4.21, t(444) = 5.15, p = $3.85*10^{-7}$, post-test: β = 6.96, t(444) = 6.39, p = $4.21*10^{-10}$). These findings replicate previous findings and demonstrate added benefit of the cue with added additional practice/exposure. For real word decoding, the group by time effect was not significant at midpoint (β = 0.95, t(445) = 1.62, p = 0.105) but significant at post-test (β = 1.64, t(445) = 2.57, p = 0.011). These results replicate previous findings (i.e. the lack of effect at session 2 (Donnelly et al., 2020) and extend these previous results to provide evidence for benefits to real word reading skills with prolonged exposure.

To examine the effect of the cue on connected-text reading for comparison A, we examined two final secondary outcome measures: passage reading accuracy and rate. For passage reading rate, the group by time effect was not significant at either time point (midpoint: $\beta = -0.10$, $t(176) = -1.57$, $p = 0.117$; post-test: $\beta = -0.07$, $t(176) = -1.01$, $p = 0.315$). For passage reading accuracy, the group by time effect was significant at both midpoint ($\beta = 0.03$, $t(208) = 2.86$, $p = 0.005$) and post-test ($\beta = 0.03$, $t(208) = 2.50$, $p = 0.013$). (See Table B.4 for full model output for each outcome measure). These results demonstrate that with prolonged practice with the *Sound it Out* cue, benefits are observed beyond single word decoding, generalizing to measures of reading fluency.

4.3.2 *A significant advantage of supervised reading for passage reading accuracy*

Our second comparison (B) focused on the two intervention groups to determine if there was an effect of type of practice (independent versus dyadic reading) on *Sound it Out* cue benefit. In this analysis, we sought to disambiguate the role of the cue as a scaffold for independent reading or an aid in empowering a shared reading experience with a parent or caregiver. Figure 4.3 shows performance across the primary and secondary measures for those in the two intervention groups compared to control.

Figure 4.3: Rate of change plots for group effect comparison.



For each outcome measure line plots depict rate of change at the group-level from the LME model for the dyadic reading group (purple), the independent reading group (green), and control (gray). Although individual intercepts are removed for visualization, model fits are calculated with the full fixed and random effects structure defined in the Methods. Individual slopes are plotted to demonstrate variability between-subjects. Error bars represent +/- 1 SEM.

For comparison B of the decoding accuracy primary outcome measure, the group (Dyadic vs. Independent) by time effect was not significant at midpoint ($\beta = 0.74$, $t(603) = 1.05$, $p = 0.296$) or post-test ($\beta = 0.66$, $t(603) = 0.75$, $p = 0.455$). This suggests that both groups benefited equally and there is no significant difference between the type of practice for measures of single word decoding. This is reflected in both component measures for pseudo word decoding (midpoint: $\beta = 0.97$, $t(298) = 0.98$, $p = 0.327$; post-test: $\beta = 0.49$, $t(298) = 0.37$, $p = 0.713$) and real word decoding (midpoint: $\beta = 0.51$, $t(299) = 0.76$, $p = 0.445$; post-test: $\beta = 0.83$, $t(299) = 1.16$, $p = 0.248$).

As with the previous comparison, we examined the effect on two measures of connected-text reading: passage reading accuracy and rate. For passage reading rate, the group by time effect was not significant at either time points (midpoint: $\beta = -0.03$, $t(116) = -0.43$, $p = 0.665$; post-test: $\beta = 0.05$, $t(116) = 0.55$, $p = 0.58$); however, for passage reading accuracy, the group by time effect was significant at both midpoint ($\beta = 0.03$, $t(134) = 2.41$, $p = 0.017$) and post-test ($\beta = 0.04$, $t(134) = 2.37$, $p = 0.019$). (See Table B.5 for full model output for each outcome measure). These results demonstrate that there is a potential, enhanced benefit to supervised, shared reading for connected-text reading accuracy.

4.3.3 *Cue may be particularly impactful for younger & more struggling readers*

In an exploratory analysis, we examined the role of five moderators (WJ BRS, TOWRE Index, CTOPP Phonological Awareness, CTOPP Rapid Naming, and chronological age) in predicting the benefit of the cue on our primary and secondary outcome measures. The goal of these analyses was to investigate potential subgroup differences that influence the impact of *Sound it Out*. For example, because the cue relies on phonological awareness skills (i.e. elision) to isolate the target vowel sounds from the symbol names, the cue may be more well-suited for those with higher scores on tests of phonological elision (CTOPP PA).

For these analyses, an LME model was fit to data for the combined intervention groups with fixed effects of (1) time (session as a categorical variable), (2) moderator, and (3) predictor by time interaction. Identical random effects structures to the main analyses were used. Using this method, we examined whether the cue response was affected by performance along the predictor continuums.

For decoding accuracy, suggestive predictor by time interactions were found for two predictor variables: TOWRE Index and chronological age. Results indicate that the cue may benefit those with lower TOWRE Index scores (midpoint: $\beta = -0.078$, $t(603) = -1.58$, $p = 0.114$; post-test: $\beta = -0.15$, $t(603) = -2.64$, $p = 0.008$) as well as younger users (midpoint: $\beta = -0.02$, $t(603)$

= -0.79, p = 0.406; endpoint = β = -0.06, t(603) = -2.41, p = 0.016). No effects were observed for CTOPP PA, CTOPP Rapid Naming or WJ BRS.

For the two component measures of decoding accuracy, results indicate that the observed associations are specific to real word decoding performance. For real word decoding, suggestive interaction effects were found for TOWRE Index (midpoint: β = -0.13, t(299) = -3.13, p = 0.002; post-test: β = -0.17, t(299) = -3.88, p = 0.0001), WJ BRS (midpoint: β = -0.07, t(299) = -2.07, p = 0.039; post-test: β = -0.04, t(299) = -1.25, p = 0.211, and chronological age (midpoint: β = -0.05, t(299) = -2.72, p = 0.007; post-test: β = -0.07, t(299) = -4.00, p = $7.89*10^{-5}$). These results indicate that the *Sound it Out* cue may be more beneficial for both more impaired and younger readers. No interaction effects were found for pseudo word decoding.

For passage reading accuracy, there was a small interaction effect for chronological age (midpoint: β = $9.8*10^{-4}$, t(134) = -2.82, p = 0.006; post-test: β = $1.2*10^{-3}$, t(134) = -2.56, p = 0.012) but no suggestive associations across reading-related predictors, compounding the results with decoding measures to suggest that the cue may have greater impact for younger readers. No significant associations were found for passage reading rate performance.

4.4 Discussion

Building on the knowledge gained from Chapter 3, we endeavored to find out if extended exposure and prescriptive practice could produce proportional gains that generalize to aspects of fluent reading skill. Using an RCT methodology, we asked: (1) Will prolonged exposure translate to added and generalized benefits to reading skill? and (2) Can motivated exposure with a parent/caregiver impact the benefit of the phonemic image cue? Participants were randomly assigned to two intervention groups - independent and dyadic readers - and a matched control to examine two comparisons of interest: (A) the intervention comparison - combined intervention groups versus control and (B) the group comparison - independent readers versus dyadic readers.

Results from comparison A demonstrate significant group by time interaction effects for the primary outcome measure, decoding accuracy, as well as the secondary measures of real word decoding, pseudo word decoding, and passage reading accuracy. These results replicate previous findings and demonstrate that observed benefits compound with prolonged usage and extend to areas of reading fluency. From comparison B, results show that although there was no significant group by time effect for measures of decoding, there was a significant advantage for the dyadic reading group for passage reading accuracy. These results confirm our hypotheses, but open additional questions about reader-to-reader variability. To examine these individual differences, we conducted an exploratory analysis on various predictor variables. Results indicate negative predictor by time interactions for real word reading and passage reading accuracy and suggest that, with future research, the cue may be ideally suited to younger and more struggling readers.

Due in part to troubling national literacy rates and the prevalence of reading-related disabilities, there has been high demand for educational technologies to support emerging readers outside the classroom (Guernsey & Levine, 2015). According to recent reports, literacy rates for American youth are declining. Since the previous iteration in 2017, the 2019 Reading Report Card reports that the percentage of U.S. fourth graders performing at or above proficiency in reading has decreased from 37% to 35% . This pattern is mirrored for eighth graders which observed a decrease from 36% to 34%, and is significantly pronounced for students in ethnic and racial minority populations (National Assessment of Educational Progress, 2019). Moreover, it is estimated that roughly 15% of the population has developmental dyslexia (Lyon et al., 2003; Wolf, 2007), the most prevalent and studied of the learning disabilities (Fletcher et al., 2018). These statistics paint a troubling landscape for youth in the U.S. as literacy rates in these foundational years are highly predictive of long term educational outcomes (Maughan, 1995; Rabiner, Murray, Skinner, & Malone, 2010; The Annie E. Casey Foundation, 2013) and overall health and longevity (Gilbert et al., 2018).

To address these poor literacy achievement rates, and regardless of etiology, many educators and caregivers have turned to digital solutions. The utility and accessibility of technology make it markedly appealing to reading education, but existing research is insufficient to inform those interested in how to utilize digital tools effectively. The goal of the present research was to address this need and contribute to the burgeoning literature on educational technologies for struggling readers with one such digital tool: Sound it Out.

With this study, we applied controlled study environments to elucidate not only the legitimacy of the cue for improving reading skill, but also contextualize the manner of use that allows for that improvement. These results support the implementation of Sound it Out as a meaningful scaffold for both independent and shared reading environments. With a brief training and motivated exposure, children who read on their own as well as those who read with an adult demonstrated measurable gains in decoding accuracy. Additionally, those who read with an adult made significant gains in fluency, particularly for younger readers. These results support consistent research in literacy education of the added value that an adult tutor or caregiver brings to fore, especially for emerging readers (Auerbach, 1989; Cheung & Slavin, 2013; Hannon & James, 1990; Kraft & Monti-Nussbaum, 2017; Lynch et al., 2006; Ronimus & Lyytinen, 2015; Tichnor-Wagner et al., 2016). Recent studies have shown that even a non-caregiver adult allows for a ‘tutor/child/computer triangle’ providing both pedagogical and socioemotional motivation (McTigue et al., 2020). Because of the role that caregiver encouragement played in our study, results also suggest that Sound it Out may also be ideally suited for peer-supported learning. Growing evidence has clarified the proximal advantage of peer learning strategies, both inside and outside the classroom (Drummond, Chinen, & Duncan, 2011; Fuchs & Fuchs, 2005; Tare et al., 2020). Future research should be conducted to explore the potential role of peer learning with Sound it Out.

Even though these findings are encouraging, there are important limitations to consider. First, we did not test the intervention groups on their performance without the cue, and control participants were never exposed to cue. As a result, we are unable to discuss the ability of the cue to benefit uncued reading. Based on results from an RCT of a related technology (Messer & Nash, 2018), we have reason to expect that, over time, the cue would help scaffold phonics knowledge and benefit reading skills more generally. But that remains to be tested. Second, we are unable to determine whether a period of caregiver-included practice without the cue would relate to comparable improvements in decoding and passage reading fluency. Third, although the groups were given clear instructions on how to practice (independently vs. dyadically), there was no ability to enforce adherence. We did, however, collect comprehensive parent-reported usage statistics which demonstrate overall coherence (see Figure B.2). Future research should study effects over the course of an academic year to explore generalizability and leverage detailed app usage data to add granularity to the impacts of the type of exposure. Finally, due to the variability in our sample, it will be necessary to conduct future studies with larger sample sizes and smaller age ranges to reduce the possibility of ceiling effects and regression to the mean in the current studies.

Considering these positive effects, the true impact of Sound it Out will depend on its ability to supplement what the reader learns in the classroom. Consistent across the literature on educational technologies is the importance of programs to integrate with and supplement existing curricula (Christ et al., 2018; Lee et al., 2013; Ok & Rao, 2019; Tare et al., 2020; Taylor et al., 2020). Namely, existing research has shown that although the degree of benefit is linked to amount of exposure with a caregiver, the degree of utilization is more significantly linked to the amount of exposure at school (Ronimus & Lyytinen, 2015). As our findings suggest that the amount of utilization is critical (the more a child was exposed to the tool, the more they improved across outcome measures), future research that involves schools is needed to reveal the true potential of Sound it Out.

During development of Sound it Out, persistent challenges were the consistency of the speech API-generated phoneme with participants' spoken dialects and the choice of image cues. Together with Microsoft, our goal was to implement a code that aligned with the International Phonetic Alphabet (IPA), but also matched most common pronunciations in Pacific Northwest regional dialect. The most consistent feedback from participants was a perceived mismatch between what they expected to see and the symbol that was supplied. This sets a challenge for future research - to modify cues to match individual dialects (Hattan & Lupo, 2020; Hoffman, Hikida, & Sailors, 2020; Noguerón-Liu, 2020; Willis, 2019). As for the image cues, our primary goals were to choose symbols that were easily recognizable and in alignment with commonly used images in curricula. As our studies have shown, the chosen symbols were quantifiably effective, but qualitatively, image favor was more mixed. In post study surveys several participants noted a desire to use alternate image cues either to align with specific instructional materials or as a matter of personal preference.

Fortunately, one of the many affordances of digital tools is their ability to adapt to individual variation. Just as word processors provide a selection of fonts for their users, technological tools like Sound it Out are inherently modular to adapt to user preference. In the future, not only will this allow schools and families to select the most meaningful images for their beginning readers, but it will also allow for the tool to adapt to dialectal and regional variation. For speakers of a minority dialect (such as African American English (AAE)), the erasure of minority dialects in mainstream education has been shown to be detrimental to reading development (Brown et al., 2015). With the ability of a tool like Sound it Out to adapt a child's home and school environment, educational technologies can help establish 'cultural synchronization' by expanding access to culturally-relevant text (Phillips Galloway, McClain, & Uccelli, 2020; Scott & Marcus, 2014; Washington, 2001). This is a promising direction for Sound it Out to provide a culturally-relevant affordance that can broaden access to multicultural knowledge (Hattan & Lupo, 2020).

To conclude, this pair of studies demonstrates that Sound it Out, a phonemic image cue annotation for vowels, can effectively support emerging readers more accurately decode complex words and passages. This tool represents a simple adaptation that is fundamentally modular, and grounded in principles of phonics, to provide a meaningful and adaptable cue to scaffold sound-letter correspondence. In the world of educational technologies, Sound it Out unites pedagogical practice and developer expertise to create an impactful, accessible aid for emerging readers. When in-person learning is not possible, and extending into the future when it is, this combination empowers families to take an active, participatory role in their child's learning.

Chapter 5. CONCLUDING REMARKS

In the course of three training studies, my coauthors and I have applied empirical methodologies to investigate the utility of *Sound it Out* as an aid for emerging readers. In the context of the growing literature on educational technologies, our results shed light on several useful conclusions for the field.

Primarily, a digital, symbolic annotation for vowel sounds is an effective tool for decoding novel words. In the short term, the visual and auditory cue helps readers decode pseudowords, and with prolonged usage, extends to real words and connected-text fluency. In contrast to other technologies which read to the child, *Sound it Out* demonstrates that a digital tool can embed aspects of pedagogy to scaffold a child's reading.

A phonemic image cue is an existing component of many curricula ("Fundations®," n.d.; M. W. Lovett et al., 2000; Moats, 1999), but our studies demonstrate the legitimacy of a digital implementation of a cue during active reading. In fact, our findings suggest a linear, dose-response relationship that parallels those observed in many in-person interventions (Donnelly, Huber, & Yeatman, 2019; Maureen W. Lovett, Borden, DeLuca, Kacerenza, & Et Al, 1994; Lyon & Moats, 1997; Skibbe, Grimm, Bowles, & Morrison, 2012; Stage, Abbott, Jenkins, & Berninger, 2003). While long terms studies are needed to achieve comparable benefits to formal intervention programs, these studies support the literature on the value of technologies to supplement the development of foundational literacy skills (i.e. phonological awareness and sound-letter correspondence) (Cheung & Slavin, 2013; Joseph K. Torgesen et al., 1999).

Additionally, these studies demonstrate the salience of technologies to support both independent and shared reading. Both intervention groups in the final study improved significantly in isolated decoding, and for those readers who were asked to read exclusively with a caregiver present, their gains in reading fluency were superior to their independent reading peers. These results echo the growing literature on the value of technology in empowering parents and caregivers to more effectively support their struggling reader (Kraft & Monti-Nussbaum, 2017; Ronimus & Lyytinen, 2015). Results, combined with parent testimonials, show that *Sound it Out* provides a practical strategy to support structure in daily reading practice. Alone, *Sound it Out* doesn't incorporate all the strategies that have been shown to empower shared reading (Maureen W. Lovett, Barron, & Benson, 2003; Moats, 1999), but our studies show that a digital implementation of one such strategy can be effective.

The ability for technology to encourage productive, shared reading experience is especially meaningful given a troubling decrease observed in the amount of time parents spend reading with their children over the past ten years ("Children, teens, and reading," n.d.). For both typical and struggling readers, enriched reading experiences with a caregiver are especially important for reading development (Chomsky, 1972; Piasta, Justice, McGinty, & Kaderavek, 2012). The rise of e-books and digital readers, particularly targeting youth, would seem to reflect a rise in access and engagement in reading. Instead, research demonstrates that reading engagement is declining (Alter, 2014), and that digital readers have adversely impacted shared reading experiences (Barzillai et al., 2017). Instead of focusing on the linguistic elements important for reading development, caregivers with an electronic reader focus on the game-like, mechanistic affordances of the "book" (Lauricella, Barr, & Calvert, 2014; Segal-Drori, Korat, Shamir, & Klein, 2010).

With these trends as a backdrop, the results of this dissertation suggest a promising direction in educational technologies to provide both content and guidance for effective use. *Sound it Out* represents a flexible and effective tool that can be utilized at-will on a medium with limitless access to decodable content.

5.1 THE FUTURE OF *SOUND IT OUT*

The three studies discussed in this dissertation represent just the beginning for *Sound it Out*. Our results provide a solid foundation for positive effect, but also point to promising future directions for the implementation of this digital tool.

First and foremost, the most pressing future direction for *Sound it Out* is a school implementation. As discussed in Chapter 4, the modularity of *Sound it Out* allows the specific symbols to match those that are used in individual classrooms. Another advantage is its ability to instantly augment text: the phonemic cues can be generated for any input text making it a tool that can be effectively applied to any existing curriculum. In a partnership with schools and teachers, and over the course of an academic year, future studies should explore the value of *Sound it Out* as a curricular extension.

Building on (or in concert with) future work in schools, a second promising research direction is the incorporation of peer-based learning frameworks. *Sound it Out* is a relatively simple tool that we have shown can be taught in a single training session. In the spirit of a Vygotskian zone of proximal development (Vygotsky, 1978), and has been shown in existing research (Fuchs, Fuchs, & Abramson, 2020), including siblings and peers with varied proficiency may encourage use of the image cue. Combined with school-based studies, this research direction can provide evidence for the ideal implementation for motivated adoption.

This relates to the third promising research direction: generalization. With longitudinal studies on the scale of months or years, research should explore the ability of cued reading to generalize to non-cued reading. The goal of *Sound it Out* is not to provide a metaphorical corrective lens to permanently change the way struggling readers engage with text; rather, the goal is to provide training wheels that can be removed once the vowel representations are automatic. A feature of *Sound it Out* is that it can be turned on and off at the level of full passages or individual words. A promising dimension of this research, then, should capitalize on this functionality to explore how use changes over time based on regularity and complexity of orthographic and morphological units.

5.2 THE IMPORTANCE OF EQUALITY IN EDTECH

Although technology has become commonplace in many American homes and classrooms, there are still persistent barriers and inequities in access to digital learning tools. The sobering statistics that show that nearly two thirds of American fourth graders are not reading with proficiency are more pronounced in ethnic and minority populations (National Assessment of Educational Progress, 2019), and inequitable access to technology is considered a major factor in worsening this trend (White, Kim, Chen, & Liu, 2015). Educational institutions place increasing importance on digital literacy, yet run counter to widening gaps along a continua including race, ethnicity, gender, and socio-economic status in digital access (Journell, 2007; Stevenson, 2009).

The gap, however, is not only one of access to devices. A significant segment of the population may have access to a device (i.e. cell phone, library computer, etc.) but lack the ability to access information about high-quality apps and resources (Jenkins, 2009). Lamentably, those experiencing a lack of digital access or lack of digital participation are often those that also have few books in the home. Thus, the issue of equity in educational technologies cannot be reduced to one of just increasing access to tools and services (Neuman & Celano, 2006). As stated in a recent report, “access is no longer just a yes/no question. The quality of families’ Internet connections, and the kinds and capabilities of devices they can access, have considerable consequences for parents and children alike” (Rideout & Katz, 2016).

These realities are essential considerations for educational technologies such as *Sound it Out*. A consistent mistake that developers make is assuming that in merely providing a service for a challenged population, there will be widespread improvement and learning. The studies discussed in this dissertation show that is not the case: only with training, motivated use, and monitored practice can we observe measurable gains. Thus, for *Sound it Out* to be successful and equitable, a concerted effort should be made to ensure that resistance to use is not an issue of digital access or participation by the many disenfranchised children and families that it has the potential to aid.

In addition to equality in access is an issue of equality in cultural representation in mainstream education (Hattan & Lupo, 2020). A great majority of standard curricula focus on majority perspectives, dialect, and content that does not account for the sociocultural and sociolinguistic diversity of the student population (Noguerón-Liu, 2020). For reading instruction, the focus on Standard American English dialect and pronunciations may be important for supporting academic knowledge, but has also been shown to be counterproductive for those emerging readers who have linguistically diverse home environments (Brown et al., 2015; Washington, 2001). This creates a cultural asynchrony between the school and the home where students cannot relate to the content of the classroom, and families cannot use the pedagogical tools of the classroom for the cultural content of the home.

This inequality is another essential consideration for digital tools such as *Sound it Out*. With the inherent modularity of digital tools, a focus of research and development should be ensuring that any affordance can adapt to cultural variability in schools, families, and individual readers. For *Sound it Out* the symbols, and the culturally determined delineation of meaningful vowel sounds they represent, are subject-to-change. This flexibility means that the tool can be used to increase access to culturally meaningful content using culturally relevant tools.

5.3 CONCLUSION

Sound it Out represents one of a growing number of educational technologies that combine pedagogical practice and technical expertise to create impactful tools for struggling readers. There is still much to be done to determine the ideal place for technology in education, but our partnership to develop *Sound it Out* brings the field one step closer. As our work has shown, the value in technology is not to revolutionize or change the way that we read, but to leverage its unique affordances to augment, equalize, and democratize the teaching of reading.

BIBLIOGRAPHY

Abrami, P. C., Lysenko, L., & Borokhovski, E. (2020). The effects of ABRACADABRA on reading outcomes: An updated meta-analysis and landscape review of applied field research. *Journal of Computer Assisted Learning*, jcal.12417.

Abu-Rabia, S. (1999). The effect of arabic vowels on the reading comprehension of second- and sixth-grade native arab children. *Journal of Psycholinguistic Research*, *28*(1), 93–101.

Abu-Rabia, S. (2001). The role of vowels in reading Semitic scripts: Data from Arabic and Hebrew. *Reading and Writing*, *14*(1–2), 39–59.

Alexander, P. A. (2020). What research has revealed about readers' struggles with comprehension in the digital age: Moving beyond the phonics versus whole language debate. *Reading Research Quarterly*, (rrq.331). doi:10.1002/rrq.331

Alter, C. (2014, May 12). Study: The number of teens reading for fun keeps declining. *Time*. Retrieved from https://time.com/94794/common-sense-media-reading-report-never-read/

Atkinson, R. C., & Hansen, D. N. (1966). Computer-Assisted Instruction in Initial Reading: The Stanford Project. *Reading Research Quarterly*, *2*(1), 5.

Auerbach, E. R. (1989). Toward a Social-Contextual Approach to Family Literacy. *Harvard Educational Review*, *59*(2), 165–182.

Barker, T. A., & Torgesen, J. K. (1995). An Evaluation of Computer-Assisted Instruction in Phonological Awareness with below Average Readers. *Journal of Educational Computing Research*, *13*(1), 89–103.

Barzillai, M., Thomson, J. M., & Mangen, A. (2017). The influence of e-books on language and literacy development. In K. Sheehy & A. Holliman (Eds.), *Education and New Technologies: Perils and Promises for Learners* (pp. 33–47). New York, NY: Routledge.

Becker, H. J. (1992). Computer-Based Integrated Learning-Systems in the Elementary and Middle Grades - A Critical-Review and Synthesis of Evaluation Reports. *Journal of Educational Computing Research*, *8*(1), 1–41.

Benton, L., Vasalou, A., Berkling, K., Barendregt, W., & Mavrikis, M. (2018). A Critical Examination of Feedback in Early Reading Games. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–12.

Benton, L., Vasalou, A., Khaled, R., Johnson, H., & Gooch, D. (2014). Diversity for Design: A Framework for Involving Neurodiverse Children in the Technology Design Process. *Proceedings of Conference on Human Factors in Computing Systems*, 3747–3756.

Berninger, V. W., Lee, Y. L., Abbott, R. D., & Breznitz, Z. (2013). Teaching children with dyslexia to spell in a reading-writers' workshop. *Annals of Dyslexia*, *63*(1), 1–24.

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-Assisted Instruction in Support of Beginning Reading Instruction: A Review. *Review of Educational Research*, *72*(1), 101–130.

Bowles, N. (2018, October 26). The Digital Gap Between Rich and Poor Kids Is Not What We Expected. *The New York Times*. Retrieved from https://www.nytimes.com/2018/10/26/style/digital-divide-screens-schools.html

Bradley L, & Bryant P E. (1983). Categorizing sounds and learning to read -- a causal connection. *Nature*, *301*(3), 419–421.

Brown, M. C., Sibley, D. E., Washington, J. A., Rogers, T. T., Edwards, J. R., MacDonald, M. C., & Seidenberg, M. S. (2015). Impact of dialect use on a basic component of learning to read. *Frontiers in Psychology*, *6*(MAR), 1–17.

Bryant, N. D. (1965). Some Principles of Remedial Instruction for Dyslexia Some Principles of Remedial Instruction for Dyslexia. *The Reading Teacher*, *18*(7), 567–572.

Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). *Effectiveness of Reading and Mathematics Software Products: Findings From Two Student Cohorts*. Retrieved from National Center for Education Evaluation and Regional Assistance. Available from: ED Pubs. P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827; Web site: http://ies.ed.gov/ncee/pubs/ website: https://eric.ed.gov/?id=ED504657

Cassidy, J., Ortlieb, E., & Grote-Garcia, S. (2020). What's Hot in Literacy: New Topics and New Frontiers are Abuzz. *Literacy Research and Instruction*, 1–12.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars : Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, *19*(1), 5–51.

Chall, J. S. (1983). *Learning to Read: The Great Debate*. New York: McGraw-Hill Book Company.

Cheung, A. C. K., & Slavin, R. E. (2011). The effectiveness of education technology for enhancing reading achievement : A meta-analysis. *Best Evidence Encyclopaedia*, *97*(January), 1–48.

Cheung, A. C. K., & Slavin, R. E. (2013). Effects of Educational Technology Applications on Reading Outcomes for Struggling Readers: A Best-Evidence Synthesis. *Reading Research Quarterly*, *48*(3), 277–299.

Children, teens, and reading. (n.d.). Retrieved November 14, 2020, from https://www.commonsensemedia.org/research/children-teens-and-reading

Chomsky, C. (1972). Stages in Language Development and Reading Exposure. *Harvard Educational Review*, *42*(1), 1–33.

Christ, T., Arya, P., & Liu, Y. (2018). Technology Integration in Literacy Lessons: Challenges and Successes. *Literacy Research and Instruction*, *00*(00), 1–18.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Review of Educational Research*, *86*(1), 79–122.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256.

Daugherty, L., Dossani, R., Johnson, E.-E., & Wright, C. (2014). *Getting on the Same Page: Identifying Goals for Technology Use in Early Childhood Education*. Retrieved from RAND Corporation website: https://www.rand.org/pubs/research_reports/RR673z1.html

de Souza, G. N., Brito, Y. P. dos S., Tsutsumi, M. M. A., Marques, L. B., Goulart, P. R. K., Monteiro, D. C., & Santana, Á. L. de. (2018). The Adventures of Amaru: Integrating Learning Tasks Into a Digital Game for Teaching Children in Early Phases of Literacy. *Frontiers in Psychology*, *9*(DEC), 2531.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262.

Delany, K. (2017). The Experience of Parenting a Child With Dyslexia: An Australian perspective. *Journal of Student Engagement: Education Matters*, *7*(1). Retrieved from http://ro.uow.edu.au/jseem/vol7/iss1/6

Dexter, S., & Richardson, J. W. (2019). What does technology integration research tell us about the leadership of technology? *Journal of Research on Technology in Education*, 1–20.

Donnelly, P. M., Huber, E., & Yeatman, J. D. (2019). Intensive Summer Intervention Drives Linear Growth of Reading Skill in Struggling Readers. *Frontiers in Psychology*, *10*, 1900.

Donnelly, P. M., Larson, K., Matskewich, T., & Yeatman, J. D. (2020). Annotating digital text with phonemic cues to support decoding in struggling readers. *PloS One*, *15*(12). doi:10.1371/journal.pone.0243435

Dowker, A. (2005). Early identification and intervention for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*(4), 324–332.

Drummond, K., Chinen, M., & Duncan, T. G. (2011). Impact of the Thinking Reader ® Software Program on Grade 6 Reading Vocabulary , Comprehension , Strategies , and Motivation. *Ncee 2010-4035*, 184.

Eagleton, M. B., & Dobler, E. (2007). Reading the Web. *Strategies for Internet Inquiry*.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, *18*(2), 116–125.

Ehri, L. C. (2014). Orthographic Mapping in the Acquisition of Sight Word Reading, Spelling Memory, and Vocabulary Learning. *Scientific Studies of Reading: The Official Journal of the Society for the Scientific Study of Reading*, *18*(1), 5–21.

Ehri, L. C. (2020). The Science of Learning to Read Words: A Case for Systematic Phonics Instruction. *Reading Research Quarterly*, *55*(S1), 175.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic Awareness Instruction Helps Children Learn to Read: Evidence From the National Reading Panel's Meta-Analysis. *Reading Research Quarterly*, *36*(3), 250–287.

Elbeheri, G., & Everatt, J. (2007). Literacy ability and phonological processing skills amongst dyslexic and non-dyslexic speakers of Arabic. *Reading and Writing*, *20*(3), 273–294.

Erickson, G. C., Foster, K. C., Foster, D. F., Torgesen, J. K., & Packer, S. (1992). DaisyQuest. *Scotts Valley, CA: Great Wave Software*.

Farkas, W. A., & Jang, B. G. (2019). Designing, Implementing, and Evaluating a School-Based Literacy Program for Adolescent Learners With Reading Difficulties: A Mixed-Methods Study. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 1–17.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fletcher, J. M., Reid Lyon, G., Fuchs, L. S., & Barnes, M. A. (2018). *Learning Disabilities, Second Edition: From Identification to Intervention*. Guilford Publications.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. D. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, *90*(1), 37–55.

Francom, G. M. (2019). Barriers to technology integration: A time-series survey study. *Journal of Research on Technology in Education*, 1–16.

Frijters, J. C., Tsujimoto, K. C., Boada, R., Gottwald, S., Hill, D., Jacobson, L. A., … Gruen, J. R. (2017). Reading-Related Causal Attributions for Success and Failure: Dynamic Links With Reading Skill. *Reading Research Quarterly*. doi:10.1002/rrq.189

Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies: Promoting word recognition, fluency, and reading comprehension in young children. *The Journal of Special Education*, *39*(1), 34–44.

Fuchs, D., Fuchs, L. S., & Abramson, R. (2020). Peer-Assisted Learning Strategies (PALS): A Validated Classwide Program for Improving Reading and Mathematics Performance. In A. L. Reschly, A. J. Pohl, & S. L. Christenson (Eds.), *Student Engagement: Effective Academic, Behavioral, Cognitive, and Affective Interventions at School* (pp. 109–120). Cham: Springer International Publishing.

Fundations®. (n.d.). Retrieved October 29, 2020, from https://www.wilsonlanguage.com/programs/fundations/

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment*, *1*(1), 20.

Gilbert, L., Teravainen, A., Clark, C., & Shaw, S. (2018). *Literacy and life expectancy: An evidence review exploring the link between literacy and life expectancy in England through health and socioeconomic factors*. London: National Literacy Trust.

Good, R. H., & Kaminski, R. A. (Eds.). (2007). *Dynamic Indicators of Basic Early Literacy Skills (6th ed.)*. Retrieved from Institute for the Development of Educational Achievement website: http://dibels.uoregon.edu/

Goodall, J. (2017). *Narrowing the achievement gap: Parental engagement with children's learning*. London: Routledge.

Grant, A. K., Wood, E., Gottardo, A., Evans, M. A., Phillips, L., & Savage, R. S. (2012). Assessing the Content and Quality of Commercially Available Reading Software

Programs: Do They Have the Fundamental Structures to Promote the Development of Early Reading Skills in Children? *NHSA Dialog*, *15*(4), 319–342.

Greenfield, P. M. (2009, January 2). Technology and informal education: What is taught, what is learned. *Science*, Vol. 323, pp. 69–71. doi:10.1126/science.1167190

Guernsey, L., & Levine, M. H. (2015). *Tap, Click, Read: Growing Readers in a World of Screens*. San Francisco: Jossey-Bass.

Haddad, L., Weiss, Y., Katzir, T., & Bitan, T. (2018). Orthographic Transparency Enhances Morphological Segmentation in Children Reading Hebrew Words. *Frontiers in Psychology*, *8*(JAN), 2369.

Hannon, P., & James, S. (1990). Parents' and Teachers' Perspectives on Preschool Literacy Development. *British Educational Research Journal*, *16*(3), 259–272.

Harris, J., & Hofer, M. (2009). Grounded Tech Integration: An Effective Approach Based on Content, Pedagogy, and Teacher Planning. *Learning & Leading with Technology*. Retrieved from https://eric.ed.gov/?id=EJ859576

Hattan, C., & Lupo, S. M. (2020). Rethinking the Role of Knowledge in the Literacy Classroom. *Reading Research Quarterly*, *55*(S1), 66.

Hindman, A. H., Morrison, F. J., Connor, C. M. D., & Connor, J. A. (2020). Bringing the Science of Reading to Preservice Elementary Teachers: Tools That Bridge Research and Practice. *Reading Research Quarterly*, *55*(S1), S197–S206.

Hoffman, J. V., Hikida, M., & Sailors, M. (2020). Contesting Science That Silences: Amplifying Equity, Agency, and Design Research in Literacy Teacher Preparation. *Reading Research Quarterly*, *55*(S1), S255–S266.

Hornung, C., Martin, R., & Fayol, M. (2017). The power of vowels: Contributions of vowel, consonant and digit RAN to clinical approaches in reading development. *Learning and Individual Differences*, *57*, 85–102.

Hurford, D. P., & Hurford, J. D. (2016). The Dyslexia Dilemma: A History of Ignorance, Complacency and Resistance in Colleges of Education. *Journal of Childhood & Developmental Disorders*, *2*(3). doi:10.4172/2472-1786.100034

Hutchison, A., Beschorner, B., & Schmidt-Crawford, D. (2012). Exploring the Use of the iPad for Literacy Learning. *The Reading Teacher*, *66*(1), 15–23.

Jenkins, H. (2009). *Confronting the challenges of participatory culture: Media education for the 21st century*. The MIT Press.

Joo, S. J., White, A. L., Strodtman, D. J., & Yeatman, J. D. (2018). Optimizing text for an individual's visual system: The contribution of visual crowding to reading difficulties. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *103*, 291–301.

Journell, W. (2007). The Inequities of the Digital Divide: is e-learning a solution? *E-Learning and Digital Media*, *4*(2), 138–149.

Kahne, J., Middaugh, E., & Evans, C. (2009). *The civic potential of video games*. The MIT Press.

Kim, M. K., McKenna, J. W., & Park, Y. (2017). The Use of Computer-Assisted Instruction to Improve the Reading Comprehension of Students With Learning Disabilities: An Evaluation of the Evidence Base According to the What Works Clearinghouse Standards. *Remedial and Special Education: RASE*, *38*(4), 233–245.

Kraft, M. A., & Monti-Nussbaum, M. (2017). Can Schools Enable Parents to Prevent Summer Learning Loss? A Text-Messaging Field Experiment to Promote Literacy Skills. *The Annals of the American Academy of Political and Social Science*, *674*(1), 85–112.

Kristen D., J. P. B., Ellen, M., Zoi A., P., Maryann, M., Paola, P., & Vintinner, J. P. (2018). Effects of a Summer Reading Intervention on Reading Skills for Low-Income Black and Hispanic Students in Elementary School. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, *34*(3), 1–18.

Kulik, C.-L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, *7*(1–2), 75–94.

Kyle, F. E., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computerassisted reading interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly*, *48*(1), 61–76.

Larson, L. C. (2010). Digital Readers: The Next Chapter in E-Book Reading and Response. *The Reading Teacher*, *64*(1), 15–22.

Lauricella, A. R., Barr, R., & Calvert, S. L. (2014). Parent–child interactions during traditional and computer storybook reading for children's comprehension: implications for electronic storybook design. *International Journal of Child-Computer*. Retrieved from https://www.sciencedirect.com/science/article/pii/S221286891400021X?casa_token=eY4vg6JwggsAAAAA:L5TIjeg-uGR_zPU1Du63DNDggf3oH8trbsTL0geqxGCDBY78eiTMXYEEU_gXI2yO7ITEH8Y3BC0

Laurillard, D. (2016). Learning “Number Sense” through Digital Games with Intrinsic Feedback. *Australasian Journal of Educational Technology*, *32*(6), 32–44.

Lee, Y.-H., Waxman, H., Wu, J.-Y., Michko, G., & Lin, G. (2013). Revisit the Effect of Teaching and Learning with Technology. *Educational Technology & Society*, *16*(1), 133–146.

Legge, G. E., Mansfield, J. S., & Chung, S. T. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, *41*(6), 725–743.

Lerner, J. W. (1989). Educational interventions in learning disabilities. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28*(3), 326–331.

Lindeblad, E., Nilsson, S., Gustafson, S., & Svensson, I. (2017). Assistive technology as reading interventions for children with reading impairments with a one-year follow-up. *Disability and Rehabilitation. Assistive Technology*, *12*(7), 713–724.

Lovett, M. W., Lacerenza, L., & Borden, S. L. (2000). Putting struggling readers on the PHAST track: a program to integrate phonological and strategy-based remedial reading instruction and maximize outcomes. *Journal of Learning Disabilities*, *33*(5), 458–476.

Lovett, Maureen W., Barron, R. W., & Benson, N. J. (2003). Effective remediation of word identification and ecoding difficulties in school-age children with reading disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of Learning Disabilities*. New York, NY: The Guilford Press.

Lovett, Maureen W., Borden, S. L., DeLuca, T., Kacerenza, L., & Et Al. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategy-based reading training programs. *Developmental Psychology*, *30*(July), 805–822.

Lovett, Maureen W., Frijters, J. C., Wolf, M., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2017). Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology*, *109*(7), 889–914.

Lovett, Maureen W., Lacerenza, L., Borden, S. L., Frijters, J. C., Steinbach, K. A., & De Palma, M. (2000). Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes. *Journal of Educational Psychology*, *92*(2), 263–283.

Lynch, J., Anderson, J., Anderson, A., & Shapiro, J. (2006). Parents' Beliefs About Young Children's Literacy Development And Parents' Literacy Behaviors. *Reading Psychology*, *27*(1), 1–20.

Lyon, G. R., & Moats, L. (1997). Critical conceptual and methodological considerations in reading intervention research. *Journal of Learning Disabilities*, *30*(6), 578.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*(1), 1–14.

Macaruso, P., Hook, P. E., & McCabe, R. (2006). The efficacy of computer-based supplementary phonics programs for advancing reading skills in at-risk elementary students. *Journal of Research in Reading*, *29*(2), 162–172.

Macgilchrist, F. (2018). Cruel optimism in edtech: when the digital data practices of educational technology providers inadvertently hinder educational equity. *Learning, Media and Technology*, *9884*. doi:10.1080/17439884.2018.1556217

Mahfoudhi, A., Everatt, J., & Elbeheri, G. (2011, October 10). Introduction to the special issue on literacy in Arabic. *Reading and Writing*, Vol. 24, pp. 1011–1018. doi:10.1007/s11145-011-9306-y

Mandel Morrow, L., & Asbury, E. (1999). What Should We Do About Phonics? In L. B. Gambrell, L. Mandel Morrow, S. B. Neuman, & M. Pressley (Eds.), *Best Practices in Literacy Instruction* (pp. 68–89). New York, NY: The Guilford Press.

Maughan, B. (1995). Annotation: long-term outcomes of developmental reading problems. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *36*(3), 357–371.

McDonald Connor, C., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., … Morrison, F. J. (2009). Individualizing Student Instruction Precisely: Effects of Child × Instruction Interactions on First Graders' Literacy Development. *Child Development*, *80*(1), 77–100.

McTigue, E. M., Solheim, O. J., Zimmer, W. K., & Uppstad, P. H. (2020). Critically Reviewing GraphoGame Across the World: Recommendations and Cautions for Research and Implementation of Computer-Assisted Instruction for Word-Reading Acquisition. *Reading Research Quarterly*, *55*(1), 45–73.

Medler, D. A., & Binder, J. R. (2005). MCWord: An Orthographic Wordform Database. Retrieved from http://www.neuro.mcw.edu/mcword/

Mesmer, H. A. E., & Williams, T. O. (2015). Examining the Role of Syllable Awareness in a Model of Concept of Word: Findings From Preschoolers. *Reading Research Quarterly*, *50*(4), 483–497.

Messer, D., & Nash, G. (2018). An evaluation of the effectiveness of a computer-assisted reading intervention. *Journal of Research in Reading*, *41*(1), 140–158.

Moats, L. C. (1999). *Teaching Reading Is Rocket Science: What Expert Teachers of Reading Should Know and Be Able To Do*. Retrieved from American Federation of Teachers website: https://eric.ed.gov/?id=ED445323

Morris, R. D., Lovett, M. W., Wolf, M., Sevcik, R. A., Steinbach, K. A., Frijters, J. C., & Shapiro, M. B. (2012). Multiple-Component Remediation for Developmental Reading Disabilities: IQ, Socioeconomic Status, and Race as Factors in Remedial Outcome. *Journal of Learning Disabilities*, *45*(2), 99–127.

National Assessment of Educational Progress. (2006). *The Nation's Report Card. Reading 2005*. Retrieved from U.S. Department of Education, Institute of Education Sciences website: http://files.eric.ed.gov/fulltext/ED486463.pdf

National Assessment of Educational Progress. (2019). *NAEP Report Card: Reading*. Retrieved from U.S. Department of Education website: https://www.nationsreportcard.gov/reading?grade=4

National Reading Panel. (2000). *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction*. Washington, D.C.: NICHD.

Neuman, S. B., & Celano, D. (2006). The knowledge gap: Implications of leveling the playing field for low-income and middle-income children. *Reading Research Quarterly*, *41*(2), 176–201.

Noguerón-Liu, S. (2020). Expanding the Knowledge Base in Literacy Instruction and Assessment: Biliteracy and Translanguaging Perspectives From Families, Communities, and Classrooms. *Reading Research Quarterly*, *55*(S1), 355.

Norton, E. S., & Wolf, M. (2012). Rapid Automatized Naming (RAN) and Reading Fluency: Implications for Understanding and Treatment of Reading Disabilities. *Annual Review of Psychology*, *63*, 427–452.

Oakland, T., Black, J. L., Stanford, G., Nussbaum, N. L., & Balise, R. R. (1998). An Evaluation of the Dyslexia Training Program: A Multisensory Method for Promoting Reading in Students with Reading Disabilities. *Journal of Learning Disabilities*, *31*(2), 140–147.

O'Brien, B. A., Habib, M., & Onnis, L. (2019). Technology-Based Tools for English Literacy Intervention: Examining Intervention Grain Size and Individual Differences. *Frontiers in Psychology*, *10*, 2625.

Ok, M. W., & Rao, K. (2019). Digital Tools for the Inclusive Classroom: Google Chrome as Assistive and Instructional Technology. *Journal of Special Education Technology*, *34*(3), 204–211.

Olson, R. K., & Wise, B. W. (1992). Reading on the computer with orthographic and speech feedback. *Reading and Writing*, *4*(2), 107–144.

Perie, M., Grigg, W., & Donahue, P. (2005). The Nation's Report Card [TM]: Reading, 2005. NCES 2006-451. *National Center for Education Statistics*. Retrieved from https://eric.ed.gov/?id=ED486463

Petscher, Y., Cabell, S. Q., Catts, H. W., Compton, D. L., Foorman, B. R., Hart, S. A., … Wagner, R. K. (2020). How the Science of Reading Informs 21st-Century Education. *Reading Research Quarterly*, *55*, 267–282.

Phillips Galloway, E., McClain, J. B., & Uccelli, P. (2020). Broadening the Lens on the Science of Reading: A Multifaceted Perspective on the Role of Academic Language in Text Understanding. *Reading Research Quarterly*, *55*(S1), S331–S345.

Piasta, S. B., Justice, L. M., McGinty, A. S., & Kaderavek, J. N. (2012). Increasing young children's contact with print during shared reading: longitudinal effects on literacy achievement. *Child Development*, *83*(3), 810–820.

Rabiner, D. L., Murray, D. W., Skinner, A. T., & Malone, P. S. (2010). A randomized trial of two promising computer-based interventions for students with attention difficulties. *Journal of Abnormal Child Psychology*, *38*(1), 131–142.

Raskind, M. H., Goldberg, R. J., Higgins, E. L., & Herman, K. L. (1999). Patterns of Change and Predictors of Success in Individuals With Learning Disabilities: Results From a Twenty-Year Longitudinal Study. *Learning Disabilities Research & Practice: A Publication of the Division for Learning Disabilities, Council for Exceptional Children*, *14*(1), 35–49.

Richardson, U., & Lyytinen, H. (2014). The GraphoGame Method: The Theoretical and Methodological Background of the Technology-Enhanced Learning Environment for Learning to Read. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, *10*(1), 39–60.

Rideout, V., & Katz, V. S. (2016). Opportunity for All? Technology and Learning in Lower-Income Families. *Joan Ganz Cooney Center at Sesame Workshop*. Retrieved from https://eric.ed.gov/?id=ED574416

Ronimus, M., & Lyytinen, H. (2015). Is School a Better Environment Than Home for Digital Game-Based Learning? The Case of GraphoGame. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, *11*(2), 123–147.

Rose, D. H., & Strangman, N. (2007, April 4). Universal Design for Learning: Meeting the challenge of individual learning differences through a neurocognitive perspective. *Universal Access in the Information Society*, *5*, 381–391. Springer.

Scanlon, D. M., & Anderson, K. L. (2020). Using Context as an Assist in Word Solving: The Contributions of 25 Years of Research on the Interactive Strategies Approach. *Reading Research Quarterly*, *55*(S1), 73.

Scanlon, D. M., & Vellutino, F. R. (1997). A Comparison of the Instructional Backgrounds and Cognitive Profiles of Poor, Average, and Good Readers Who Were Initially Identified as At Risk for Reading Failure. *Scientific Studies of Reading: The Official Journal of the Society for the Scientific Study of Reading*, *1*(3), 191–215.

Schiff, R., & Saiegh-Haddad, E. (2018). Development and Relationships Between Phonological Awareness, Morphological Awareness and Word Reading in Spoken and Standard Arabic. *Frontiers in Psychology*, *9*(APR), 356.

Scott, J. C., & Marcus, C. D. (2014). Emergent literacy: Home--school connections. In *Literacy in African American communities* (pp. 105–126). Routledge.

Segal-Drori, O., Korat, O., Shamir, A., & Klein, P. S. (2010). Reading electronic and printed books with and without adult instruction: effects on emergent reading. *Reading and Writing*, *23*(8), 913–930.

Seidenberg, M. S. (2005). Connectionist Models of Word Reading. *Current Directions in Psychological Science*, *14*(5), 238–242.

Seidenberg, M. S. (2017). *Language at the speed of sight: how we read, why so many can't, and what can be done about it*. New York, NY: Basic Books.

Seidenberg, M. S., Cooper Borkenhagen, M., & Kearns, D. M. (2020). Lost in Translation? Challenges in Connecting Reading Science and Educational Practice. *Reading Research Quarterly*, *55*(S1), S119–S130.

Seward, R., O'Brien, B., Breit-Smith, A. D., & Meyer, B. (2014). Linking Design Principles with Educational Research Theories to Teach Sound to Symbol Correspondence with Multisensory Type. *Visible Language*, *48*(3), 87–108.

Shaheen, N. L., & Lohnes Watulak, S. (2019). Bringing Disability Into the Discussion: Examining Technology Accessibility as An Equity Concern in the Field of Instructional Technology. *Journal of Research on Technology in Education*, *51*(2), 187–201.

Shany, M., Bar-On, A., & Katzir, T. (2012). Reading different orthographic structures in the shallow-pointed Hebrew script: a cross-grade study in elementary school. *Reading and Writing*, *25*(6), 1217–1238.

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218; discussion 219-26.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., … Gore, J. C. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences*, *95*(5), 2636–2641.

Skibbe, L. E., Grimm, K. J., Bowles, R. P., & Morrison, F. J. (2012). Literacy Growth in the Academic Year Versus Summer From Preschool Through Second Grade: Differential Effects of Schooling Across Four Skills. *Scientific Studies of Reading: The Official Journal of the Society for the Scientific Study of Reading*, *16*(2), 141–165.

Slavin, R. E., Cheung, A. C. K., Groff, C., & Lake, C. (2008). Effective Reading Programs for Middle and High Schools: A Best-Evidence Synthesis. *Reading Research Quarterly*, *43*(3), 290–322.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2010). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, *6*, 1–26.

Snow, C. E., Burns, M. S., & Griffin, P. (1999). *Preventing reading difficulties in young children* (pp. 148–155).

Snowling, M. J. (2004). The Science of Dyslexia: A Review of Contemporary Approaches. In M. Turner & J. Rack (Eds.), *The Study of Dyslexia* (pp. 77–90). Boston, MA: Springer US.

Soe, K., Koki, S., & Chang, J. M. (2000). Effect of Computer-Assisted instruction (CAI) on reading Achievement: A Meta-Analysis. *Pacific Resources for Education and Learning*. Retrieved from https://eric-ed-gov.offcampus.lib.washington.edu/?id=ED443079

Solheim, O. J., Frijters, J. C., Lundetræ, K., & Uppstad, P. H. (2018). Effectiveness of an early reading intervention in a semi-transparent orthography: A group randomised controlled trial. *Learning and Instruction*, *58*(July 2017), 65–79.

Stage, S. A., Abbott, R. D., Jenkins, J. R., & Berninger, V. W. (2003). Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word readings discrepancy: Failure to validate discrepancy method. *Journal of Learning Disabilities*, *36*(1), 24–33.

Stetter, M. E., & Hughes, M. T. (2015). Computer-Assisted Instruction to Enhance the Reading Comprehension of Struggling Readers: A Review of the Literature. *Journal of Special Education Technology*, *25*(4), 1–16.

Stevenson, S. (2009). Digital Divide: A Discursive Move Away from the Real Inequities. *The Information Society*, *25*(1), 1–22.

Storch, S. a., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental Psychology*, *38*(6), 934–947.

Suggate, S. P. (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental Psychology*, *46*(6), 1556–1579.

Sum, A., Kirsch, I. S., & Yamamoto, K. (2004). *A human capital concern: The literacy proficiency of US immigrants*. Policy Information Center, Educational Testing Service Princeton, NJ, USA.

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and Pitfalls of Multimedia and Interactive Features in Technology-Enhanced Storybooks. *Review of Educational Research*, *85*(4), 698–739.

Tare, M., Shell, A. R., & Jackson, S. R. (2020). Student engagement with evidence-based supports for literacy on a digital platform. *Journal of Research on Technology in Education*, *0*(0), 1–11.

Taylor, D. B., Handler, L. K., FitzPatrick, E., & Whittingham, C. E. (2020). The device in the room: Technology's role in third grade literacy instruction. *Journal of Research on Technology in Education*. doi:10.1080/15391523.2020.1747577

The Annie E. Casey Foundation. (2013). *Early warning confirmed: a research update on third-grade reading*. Retrieved from https://www.aecf.org/resources/early-warning-confirmed/

The MathWorks, I. (2017). *MATLAB Statistics and Machine Learning Toolbox*. Natick, MA, USA.

Tichnor-Wagner, A., Garwood, J. D., Bratsch-Hines, M., & Vernon-Feagans, L. (2016). Home Literacy Environments and Foundational Literacy Skills for Struggling and Nonstruggling Readers in Rural Early Elementary Schools. *Learning Disabilities Research & Practice: A Publication of the Division for Learning Disabilities, Council for Exceptional Children*, *31*(1), 6–21.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, *34*(1), 33–58, 78.

Torgesen, Joseph K., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia*, *60*(1), 40–56.

Torgesen, Joseph K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*, *91*(4), 579–593.

Vaughn, S., Fletcher, J. M., Francis, D. J., Denton, C. A., Wanzek, J., Wexler, J., … Romain, M. A. (2008). Response to intervention with older students with reading difficulties. *Learning and Individual Differences*, *18*(3), 338–345.

Vellutino, F. R., & Scanlon, D. M. (2002). The Interactive Strategies approach to reading intervention. *Contemporary Educational Psychology*, *27*(4), 573–635.

Vygotsky, L. S. (1978). *Mind in society : the development of higher psychological processes* (p. 159; M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N. K., Vaughn, S., & Sargent, K. (2018). Current Evidence on the Effects of Intensive Early Reading Interventions. *Journal of Learning Disabilities*, 002221941877511.

Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive Reading Interventions for Students With Reading Difficulties After Grade 3. *Review of Educational Research*, *83*(2), 163–195.

Washington, J. A. (2001). Early Literacy Skills in African-American Children: Research Considerations. *Learning Disabilities Research & Practice: A Publication of the Division for Learning Disabilities, Council for Exceptional Children*, *16*(4), 213–221.

Weiss, Y., Katzir, T., & Bitan, T. (2015a). Many ways to read your vowels--Neural processing of diacritics and vowel letters in Hebrew. *NeuroImage*, *121*, 10–19.

Weiss, Y., Katzir, T., & Bitan, T. (2015b). The effects of orthographic transparency and familiarity on reading Hebrew words in adults with and without dyslexia. *Annals of Dyslexia*, *65*(2), 84–102.

White, S., Kim, Y. Y., Chen, J., & Liu, F. (2015). Performance of Fourth-Grade Students in the 2012 NAEP Computer-Based Writing Pilot Assessment: Scores, Text Length, and Use of Editing Tools. Working Paper Series. NCES 2015-119. *National Center for Education Statistics*. Retrieved from https://eric.ed.gov/?id=ED562627

Willis, A. I. (2019). Race, Response to Intervention, and Reading Research. *Journal of Literacy Research: JLR*, *51*(4), 394–419.

Wise, B. W., & Olson, R. K. (1995). Computer-based phonological awareness and reading instruction. *Annals of Dyslexia*, *45*(1), 97–122.

Wise, B. W., & Olson, R. K. (1998). Studies of computer-aided remediation for reading disabilities. In C. Hulme & R. M. Joshi (Eds.), *Reading and Spelling: Development and Disorders* (pp. 473–487). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Wolf, M. (2007). *Proust and the Squid: The Story and Science of the Reading Brain*. New York: HarperCollins.

Wolf, M. (2018). *Reader, Come Home*. New York, NY: HarperCollins.

Wolf, M., Barzillai, M., Gottwald, S., Miller, L., Spencer, K., Norton, E. S., … Morris, R. D. (2009). The RAVE-O intervention: Connecting neuroscience to the classroom. *Mind, Brain and Education: The Official Journal of the International Mind, Brain, and Education Society*, *3*(2), 84–93.

Wolf, M., & Bowers, P. G. (1997). Naming-Speed Processes and Developmental Reading Disabilities: An Introduction to the Special Issue on the Double-Deficit Hypothesis. *Journal of Learning Disabilities*, *33*(4), 322–324.

Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-Speed Processes, Timing, and Reading. *Journal of Learning Disabilities*, *33*(4), 387–407.

Wolf, M., Gottwald, S., Breazeal, C., Galyean, T., & Morris, R. (2017). "I Hold Your Foot:" Lessons from the Reading Brain for Addressing the Challenge of Global Literacy. In *Children and Sustainable Development* (pp. 225–238). Cham: Springer International Publishing.

Wolf, M., Gottwald, S., Galyean, T., Morris, R., & Breazeal, C. (2014). The Reading Brain, Global Literacy, and the Eradication of Poverty. *Proceedings of Bread and Brain, Education and Poverty*, 1–22. Vatican City: Pontificail Academy of Sciences.

Wolf, M., & Katzir-Cohen, T. (2001). Reading Fluency and Its Intervention. *Scientific Studies of Reading: The Official Journal of the Society for the Scientific Study of Reading*, *5*(3), 211–239.

Zuckerman, B. (2017). A guiding framework for considering touchscreens in children under two. *International Journal of Child-Computer Interaction*, *12*, 46–49.

APPENDIX A

A.1 SUPPLEMENTARY MATERIAL FOR DONNELLY, ET AL. 2020 (IN REVIEW)

A.1.1 Role of Age and Phonological Awareness as Covariates to Mixed Effects Models

Due to the heterogeneity of our struggling reader population in both reading ability and age, we performed an exploratory analysis of these results testing a model that adds covariates for participant age and initial phonological awareness ability (as measured using the CTOPP phonological awareness composite measure).

Model fits were compared using AIC/BIC values and revealed that in no case was the model with added covariates a superior model to the simpler model used in the manuscript. Looking at the results of these models, moreover, revealed no significant effects of age or phonological awareness. Interaction effects were unchanged from the simpler models reported in the manuscript.

The model fit workflow and code associated with this analysis can be found in the associated project repository on GitHub.

Table A.1: Post-study questionnaire for intervention group

	Yes	No	Maybe	Did not respond	Total
[Child] "Did you like the app?"	17	2	0	1	20
[Child] "Would you use the app again in the future?"	17	1	1	1	20
[Adult] "Did you like the app?"	16	0	0	4	20

Listed are the responses to the post-study questionnaire for the intervention group participants and their parent. After completing the study, children were asked to answer honestly to the following questions: Did you like the app? And would you like to use the app again in the future? Parents were then asked if they enjoyed using the app. Those adults who did not respond did not participate in the practice to comment on the app.

Table A.2: Real word frequency statistics

Word List 1		Word List 2		Word List 3		Word List 4	
Word	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**
has	1809.24	his	5828.33	had	6491.25	can	1954.34
you	0.18	now	1683.88	she	4270.81	one	3302.38
was	11050.2	who	4754.07	the	61445.7	for	8319.47
not	5082.53	are	4254.92	all	3510.01	her	4032.24
but	5240.18	him	2617.38	and	28470.7	out	2521.24
then	1820	were	3589.55	this	4504.73	only	1774.31
from	4229.52	time	1770.21	when	2560.63	into	2006.57
they	4734.61	what	2581.57	that	11437.3	have	4529.54
been	2716.62	more	2427.66	will	2105.87	them	2305.28
said	2838.7	like	1884.02	with	7066.31	some	1859.62
again	766.802	other	1596.96	about	2322.72	right	758.533
could	1906.39	world	739.197	where	1034.58	those	866.215
their	2947.33	after	1185.51	which	3330.76	would	2913.48
first	1210.08	still	923.923	years	899.53	never	906.253
these	1223.29	being	850.746	there	3232.06	think	800.297
called	416.152	course	604.387	always	651.267	social	408.299
though	639.428	during	390.391	enough	500.87	mother	446.196
little	972.588	around	514.136	people	1317.05	should	956.941
looked	531.865	really	484.271	rather	471.718	almost	481.177
things	535.018	seemed	437.272	before	987.282	better	435.784
looking	292.942	without	595.76	another	690.83	certain	279.437
against	600.996	society	270.752	however	435.546	whether	308.41
between	742.529	perhaps	444.292	because	1166.95	thought	707.012
nothing	520.74	already	373.02	country	335.123	british	353.327
himself	529.664	morning	305.257	through	956.346	general	300.081
happened	187.997	american	248.977	although	304.841	students	178.002
probably	263.731	thinking	173.124	business	237.733	position	194.601
together	364.691	remember	238.447	children	647.817	possible	345.593
question	238.685	problems	221.194	economic	197.337	interest	211.437
national	240.351	suddenly	192.757	anything	380.575	movement	192.697
MEAN:	1821.7683	**MEAN:**	1406.065433	**MEAN:**	5065.474833	**MEAN:**	1488.292067

Frequency information for real word stimuli. Words were retrieved from MCWord Orthographic Wordform Database (http://www.neuro.mcw.edu/mcword/). According to the database, the frequency is a measure of how often the wordform occurred in 1,000,000 presentations in the CELEX database.

Table A.3: Pseudoword frequency statistics

Pseudo List 1		Pseudo List 2		Pseudo List 3		Pseudo List 4	
Word	Constrained Bigram Frequency	Word	Constrained Bigram Frequency	Word	Constrained Bigram Frequency	Word	Constrained Bigram Frequency
cru	26.27	hov	641.66	ope	3	lan	1792.94
ers	199.98	hal	4326.94	sed	1185.45	hel	2032.99
oth	0.09	med	730.3	ery	220.18	hin	4364.75
nup	38.46	ute	21.42	urt	83.74	nef	558.55
por	4284.4	ent	231.52	wim	1435.05	irt	94.5
mish	818.92	deen	5215.02	fass	1629.96	tast	3249.51
thub	9045.93	clim	99.06	fust	2993.5	pind	1746.77
flad	764.48	pess	894.12	cark	1815.14	dard	1288.16
feen	5399.92	wilm	4372.96	wilk	4509.2	felp	1568.29
trag	262.78	fack	2171.32	surn	897.63	thip	10553.53
bress	1356.15	yound	3047.7	grong	2267.25	keeds	695.47
stath	1763.17	sweel	538.04	doint	2120.59	prent	1132.06
parge	1075.33	lorks	1110.74	wouth	3990.52	terve	535.82
tound	3095.77	frugs	361.54	twisp	192.3	blace	1631.47
draff	593.23	thich	6451.07	greep	1211.39	dring	2763.72
calent	1050.81	plampy	320.15	balket	1074.27	prined	2646.82
empand	560.47	twelds	283.03	freety	536.92	strust	846.43
skults	328.28	moddle	1412.67	strope	813.88	hormal	872.82
befort	1546.42	rember	2182.66	peemed	2620.72	garked	2695.43
natual	939.25	insing	2129.97	degair	411.63	danted	3001.98
finands	672.77	harning	3659.92	cappors	934.04	subbery	920.38
himsent	968.93	plastly	792.64	musteme	631.52	gettled	2100.96
traffit	319.36	spanged	1770.06	bastily	1244.5	kithods	744.23
propind	2381.8	sibbing	3107.28	tobbing	3086.9	grovice	838.9
putsins	2074.72	ordular	237.39	groblex	788.46	pervice	888.55
contrage	686.77	crinking	1868.92	chamater	860.61	swincher	642.34
identant	435.02	natching	1896.4	marasand	455.46	dwending	1719.41
probling	2309.61	offector	298.51	amervise	372.2	inforway	380.47
retraver	933.99	pastrows	301.7	haparder	653.43	felegral	446.09
preaming	2153.14	interter	1186.8	feetched	1285.6	salcatic	263.92
MEAN:	1536.207333	**MEAN:**	1722.050333	**MEAN:**	1344.168	**MEAN:**	1767.242

Frequency information for pseudoword stimuli. Pseudowords were retrieved from MCWord Orthographic Wordform Database (http://www.neuro.mcw.edu/mcword/). According to the database, the constrained bigram frequency is a measure of how often the bigram wordform occurred in 1,000,000 presentations in the CELEX database.

APPENDIX B

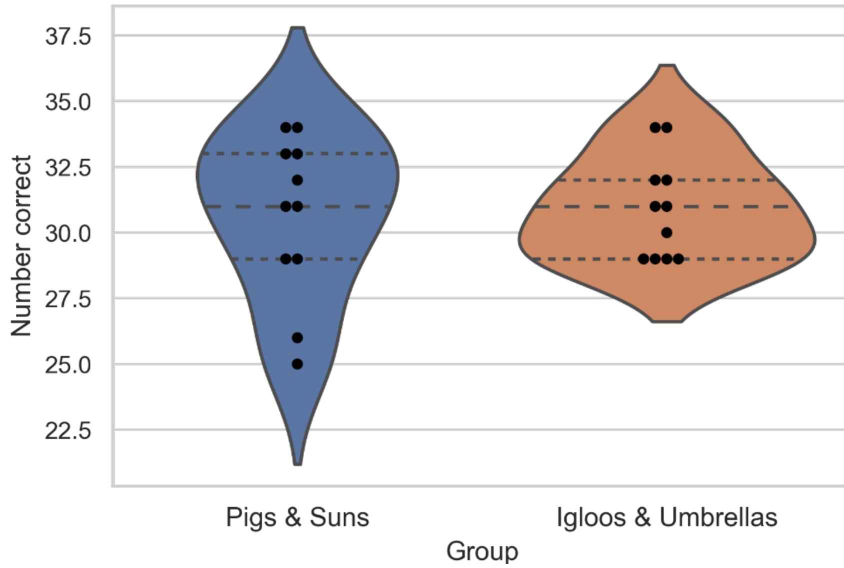
B.1 Supplementary material for Donnelly, et al. 2020 (submitted)

Table B.1: Side-by-side image cues for validation study.

Teacher Name	IPA	Pigs & Suns	Igloos & Umbrellas
Long a	/eɪ/	Grape	Alien
Long e	/i/	Key	Eagle
Long i	/aɪ/	Eye	Eye
Long o	/oʊ/	Bone	Ogre
Long u	/u/	Moon	Ooze
Short a	/æ/	Ant	Apple
Short e	/e/	Bed	Elephant
Short i	/ɪ/	Pig	Igloo
Short o	/ɑ/ & /ɔ/	Dog	Octopus
Schwa	/ə/	Sun	Umbrella
Short u	/ʊ/	Book	Book
R controlled A	/ɑr/	Heart	Arm
R controlled E	/ɝ/	Earth	Earth
R controlled O	/ɔr/	Corn	Orange
Consonant le	/əl/	Wolf	Wolf
Diphthong OI	/ɔɪ/	Coin	Oil
Diphthong OU	/aʊ/	House	Owl

Symbols for the 'Pigs and Suns' and 'Igloos and Umbrellas' phonemic image cue sets as used in the small-scale validation study in the final study (4.2.2 App Design).

Figure B.1: Between-subjects comparison of image cue sets.



Violin density plots show distributions of performance for the 'Pigs & Suns' (blue) and 'Igloos & Umbrellas' (orange) cue conditions. Points represent individual performance within each group, with dashed and dotted lines representing the mean and interquartile range, respectively. No significant difference in performance was found between the two sets (t(20) = -0.25, p = 0.806).

Table B.2: Fixed effects from mixed effects analysis for crossover study

Outcome Measure	Parameter	Estimate	SE	tStat	DF	pValue	95% CI
Real word decoding	Intercept	25.68	0.77	33.25	108	<0.001	(24.15 , 27.22)
	Condition	0.23	0.39	0.59	108	0.557	(-0.54 , 1)
Pseudo word decoding	Intercept	18.63	0.96	19.45	110	<0.001	(16.73 , 20.52)
	Condition	-0.11	0.63	-0.17	110	0.865	(-1.35 , 1.14)
Passage reading accuracy	Intercept	0.91	0.01	66.02	52	<0.001	(0.88 , 0.94)
	Condition	0.03	0.01	2.48	52	0.016	(0.01 , 0.05)
Passage reading rate	Intercept	1.25	0.14	8.76	52	<0.001	(0.97 , 1.54)
	Condition	-0.05	0.06	-0.91	52	0.367	(-0.16 , 0.06)

Table B.3: Real word frequency statistics

Words List 1		Words List 2		Words List 3		Words List 4		Words List 5		Words List 6	
Word	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**
has	2077.43	his	5828.33	had	6491.25	can	1954.34	any	1278.8	may	1040.35
you	6334.67	now	1683.88	she	4270.81	one	3302.38	did	1149.64	new	1082.35
was	11050.2	who	4754.07	the	61445.7	for	8319.47	get	1056.35	our	1258.63
not	5082.53	are	4254.92	all	3510.01	her	4032.24	how	1129.05	own	921.01
but	5240.18	him	2617.38	and	28470.7	out	2521.24	man	1024.11	see	1061.35
then	1820	were	3589.55	this	4504.73	only	1774.31	also	912.86	many	959.8
from	4229.52	time	1770.21	when	2560.63	into	2006.57	come	845.99	much	1122.69
they	4734.61	what	2581.57	that	11437.3	have	4529.54	even	1334.3	over	1365.06
been	2716.62	more	2427.66	will	2105.87	them	2305.28	just	1236.44	than	1662.82
said	2838.7	like	1884.02	with	7066.31	some	1859.62	made	1016.43	your	1579.71
again	766.8	other	1596.96	about	2322.72	right	758.53	going	709.45	since	513.01
could	1906.39	world	739.2	where	1034.58	those	866.22	great	639.73	small	531.63
their	2947.33	after	1185.51	which	3330.76	would	2913.48	house	563.93	three	629.61
first	1210.08	still	923.92	years	899.53	never	906.25	might	731.17	under	569.05
these	1223.29	being	850.75	there	3232.06	think	800.3	place	544.72	while	640.68
called	416.15	course	604.39	always	651.27	social	408.3	change	266.59	person	216.55
though	639.43	during	390.39	enough	500.87	mother	446.2	either	291.63	pounds	218.16
little	972.59	around	514.14	people	1317.05	should	956.94	father	343.99	reason	227.56
looked	531.87	really	484.27	rather	471.72	almost	481.18	moment	315.73	street	261.77
things	535.02	seemed	437.27	before	987.28	better	435.78	myself	268.97	turned	340.89
looking	292.94	without	595.76	another	690.83	certain	279.44	britain	225.6	members	206.97
against	601	society	270.75	however	435.55	whether	308.41	changed	233.21	outside	257.31
between	742.53	perhaps	444.29	because	1166.95	thought	707.01	example	221.25	present	243.5
nothing	520.74	already	373.02	country	335.12	british	353.33	friends	209.18	several	248.68
himself	529.66	morning	305.26	through	956.35	general	300.08	getting	229.46	started	200.19
happened	188	american	248.98	although	304.84	students	178	building	160.33	language	109.76
probably	263.73	thinking	173.12	business	237.73	position	194.6	evidence	155.63	realized	105.3
together	364.69	remember	238.45	children	647.82	possible	345.59	followed	137.43	returned	116.61

question	238.69	problems	221.19	economic	197.34	interest	211.44	hospital	107.15	straight	122.2
national	240.35	suddenly	192.76	anything	380.57	movement	192.7	instance	101.73	training	104.17
MEAN:	2041.858	**MEAN:**	1406.06567	**MEAN:**	5065.475	**MEAN:**	1488.29233	**MEAN:**	581.361667	**MEAN:**	597.245667

Frequency information for real word stimuli. Words were retrieved from MCWord Orthographic Wordform Database (http://www.neuro.mcw.edu/mcword/). According to the database, the frequency is a measure of how often the wordform occurred in 1,000,000 presentations in the CELEX database.

Table B.4: Pseudoword frequency statistics

Pseudo List 1		**Pseudo List 2**		**Pseudo List 3**		**Pseudo List 4**		**Pseudo List 5**		**Pseudo List 5**	
Word	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**
cru	26.27	hov	641.66	ope	3	lan	1792.94	bof	187.46	hed	2345.68
ers	199.98	hal	4326.94	sed	1185.45	hel	2032.99	tal	81.33	ist	1.01
oth	0.09	med	730.3	ery	220.18	hin	4364.75	ean	1700.1	eng	229.31
nup	38.46	ute	21.42	urt	83.74	nef	558.55	sha	2145.85	ast	132.25
por	4284.4	ent	231.52	wim	1435.05	irt	94.5	sil	328.34	plo	0.71
mish	818.92	deen	5215.02	fass	1629.96	tast	3249.51	fres	2216.98	verk	2896.25
thub	9045.93	clim	99.06	fust	2993.5	pind	1746.77	inth	3985.06	winy	4546.46
flad	764.48	pess	894.12	cark	1815.14	dard	1288.16	dery	2796.82	raly	1449.58
feen	5399.92	wilm	4372.96	wilk	4509.2	felp	1568.29	dong	1870.67	rell	3126.67
trag	262.78	fack	2171.32	surn	897.63	thip	10553.53	besh	1827.24	leen	5401.57
bress	1356.15	yound	3047.7	grong	2267.25	keeds	695.47	trind	2228.61	stath	1763.17
stath	1763.17	sweel	538.04	doint	2120.59	prent	1132.06	thest	5895.63	mives	2118.95
parge	1075.33	lorks	1110.74	wouth	3990.52	terve	535.82	laint	1921.62	litet	1206.39
tound	3095.77	frugs	361.54	twisp	192.3	blace	1631.47	exing	2238.33	siver	3007.94
draff	593.23	thich	6451.07	greep	1211.39	dring	2763.72	thess	5579.48	whate	2390.36
calent	1050.81	plampy	320.15	balket	1074.27	prined	2646.82	maling	2314.38	drince	899.45

empand	560.47	twelds	283.03	freety	536.92	strust	846.43	sareer	2196.18	jetter	2718.37
skults	328.28	moddle	1412.67	strope	813.88	hormal	872.82	moorer	2064.91	enlare	809.84
befort	1546.42	rember	2182.66	peemed	2620.72	garked	2695.43	satted	3409.39	borend	1333.01
natual	939.25	insing	2129.97	degair	411.63	danted	3001.98	lither	2920.75	shaned	2758.99
finands	672.77	harning	3659.92	cappors	934.04	subbery	920.38	fresing	3487.39	berving	3901.87
himsent	968.93	plastly	792.64	musteme	631.52	gettled	2100.96	banting	3730.87	lenerad	880.24
traffit	319.36	spanged	1770.06	bastily	1244.5	kithods	744.23	glinder	960.29	expoust	834.8
propind	2381.8	sibbing	3107.28	tobbing	3086.9	grovice	838.9	jerning	3584.11	eveling	3459.72
putsins	2074.72	ordular	237.39	groblex	788.46	pervice	888.55	croyest	827.55	knoited	1925.58
contrage	686.77	crinking	1868.92	chamater	860.61	swincher	642.34	restress	1049.93	hathting	2154.82
identant	435.02	natching	1896.4	marasand	455.46	dwending	1719.41	anasting	2084.82	oreation	1173.16
probling	2309.61	offector	298.51	amervise	372.2	inforway	380.47	plerated	1569.78	pathered	1750.03
retraver	933.99	pastrows	301.7	haparder	653.43	felegral	446.09	strinong	1232.75	atanking	1812.34
preaming	2153.14	interter	1186.8	feetched	1285.6	salcatic	263.92	reanding	2241.9	infender	874.62
MEAN:	1536.207333	**MEAN:**	1722.050333	**MEAN:**	1344.168	**MEAN:**	1767.242	**MEAN:**	2289.284	**MEAN:**	1930.104667

Frequency information for pseudoword stimuli. Pseudowords were retrieved from MCWord Orthographic Wordform Database (http://www.neuro.mcw.edu/mcword/). According to the database, the constrained bigram frequency is a measure of how often the bigram wordform occurred in 1,000,000 presentations in the CELEX database.

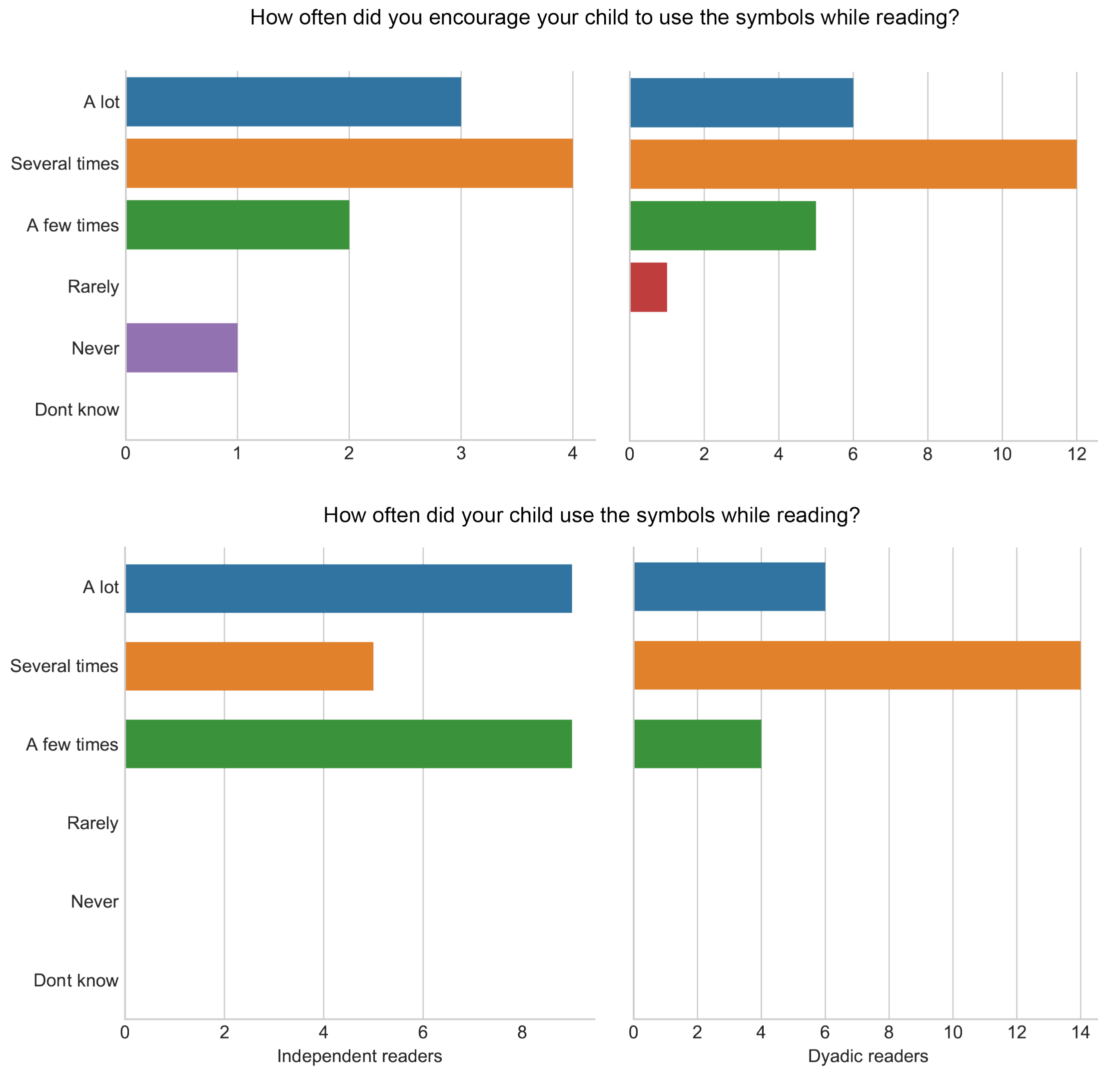
Table B.5: Fixed effects from mixed effects analysis for Comparison A

Outcome Measure	Parameter	Estimate	SE	tStat	DF	pValue	95% CI
Decoding Accuracy	Intercept	20.39	2.66	7.68	895	< 0.001	(15.18, 25.6)
	Session_2	-0.19	0.50	-0.39	895	0.698	(-1.17, 0.78)
	Session_3	0.25	0.61	0.41	895	0.680	(-0.95, 1.46)
	Group	-1.91	1.39	-1.38	895	0.169	(-4.64, 0.82)
	Session_2*group	2.58	0.60	4.26	895	< 0.001	(1.39, 3.76)
	Session_3*group	4.28	0.75	5.73	895	< 0.001	(2.82, 5.75)
Real word decoding	Intercept	24.36	1.33	18.33	445	< 0.001	(21.75, 26.97)
	Session_2	0.57	0.48	1.17	445	0.241	(-0.38, 1.51)
	Session_3	0.70	0.52	1.34	445	0.180	(-0.33, 1.74)
	Group	-0.86	1.43	-0.60	445	0.548	(-3.67, 1.95)
	Session_2*group	0.95	0.59	1.63	445	0.105	(-0.2, 2.11)
	Session_3*group	1.64	0.64	2.57	445	0.011	(0.38, 2.9)
Pseudo word decoding	Intercept	16.42	1.36	12.03	444	< 0.001	(13.74, 19.1)
	Session_2	-0.96	0.67	-1.43	444	0.153	(-2.28, 0.36)
	Session_3	-0.22	0.89	-0.24	444	0.809	(-1.97, 1.54)
	Group	-2.97	1.58	-1.88	444	0.061	(-6.07, 0.13)
	Session_2*group	4.21	0.82	5.15	444	< 0.001	(2.6, 5.81)
	Session_3*group	6.96	1.09	6.39	444	< 0.001	(4.82, 9.1)
Passage reading accuracy	Intercept	0.91	0.02	41.88	208	< 0.001	(0.87, 0.95)
	Session_2	0.00	0.01	-0.10	208	0.918	(-0.02, 0.01)
	Session_3	0.01	0.01	1.00	208	0.320	(-0.01, 0.03)
	Group	-0.02	0.03	-0.73	208	0.463	(-0.07, 0.03)
	Session_2*group	0.03	0.01	2.86	208	0.005	(0.01, 0.04)
	Session_3*group	0.03	0.01	2.51	208	0.013	(0.01, 0.05)
Passage reading rate	Intercept	1.39	0.11	12.95	176	< 0.001	(1.17, 1.6)
	Session_2	0.02	0.05	0.39	176	0.695	(-0.09, 0.13)
	Session_3	-0.01	0.06	-0.16	176	0.876	(-0.12, 0.1)
	Group	-0.09	0.13	-0.70	176	0.487	(-0.35, 0.17)
	Session_2*group	-0.10	0.07	-1.57	176	0.117	(-0.23, 0.03)
	Session_3*group	-0.07	0.07	-1.01	176	0.315	(-0.2, 0.07)

Table B.6: Fixed effects from mixed effects analysis for Comparison B

Outcome Measure	Parameter	Estimate	SE	tStat	DF	pValue	95% CI
Decoding Accuracy	Intercept	19.09	2.62	7.29	603	< 0.001	(13.94, 24.23)
	session_2	2.02	0.50	4.07	603	< 0.001	(1.04, 2.99)
	session_3	4.22	0.61	6.89	603	< 0.001	(3.02, 5.42)
	group	-1.25	1.72	-0.73	603	0.469	(-4.62, 2.13)
	session_2*group	0.74	0.71	1.05	603	0.296	(-0.65, 2.13)
	session_3*group	0.66	0.88	0.75	603	0.455	(-1.07, 2.38)
Real word decoding	Intercept	24.06	1.38	17.46	299	< 0.001	(21.35, 26.77)
	session_2	1.27	0.47	2.71	299	0.007	(0.35, 2.19)
	session_3	1.94	0.50	3.88	299	< 0.001	(0.96, 2.93)
	group	-1.14	1.79	-0.64	299	0.525	(-4.66, 2.38)
	session_2*group	0.51	0.67	0.76	299	0.445	(-0.8, 1.83)
	session_3*group	0.83	0.72	1.16	299	0.248	(-0.58, 2.24)
Pseudo word decoding	Intercept	14.12	1.41	10.01	298	< 0.001	(11.34, 16.89)
	session_2	2.77	0.69	3.99	298	< 0.001	(1.41, 4.13)
	session_3	6.50	0.94	6.95	298	< 0.001	(4.66, 8.34)
	group	-1.36	1.96	-0.69	298	0.489	(-5.21, 2.5)
	session_2*group	0.97	0.99	0.98	298	0.328	(-0.98, 2.92)
	session_3*group	0.49	1.34	0.37	298	0.713	(-2.15, 3.13)
Passage reading accuracy	Intercept	0.90	0.03	35.49	134	< 0.001	(0.85, 0.95)
	session_2	0.01	0.01	1.67	134	0.097	(0, 0.03)
	session_3	0.02	0.01	1.97	134	0.051	(0, 0.04)
	group	-0.01	0.04	-0.20	134	0.842	(-0.08, 0.06)
	session_2*group	0.03	0.01	2.41	134	0.017	(0, 0.05)
	session_3*group	0.04	0.02	2.37	134	0.019	(0.01, 0.07)
Passage reading rate	Intercept	1.32	0.10	13.54	116	< 0.001	(1.13, 1.51)
	session_2	-0.07	0.05	-1.34	116	0.183	(-0.16, 0.03)
	session_3	-0.09	0.06	-1.49	116	0.140	(-0.2, 0.03)
	group	-0.06	0.15	-0.40	116	0.689	(-0.35, 0.23)
	session_2*group	-0.03	0.08	-0.43	116	0.665	(-0.18, 0.12)
	session_3*group	0.05	0.09	0.55	116	0.580	(-0.12, 0.22)

Figure B.2: Parent-reported usage data by participant group.



Survey responses of usage data for participants in both the independent and dyadic reading groups.

VITA

Patrick joined the Brain Development and Education Lab at the Institute for Learning & Brain Sciences in July 2015. He received his B.A. in Child Development from Tufts University, focusing on dyslexia and the reading brain under the direction of Dr. Maryanne Wolf at the Center for Reading and Language Research. Patrick is particularly interested in the influence and impact of reading interventions/curricula, both print and digital, in helping struggling readers acquire fluency. He began his PhD studies in the Department of Speech & Hearing Sciences in September 2016.