AUSTRALIAN JOURNAL OF LEARNING DIFFICULTIES
2023, VOL. 28, NO. 1, 1–26
https://doi.org/10.1080/19404158.2023.2208135

OPEN ACCESS

Targeting phonological recoding to support orthographic learning: effectiveness of WordDriver delivered via telehealth

Antonette Seiler[a] and Suze Leitão[b]

[a]Curtin School of Allied Health, University Associate, Curtin University, Perth, Western Australia; [b]Curtin School of Allied Health, Curtin University, Perth, Western Australia

ABSTRACT

Fluent word recognition is an essential component of skilled reading, yet most children with reading difficulty have impaired word recognition. We developed and evaluated a web app, WordDriver, delivered via teletherapy, which targets phonological recoding to support orthographic learning and efficient word recognition. Participants were five children (aged 7–10 years) who, despite previous intervention using a systematic, synthetic phonics approach, demonstrated persistent word recognition impairment. Two studies, each using a single case experimental design examined changes in decoding accuracy (study 1) and orthographic learning (study 2) as measured by researcher-developed nonword reading and spelling lists, and standardised nonword reading assessments. Results suggested that all participants, irrespective of oral language and phonological processing profiles, made significant gains in decoding accuracy and orthographic learning for targeted vowel digraphs on researcher-developed nonword reading assessments, with clinically significant gains on standardised measures of decoding, and trends for generalisation in spelling.

ARTICLE HISTORY
Received 9 July 2022
Accepted 21 March 2023

Introduction

Accurate word recognition is essential to support reading comprehension – the ultimate aim of reading. The reciprocal relationship between word recognition and reading comprehension has been demonstrated by research supporting the simple view of reading (Gough & Tunmer, 1986), which states that reading comprehension is the product of accurate word recognition and listening comprehension (the ability to interpret the meanings of spoken words, sentences, and discourse – an oral language skill). Accurate context-free word recognition in the early years has been shown to predict oral language, reading comprehension, and general knowledge in older students (Sparks, Patton, & Murdoch, 2014), and to be a key predictor of reading comprehension across the life span (García & Cain, 2014). However, many children with reading comprehension difficulties have problems with accurate word recognition (Catts, Adlof, & Weismer, 2006).

CONTACT Suze Leitão S.Leitao@Curtin.edu.au Curtin School of Allied Health, Curtin University, Perth, Western Australia; Antonette Seiler seiler@elr.com.au Curtin School of Allied Health, University Associate, Curtin University, Perth, Western Australia

WordDriver was developed to target word recognition for children with persistent word recognition impairment. In the research reported here, we evaluated the efficacy of WordDriver and explored the use of teletherapy as a service delivery model. The WordDriver web-apps are based on research which has shown that (a) there are two main pathways to skilled word recognition (Coltheart, 2006); (b) accurate word recognition develops in phases (Ehri, 2005); (c) phonological recoding (sounding out and blending) results in efficient orthographic learning of novel words (Share, 1995), and (d) decoding tasks are a critical component within interventions for word recognition impairment (Pullen & Lane, 2014).

Theoretical basis of WordDriver

The dual-route model of skilled word reading (Coltheart, 2006) states that there are two pathways involved in context-free word recognition. The first is via the lexical route – automatic recognition of a known word. Using this process, skilled readers access a large bank of previously learned words that are automatically recognised, pronounced, and their meaning understood (due to well-established orthographic, phonological and semantic representations). The second pathway (the non-lexical route) is employed when an unknown word is encountered; in this case, grapheme-phoneme knowledge is used to sound out and blend to decipher the word (here referred to as *decoding)*. Research has shown that most children with word recognition impairment have problems with the non-lexical route – decoding (Herrmann, Matyas, & Pratt, 2006; Ouellette & Beers, 2010). While the dual-route model identifies key goals for early word reading instruction, it does not address the developmental progression of, and effective intervention approaches for, word recognition skills.

Ehri's well-supported phase theory (e.g. Hudson, Torgesen, Lane, & Turner, 2012) describes a progression through four phases in the acquisition of fluent word recognition. Initially, a few words are recognised within context (e.g. "EXIT" on an exit sign) - the *pre-alphabetic* phase. This is followed by emerging grapheme-phoneme knowledge, often with inaccurate decoding (the *partial alphabetic* phase); and eventually full mastery of most grapheme-phoneme correspondences supporting decoding of unfamiliar words (the *full alphabetic* phase). Finally, in the *consolidated* phase, knowledge of grapheme-phoneme connections expands to include larger units (e.g. rimes, syllables, morphemes, and whole words) allowing accurate decoding of multi-syllabic words, development of an increased bank of orthographic representations (stored mental images of the spellings of words), and wider knowledge of the spelling conventions of a language. This progression has been characterised as a connection-forming process, which employs phoneme awareness (knowledge of the sounds in words), decoding, and existing oral language to form connections that link written words to their pronunciations and meanings. While Ehri's theory describes the developmental phases in word reading and allows identification of the level of breakdown for a struggling reader, it does not inform teachers about effective strategies for instruction and intervention.

The Phonological Recoding theory (Share, 1995) describes a self-teaching mechanism that is an essential requirement for orthographic learning, resulting in the development of automatic word recognition. *Phonological recoding* (used interchangeably with the term *decoding*) occurs when the child attends to the internal structure of an unknown word by sounding out and blending to read the word.

Studies examining the phonological recoding theory have shown that, in typically developing children, the development of efficient word recognition of novel word forms is:

- Achieved after six presentations (Cunningham, Perry, Stanovich, & Share, 2002).
- Affected by dose rate: eight presentations were better than four (Bowey & Muller, 2005).
- Reduced when phonological recoding is prevented, for example, by being asked to say repetitive syllables ("la la la") as they read the words (Kyte & Johnson, 2006).
- More efficient when the words are presented in isolation (compared to reading words in context) and corrective feedback is provided (Martin-Chang, Ouellette, & Bond, 2017).

Other research has specifically examined the developmental progression and requirements for orthographic learning within opaque languages (Binamé & Poncelet, 2016). Two hundred and four typically developing children (second to sixth grade) who had been taught using a phonics method, were presented with an orthographic learning task, which was described as learning an extra-terrestrial language: participants learned novel words representing common objects – they decoded the word forms out loud, were provided with corrective feedback to ensure accurate decoding, and wrote the nonwords to dictation 10 times to optimise accurate formation of orthographic representations. Orthographic learning was assessed using a spelling test (participants were asked to write the target nonwords); and the relationships among the influencing factors (phonological recoding, short-term memory, orthographic sensitivity, and pre-existing orthographic knowledge) and the scores on the spelling test (assessing orthographic learning) were examined. The results showed that phonological recoding was highly significant in its independent contribution to orthographic learning in opaque languages, confirming "the undeniable involvement of phonological recoding in the formation and retention of novel orthographic representations" (Binamé & Poncelet, 2016, p. 23). Regarding the developmental trajectory, it was found that the second-grade children significantly underperformed compared with all other grades on both initial creation and long-term retention of novel word forms. It was proposed that this group was in a transition period in their mastery of the alphabetic principle due to less automation of grapheme–phoneme relationships. These results demonstrate that, within opaque languages, accurate decoding is an essential component for orthographic learning, and further, that a weakness in grapheme-phoneme knowledge results in reduced orthographic learning.

Children with dyslexia, the population targeted in our studies, have weaknesses in these two key skills: phonological recoding (the non-lexical route) and grapheme-phoneme knowledge for the extended code of an opaque language (e.g. vowel digraphs). Biname, Danzio, and Poncelet (2015) examined orthographic learning and long-term retention of novel word forms in an opaque language for dyslexic children (aged 9–13 years) compared to typically developing readers of the same age, and a group of younger children matched for reading age. Assessments of orthographic learning following the orthographic learning task described above (Binamé & Poncelet, 2016) revealed that children with dyslexia had significantly poorer decoding accuracy than both other groups; they required increased repetition to form orthographic representations of novel word forms; and long-term retention (one week post training) was significantly poorer than both other groups. The authors suggested (a) that the decoding weakness and the decreased grapheme-phoneme knowledge in this population places them at a significant disadvantage in initial orthographic learning and long-term retention of word forms, and (b) further research needs to examine the impact of "over training" in this population, that is, to determine if increased intensity improved long-term retention of orthographic representations.

Intervention studies

Studies investigating interventions for word reading impairment have also highlighted the important role of accurate decoding. In recent years, the Response to Intervention approach (Hempenstall, 2012) is commonly used as a framework to guide initial evidence-based reading instruction (Tier 1), provision of additional support for students at-risk of literacy delay (Tier 2), and use of diagnostic assessment and specific intervention for those who have not responded to previous intervention (Tier 3). There is well-substantiated evidence that effective Tier 1 early reading instruction should target phonemic awareness, phonics, fluency, and comprehension (Department of Education, 2005; National Reading Panel, 2000); and many studies have demonstrated that interventions that provide a greater focus on these areas (e.g. small group sessions) with particular emphasis on decoding (phonemic awareness combined with grapheme-phoneme knowledge) are effective for children requiring Tier 2 intervention (Berninger, Vermeulen, Abbott, & McCutchen, 2003; Bus & van Ijzendoorn, 1999; Gillon, 2002; Hatcher et al., 2006; Torgerson, Brooks, & Hall, 2006; Wheldall & Beaman, 1999).

While fewer studies have examined Tier 3 interventions, significant levels of non-response (Torgesen, 2001) or highly variable response (Denton et al., 2013) have been reported. Additionally, while multi-component interventions are unable to determine the essential element(s) within an intervention, studies that have attempted to isolate active ingredients (Austin, Vaughn, & McClelland, 2017; Lane, Pullen, Hudson, & Konold, 2009; Pullen & Lane, 2014) have suggested that activities targeting accurate decoding and phonics are a key component.

Our intervention is based on theoretical models demonstrating the role of phonological recoding (the non-lexical route) in the development of automatic word recognition (the lexical route), especially as it relates to children with severe reading difficulties learning to read in an opaque language. It aims to establish accuracy in the decoding process, and provide intensive intervention (many repetitions targeting identified impairments in grapheme-phoneme knowledge) to support orthographic learning of words and word parts. Our research examines the efficacy of a component, which has been shown to be an essential element within reading interventions for this cohort, and which may need to be delivered with a high dosage - "over training."

Design principles of the WordDriver apps

WordDriver is a web-app that targets accurate decoding to support orthographic learning – acquiring orthographic representations. It comprises two stages: WordDriver-1 establishes accuracy in the decoding process by presenting items with 1:1 grapheme-phoneme correspondences, while WordDriver-2 supports orthographic learning by targeting identified weaknesses in grapheme-phoneme knowledge – in this cohort, vowel digraphs.

A number of factors that have been shown to impact on orthographic learning are incorporated into the design of WordDriver:

- Both versions target items (randomly presented words and legal nonwords) that are presented without context because, while context has been shown to increase initial decoding accuracy, the ability to use phonological recoding to support orthographic learning is not affected by context (Martin-Chang, Ouellette, & Bond, 2017).
- The words and nonwords are matched and organised according to their orthotactic probability (the frequency with which a word's graphemes and bigraphs appear in English) as this linguistic feature has been found to influence orthographic learning – children at risk of literacy delay develop orthographic representations more efficiently with items of high orthotactic probability (Apel, Thomas-Tate, Wilson-Fowler, & Brimo, 2012).
- The child reads the items out loud and receives corrective feedback about decoding accuracy, as this has been shown to result in efficient orthographic learning (Martin-Chang, Ouellette, & Bond, 2017).
- The presentation of items is matched to the level of orthographic knowledge as orthographic learning is predicted by prior orthographic knowledge (Cunningham, Perry, Stanovich, & Share, 2002). In WordDriver-1, all items have one-to-one grapheme-phoneme correspondences because most young children with reading delay have not mastered accurate decoding of short consonant-vowel-consonant words (McCandliss, Beck, Sandak, & Perfetti, 2003). Once the child has mastered accuracy in the decoding process, WordDriver-2 presents items with consonant and vowel digraphs to promote orthographic learning.
- The intervention provides high levels of repetition of the target skill (phonological recoding), as repetition has been shown to optimise the development of orthographic representations (Biname, Danzio, & Poncelet, 2015; Bowey & Muller, 2005).

Initial investigations of WordDriver-1

The first stage in our programme of research (Seiler, Leitão, & Blosfelds, 2013, 2018) developed and examined the efficacy of WordDriver-1. Delivered in face-to-face sessions, its aim was to improve use of the decoding process (being able to accurately sound out and blend) for Year 2 children who continued to have word reading delays despite previous Tier 2 intervention.

A single-subject crossover research design with multiple treatments was used. Participants were eight Year 2 students (aged 7:6 to 8:11 years) with average cognitive skills and no other developmental issues. Despite previous reading interventions and having grapheme-phoneme knowledge for single consonants and short vowels in the average range, automatic word recognition and decoding skills were more than one standard deviation below the mean: they were unable to accurately decode 3-letter words that contained known grapheme-phoneme relationships.

Each participant received three, 15 minute sessions per week over two school terms – a total of 54 sessions. They were randomly assigned to one of two intervention sequences. Four participants received eight first-baseline sessions, 15 decoding sessions, eight second-baseline sessions, 15 language intervention sessions, and a final eight third-baseline sessions (i.e. baseline-decoding-baseline-language-baseline). The second four participants received the same content but with reversed order for the decoding and language interventions (i.e. baseline-language-baseline-decoding-baseline). Pre-intervention standardised assessments of oral language, cognitive skills, and phonological processing were used to gain insight into their language and cognitive profiles. To examine the efficacy and generalisation of the decoding intervention, researcher-developed nonword lists were administered during each baseline and intervention session, and standardised assessments of word and nonword reading efficiency, text reading accuracy, and reading comprehension were administered during the three baseline sessions.

The results suggested that, irrespective of language and cognitive profile, all participants made significant gains in nonword reading accuracy and efficiency with trends for gains on standardised measures of word-reading efficiency, text-reading accuracy, and reading comprehension. Further, for all participants, the significant gains in decoding only occurred following the decoding intervention – an intervention that involved approximately four hours (15 × 15 minute sessions). However, though use of WordDriver-1 resulted in mastery of the decoding process using items with known grapheme-phoneme relationships, examination of participant decoding responses showed that the relatively weak generalisation to other measures of word reading was likely due to the significant delays in participant grapheme-phoneme knowledge, particularly vowel digraphs that occur within opaque languages.

The current research

The current research aimed to (a) extend the results of our previous investigations by evaluating the efficacy of the WordDriver apps delivered through teletherapy: a service delivery model that has been shown to be effective, feasible, and acceptable to parents and teachers (Wales, Skinner, & Hayman, 2017) particularly in the rural locations our participants were drawn from, and (b) examine if the extended version, WordDriver-2, resulted in orthographic learning for vowel digraphs. The research questions were, for children with persistent word reading impairment in Years 2–4:

(1) Does WordDriver-1, delivered via teletherapy, increase nonword reading skills of items with known grapheme-phoneme correspondences measured by researcher-developed nonword reading lists?
(2) Does an extended version of this intervention (WordDriver-2), also delivered via teletherapy, increase nonword reading accuracy of items with unknown

grapheme-phoneme correspondences measured by researcher-developed nonword lists?

(3) Following WordDriver-1 and WordDriver-2, do any gains on the researcher-developed nonword reading lists generalise to a standardised test of nonword reading?

(4) Are gains in nonword reading reflected in nonword spelling measured by researcher-developed nonword spelling lists?

Materials and methods

Research design

There were two studies, each using a single case experimental research design with two phases. The first study, WordDriver-1, comprised a baseline and an intervention phase, while the second, WordDriver-2, included a baseline and intervention phase for two treated digraphs introduced in a staggered baseline, and one untreated digraph that served as control.

Participants

Following ethics approval from Curtin University (HRE2018–0556) and the Victorian Department of Education, two Department of Education and Training schools in East Gippsland who used a systematic synthetic phonics approach in the early years were recruited and agreed to take part. Systematic synthetic phonics is characterised by initial teaching of grapheme-phoneme correspondences in an incremental sequence (starting with single letter graphemes and then phonemes represented by digraphs), a focus on blending of phonemes for reading and segmenting of phonemes for spelling, and provision of decodable books (those that contain previously mastered grapheme-phoneme correspondences) to encourage use of phonemic strategies as a first approach to decipher unfamiliar words. An evaluation of this approach in a national rollout in the UK (Machin, McNally, & Viarengo, 2018) found strong effects for literacy acquisition in the early grades with long-term effects at age 11. Inclusion criteria were that participants:

- Had received previous evidence-based Tier 2 word reading intervention (i.e. systematic synthetic phonics);
- Were currently aged 7–10 years;
- Scored more than 1 standard deviation below the mean on the Phonemic Decoding Efficiency subtest of the Test of Word Reading Efficiency 2: TOWRE-2 (Torgesen, Wagner, & Rashotte, 2012);
- Had no developmental or sensory impairment, as screened using a parent questionnaire (Claessen, Leitão, & Barrett, 2010) and school records of previous assessments;
- Had hearing and vision in the normal range (school nurse screening);
- Had mastered grapheme-phoneme knowledge of short vowels and single consonants as evidenced by scores on the Grapheme subtest of the Phonological Awareness Test 2:PhAT2 (Robertson & Salter, 2007);

- Demonstrated speech sound production within average range as assessed by The Quick Screener (Bowen, 1996)

Additional measures were used to profile participant oral language and phonological processing skills:

- Receptive Vocabulary subtest of the WIAT-III (Wechsler, 2016)
- The Comprehensive Test of Phonological Processing Second Edition: CTOPP2 (Wagner, Torgesen, Rashotte, & Pearson, 2013)

Seventeen students were screened for inclusion using the TOWRE-2: seven Year 2 students in a larger school (school A) who had failed a school-based nonword reading assessment using DIBELS (2018), and all Year 2 (eight) and two Year 3–4 students in a small rural school (school B). Seven students matched the inclusion criteria. Of those, two were unable to complete the intervention programme: one due to frequent absences and the second student had behaviour difficulties that prevented consistent engagement with teletherapy. Table 1 provides participant performance on the selection measures for the five participants who took part.

Outcome measures

The primary measures of effectiveness were two sets of researcher-developed Assessment Nonword Lists (AxNW Lists). Similar to Seiler, Leitão, and Blosfelds (2018), one list was administered during each baseline session and every second intervention session. To address the first research question, the first set (WordDriver-1 AxNW Lists) measured change in decoding accuracy before and during WordDriver-1. There were 11 different lists, each of 35 nonwords with 1:1 grapheme – phoneme correspondence, starting with

Table 1. Standard scores of participant performance on pre-intervention assessments.

	School A			School B	
	P1	P2	P3	P4	P5
Age	7:10	8:5	7:11	10:1	10:7
Year level	2	2	2	3	4
TOWRE PDE	72	68	81	60	59
TOWRE SWE	62	73	62	55	55
PhAT G-P	91	89	96	99	110
Bowen Artic	WNL	WNL	WNL	WNL	WNL
WIAT Vocab	86	76	109	78	103
CTOPP-2					
Ph Aware	82	84	100	94	86
Ph Mem	82	61	85	55	70
RAN	82	76	92	79	79

Note: TOWRE = Test of Word Reading Efficiency; PDE = Phonemic Decoding Efficiency; SWE = Sight Word Efficiency; PhAT G-P = Decoding Subtests for short vowels and single consonants of the Phonological Awareness Test; WIAT Vocab = Wechsler Individual Achievement Test Receptive Vocabulary; CTOPP-2 = Comprehensive Test of Phonological Processing 2nd edition; Ph Aware = Phonological Awareness; Ph Mem = Phonological Memory; RAN = Rapid Automatic Naming; WNL = within normal limits. Standard scores between 86 and 115 are average; below 70 is in the severe range.

two and progressing to six-letter items. The outcome measure was WordDriver-1 NW Total – the number of nonwords accurately read out loud. To address the second research question, the second set (WordDriver-2 AxNW Lists) was developed for this study to measure decoding accuracy before and during WordDriver-2. There were 12 different lists. Each list comprised 30 randomly organised nonwords to examine decoding accuracy of three vowel digraphs: 10 for the digraph that was targeted first, 10 for the second target, and 10 for the untreated vowel digraph. The measure was WordDriver-2 NW Total – the number of nonwords accurately read out loud. Interobserver reliability for WordDriver-2 AxNW Lists, collected on a random selection of 20% of the responses and independently scored by a speech pathologist unfamiliar with the participants and whether the data were pre- or post-intervention, showed 97% agreement. WordDriver-1 was a replication of our previous research and interobserver reliability was not collected within this study due to resourcing limitations.

Two additional outcome measures were administered to assess generalisation. The first was a standardised assessment of nonword decoding – the decoding subtests of the Phonological Awareness Test-2: PhAT2 (Robertson & Salter, 2007) which has seven subtests assessing decoding of a range of word types, starting with items with 1:1 grapheme-phoneme correspondences and progressing to items with vowel digraphs. The PhAT2 was administered on three occasions: before and after WordDriver-1 by the researcher, and after WordDriver-2 by an independent speech pathologist who was unfamiliar with the participants and the goals of the intervention. The second measure, a researcher-developed nonword spelling assessment (WordDriver-2 AxNW Spelling Lists) was used to investigate the fourth research question. Four lists were constructed using the same method as the WordDriver-2 AxNW Lists. Each list comprised 30 nonwords –10 nonwords for each of the three vowel digraphs. Two were administered before and two after WordDriver-2. The measure was the number of digraphs correctly spelled out of 20, ignoring spelling errors of consonants.

Intervention

The WordDriver web-apps were used to deliver the intervention materials and the primary outcome measures (the researcher-developed WordDriver-1 and WordDriver-2 AxNW Lists). These web-apps use the analogy of learning to drive a car (Figure 1), in which the learner progresses through three stages (L-Plate – learner; P-Plate – practice; D-Plate – driver) in mastering accurate decoding of randomly presented words and nonwords. The words and legal nonwords were drawn from the second version of the MRC psycholinguistic database (Coltheart, 1981) and the ARC database (Rastle, Harrington, & Coltheart, 2002) respectively.

WordDriver-1 intervention targeted accurate phonological recoding of items with 1:1 grapheme-phoneme correspondence. Four levels of difficulty are presented (two-, three-, four-, and five-letter items). At each level, the L- and P-Plate items are organised in a predetermined sequence – initially, the first letter changes, then the last, then the middle, and then all letters. This draws attention to each letter and enables specific teaching (in the case of the L-Plate) and practice (P-Plate) of phonological recoding. The D-Plate items are organised according to orthotactic probability and are presented adaptively in response to participant error: easier items (higher orthotactic probability)

Figure 1. Screen shot of WordDriver-2. An example of D-Plate screen used to present WordDriver items.

following an incorrect response and more difficult (lower orthotactic probability) following a correct response. The researcher-developed WordDriver-1 AxNW Lists are presented with the T-Plate (test) using a similar graphical interface. Apart from two-letter items, there is no repetition of items between or within the decoding intervention and the T-Plates; thus minimising the possibility that any gains in decoding accuracy are due to practice effect.

Orthographic learning of vowel digraphs is targeted in WordDriver-2. The instructor selects the targets for the session using the Loader page (see Figure 2) - one or two vowel digraph/s with or without foil items (words and nonwords with short vowels and

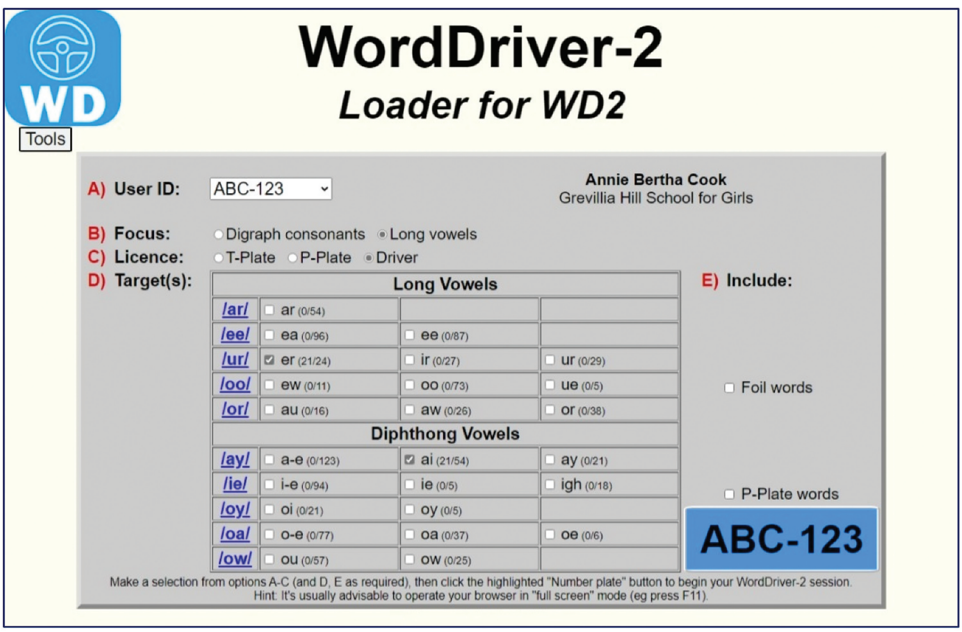

Figure 2. Screen shot of WordDriver-2 loader page. WordDriver-2 selector for target digraphs

consonant digraphs). Foil items are included to encourage the development of well-specified orthographic representations for vowel digraphs, that is, that the child accurately discriminates between easily confused short vowels and vowel digraphs (e.g. between "o" and "oa").

The items are randomly presented with increasing difficulty according to phoneme length and orthotactic probability. For example, when targeting the "ai" digraph, an example progression may be, "aim, thail, laid, waice, maint, stain, strait, spraice," while if foil words are included, the progression may include, "aim, whip, thail, sung" and so on. For each target digraph, the learner is presented with a P-Plate as a short practice, followed by a D-Plate (Driver).

Procedures

The sessions were delivered by the researcher three times per week over two school terms. Apart from an initial session at the school to familiarise participants with the researcher, and two sessions delivered face-to-face at the school due to technical difficulties, all WordDriver sessions were delivered via telehealth. During each session, the child was situated in front of a computer in a quiet room at school, accompanied by a support person or in close view of the teacher. The support person established and monitored the telehealth connection with the researcher, but was not actively involved in the intervention.

In Study 1, all participants completed four baseline and up to 15 intervention sessions. During each baseline session, a T-Plate (WordDriver-1 AxNW List) was administered first, following by one or more of the standardised outcome measures so that each session was completed within 20 minutes. On each T-Plate trial, the child touched the Go button, read out loud the nonword letter string and touched the Go button to view the next item. No feedback about accuracy of response was provided. The researcher stopped the child once six errors in eight consecutive items had occurred.

There were 15 Study 1 intervention sessions using WordDriver-1, with a T-Plate administered at the start of every second session. For WordDriver-1, all participants began at the level of three-letter strings as they had all made errors at this level on the pre-intervention assessments. The L-Plate was the starting point at all levels (three-, four-, five-letter strings), followed by the P-Plate, and finally the D-Plate. While the L-Plate was used to explicitly teach phonological recoding (i.e. the researcher performed all of the actions), during the P-Plate and D-Plate the child performed more of the actions. The child touched the Go button, and read out loud a randomly presented word or nonword. The researcher told the child whether it was a word or a nonword, and provided corrective feedback about the accuracy of response (by touching the Correct or Help button following a correct and incorrect response respectively). To support orthographic learning for words, the meaning of real words was highlighted and the child was encouraged to use the word in a sentence. In the case of nonwords, the researcher used a sentence explaining that the item "is not a real word; it has no meaning." The child then put the item in the Book or the Bin (for words or nonwords respectively) by touching either graphic, and touched the Go button to start the next trial. A criterion of 90% accuracy was required on the P-Plate to move to the D-Plate (within each level), and on the D-Plate to progress to the next level (e.g. from 3- to 4- letter items).

Study 2 comprised four baseline sessions and 15 intervention sessions. In the first baseline session, the PhAT2 was administered; and in this and each of the remaining baseline sessions, a T-Plate (WordDriver-2 AxNW List) and fifteen items from the WordDriver-2 AxNW Spelling List were delivered.

WordDriver-2 was used in each of the Study 2 intervention sessions, with a T-Plate delivered at the start of every second session. Pre-assessment testing revealed that while participants made errors on many vowel digraphs, the three digraphs common to all participants were "oa" (e.g. boat), "ai" (e.g. main), and "ou" (e.g. loud). For each participant, the two digraphs with the lowest scores were targeted for intervention and one digraph was untreated. During the first seven sessions, the first digraph was targeted: the P-Plate was used to introduce the digraph; one or more D-Plates consolidated accuracy of response; and then a D-Plate including foil words was used to ensure mastery. On the eighth session, the second digraph was introduced, and for the remaining sessions, both digraphs were targeted using a similar progression of P- D- and D-Plate with foil items.

Finally, the post-intervention outcome measures were delivered over two sessions. The researcher administered the WordDriver-2 AxNW Spelling lists and a speech pathologist, unfamiliar with the research and the participants, delivered the decoding subtests of the PhAT2. Table 2 summarises the timing and procedures in each phase.

Results

This research comprised two studies investigating the impact of an intervention called WordDriver, which targets accurate phonological recoding to support orthographic learning – Study 1 (WordDriver-1) and Study 2 (WordDriver-2). There were four research questions (RQs). The first question examined the impact of WordDriver-1 delivered via teletherapy on decoding accuracy measured by researcher-developed nonword lists (WordDriver-1 AxNW Lists); the second assessed whether WordDriver-2, also delivered via teletherapy, resulted in orthographic learning of targeted vowel digraphs measured by researcher-developed nonword lists (WordDriver-2 AxNW Lists); the third examined whether any gains in decoding and orthographic learning following WordDriver-1 and WordDriver-2 were reflected on standardised measures of nonword reading; and finally, the fourth research question used researcher-

Table 2. Timetable of assessments and intervention.

Study 1: Baseline 4 sessions	One WordDriver-1 AxNW List per session; PhAT2, CTOPP-2, and the Receptive Vocabulary subtest of the WIAT-3
Study 1: WordDriver-1 Up to 15 sessions	WordDriver-1 Decoding intervention One WordDriver-1 AxNW List on alternate sessions
Study 2: Baseline 4 sessions	One WordDriver-2 AxNW List and half a WordDriver-2 AxNW Spelling List per session; PhAT2
Study 2: WordDriver-2 15 sessions	WordDriver-2 Orthographic learning intervention One WordDriver-2 AxNW List on alternate sessions
Post-intervention Ax 2 sessions	WordDriver-2 AxNW Spelling Lists; PhAT2 by an independent speech pathologist

Note: PhAT2= Decoding subtests of the Phonological Awareness Test-2; CTOPP-2 = Comprehensive Test of Phonological Processing-2; WIAT-3 = the Receptive Vocabulary subtest of the Wechsler Individual Achievement Test-3.

developed nonword spelling lists (WordDriver-2 AxNW Spelling Lists) to determine if gains in orthographic learning of targeted vowel digraphs generalised to encoding of those digraphs.

Research question 1

The primary outcome measures addressing RQ1 were participant responses on the WordDriver-1 AxNW Lists during the four baseline sessions and every second intervention session in Study 1. Analyses of repeated measures within single subject design research involve examination of the relationship between dependent and independent variables to determine within-phase changes in level, trend, and variability; and between-phase changes, such as the immediacy of change, and overlap and consistency of scores (Kratochwill et al., 2013).

For this study, analyses of treatment effects involved visual inspection of the graphed responses (to examine individual participant responses within and between phases) and use of the Tau-U (which statistically evaluates within-phase trend and across-phase differences as well as overlap between phases). The Tau-U statistic provides the Tau coefficient (a non-parametric equivalent to a Pearson's correlation) which, as a rank order correlation, has minimal distributional assumptions and is relatively robust to autocorrelation. Tau varies between −1 and+1, and reflects the strength or consistency of the trend (the closer Tau is to 1, the stronger or more consistent is the trend). The sign of Tau reflects the direction of the trend (negative is decreasing and positive is increasing). A Tau value close to zero indicates minimal systematic trend in the time series. A statistically significant Tau (e.g. $p < .05$) confirms that the null hypothesis is rejected, that is, the trend is due to the intervention and not a chance occurrence. To calculate Tau-U, first the baseline sessions are analysed to determine if a trend occurred, and then the two phases (baseline and treatment) are compared statistically (while correcting for any trend in the baseline phase) to determine if there is a significant difference between baseline and intervention scores.

The graphed responses for each participant are shown in Figure 3.

Visual inspection of this graph suggests that while the final baseline scores for P3 and P5 increased prior to intervention, the baseline scores for all participants were low or inconsistent. Following commencement of intervention in session four, scores for all participants began to increase (apart from an unexpectedly low score for P1 in session eight), and remained above baseline values throughout intervention.

The results of the Tau-U analysis (Table 3) show that apart from P4 (who had a strong negative trend*), there was no significant trend during the baseline phase. Participants 2–5 showed a significant change following intervention with Tau all above 0.8, representing a moderate to strong effect. Participant 1 did not show a significant change following intervention (probably due to the unexpectedly low score), though a Tau of 0.7 shows a moderate effect.

These results suggest that WordDriver-1, when delivered via teletherapy, resulted in significant gains in use of phonological recoding to decode items with known letter-sound relationships – items with 1:1 letter-sound correspondence, as was reported in Seiler, Leitão, and Blosfelds (2018).

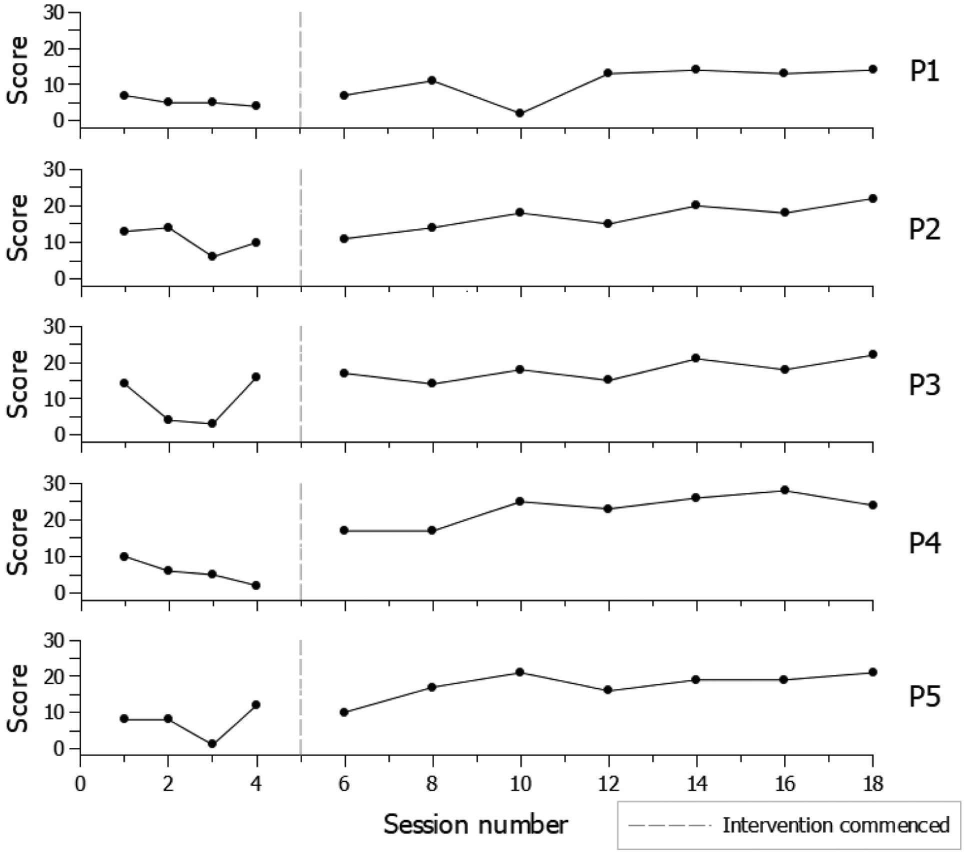

Figure 3. P1-P5 WordDriver-1 AxNW lists graphed assessment scores.

Table 3. P1-P5 Tau-U repeated measures WordDriver-1 AxNW lists.

Participant	Phase	s score	z score	*p* value	Tau	90%CI
P1	Baseline trend	−5	−1.70	0.09	−0.83	−1,-0.03
	Intervention	19	1.80	0.07	0.70	0.06,1
P2	Baseline trend	−2	−0.70	0.50	−0.34	−1,0.50
	Intervention	23	2.17	0.03	0.82	0.20,1
P3	Baseline trend	0	0	1	0	−0.81,0.81
	Intervention	23	2.17	0.03	0.82	0.20,1
P4	Baseline trend	−6	−2.04	0.04	−1*	−1,-0.19
	Intervention	28	2.17	0.00	1.2	0.59,1
P5	Baseline trend	1	0.34	0.73	0.17	−0.64,0.97
	Intervention	26	2.46	0.01	0.93	0.31,1

Research question 2

The impact of WordDriver-2 on orthographic learning of targeted vowel digraphs was measured by responses on researcher-developed nonword lists (WordDriver-2 AxNW Lists), administered during each baseline session and every second Study 2 intervention session. For each participant, the graphed responses and the results of the Tau-U analysis are presented.

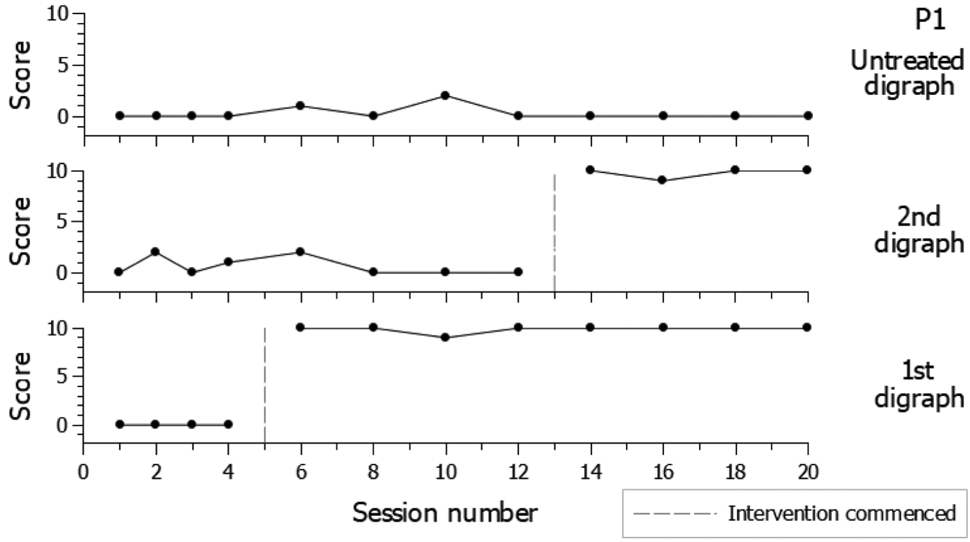

Figure 4. P1 graphed responses WordDriver-2 AxNW lists.

Participant 1

The graphed responses on WordDriver-2 AxNW Lists for P1 are shown in Figure 4.

Figure 4 shows that during the baseline phase, accuracy for all three digraphs was low. Response accuracy for the first targeted digraph increased following intervention at session four, with response accuracy for the other two non-targeted digraphs remaining low. Following intervention for the second digraph at session eight, response accuracy for both targeted digraphs was high while the untargeted digraph remained low.

The Tau-U analysis of the responses on the WordDriver-2 AxNW Lists for P1 is shown in Table 4.

These results show no trend in the baseline phase for any digraph, significant gains in the two treated digraphs, and no trend in the untreated digraph.

Participant 2

The graphed responses on WordDriver-2 AxNW Lists for P2 are shown in Figure 5.

This graph shows that during the baseline phase, accuracy for all three digraphs was low. Response accuracy for the first targeted digraph increased following intervention at session four, with response accuracy for the untreated digraph remaining low. Though there was an unexpected higher score for the second targeted digraph in session six, it

Table 4. P1 Tau-U repeated measures WordDriver-2 AxNW lists.

Phase	s score	z score	*p* value	Tau	90%CI
Target 1					
Baseline trend	0	0	1	0	−0.81,0.81
Intervention	32	2.71	0.007	1	0.39,1
Target 2					
Baseline trend	−5	−0.62	0.54	−0.18	−0.65,0.30
Intervention	32	2.71	0.006	1	0.40,1
Untreated					
Baseline	−0.02	−0.07	0.95	−0.02	−0.38,0.35

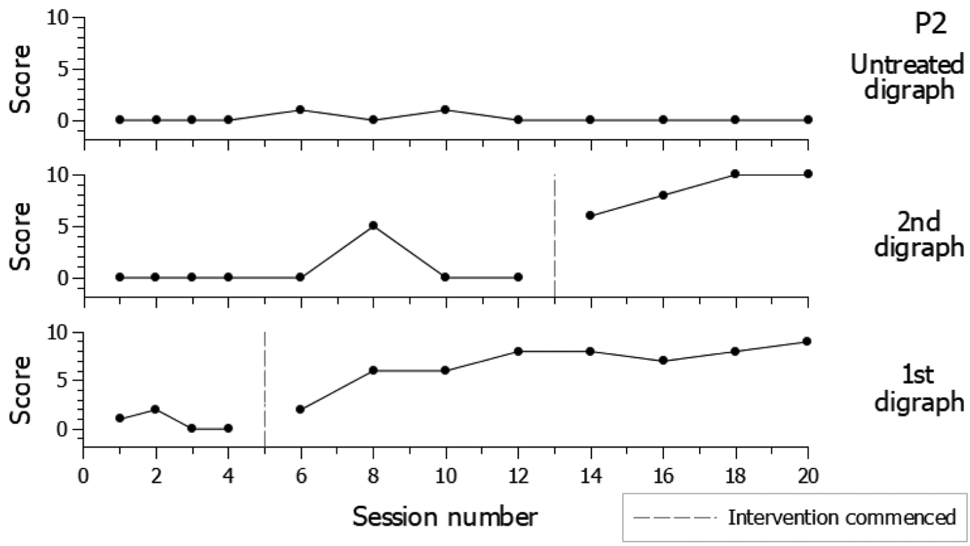

Figure 5. P2 Graphed responses WordDriver-2 AxNW lists.

wasn't until intervention for the second targeted digraph started in session eight that scores on this digraph remained consistently high. The untargeted digraph remained low throughout intervention.

The Tau-U analysis of the responses on the WordDriver-2 AxNW Lists for P2 is shown in Table 5.

These results show no trend in the baseline phase for any digraph, significant gains in the two treated digraphs, and no trend in the untreated digraph.

Participant 3

The graphed responses on WordDriver-2 AxNW Lists for P3 are shown in Figure 6.

This graph shows that during the baseline phase, accuracy for all three digraphs was low with higher scores for two digraphs (the second targeted and the untreated), in session four. Response accuracy for the first targeted digraph increased following intervention at session four, with response accuracy for second targeted and the untreated digraph remaining low. Following intervention for the second digraph at session eight, response accuracy for both targeted digraphs was high, while the untargeted digraph remained low.

Table 5. P2 Tau-U repeated measures WordDriver-2 AxNW lists.

Phase	s score	z score	*p* value	Tau	90%CI
Target 1					
Baseline trend	−3	−1.02	0.31	−0.5	−1,0.31
Intervention	31	2.63	0.009	.097	0.36,1
Target 2					
Baseline trend	3	0.37	0.71	0.11	−0.37,0.58
Intervention	32	2072	0.007	1	0.40,1
Untreated					
Baseline	−2	−0.14	0.89	−0.03	−0.39,0.33

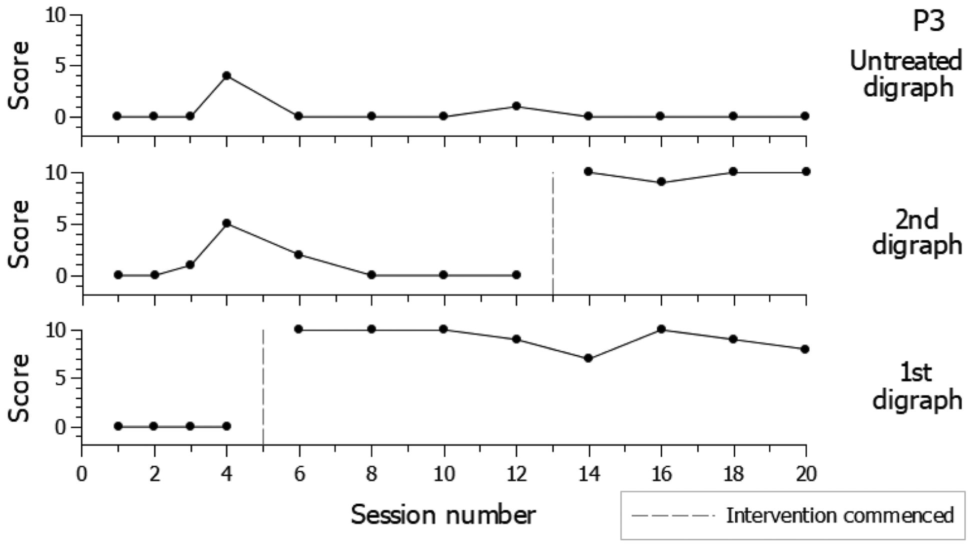

Figure 6. P3 graphed responses WordDriver-2 AxNW lists.

Table 6. P3 Tau-U repeated measures WordDriver-2 AxNW lists.

Phase	s score	z score	*p* value	Tau	90%CI
Target 1					
Baseline trend	0	0	1	0	−0.81,0.81
Intervention	32	2.72	0.007	1	0.40,1
Target 2					
Baseline trend	−2	−0.24	0.80	−0.07	−0.55,0.40
Intervention	32	2.72	0.007	1	0.40,1
Untreated					
Baseline	−3	−0.21	0.84	−0.06	−0.41,0.32

The Tau-U analysis of the responses on the WordDriver-2 AxNW Lists for P3 is shown in Table 6.

These results show no trend in the baseline phase for any digraph, significant gains in the two treated digraphs, and no trend in the untreated digraph.

Participant 4

The graphed responses on WordDriver-2 AxNW Lists for P4 are shown in Figure 7.

This graph shows that during the baseline phase, accuracy for the two treated digraphs was low, with higher scores for the untreated digraph. Response accuracy for the first targeted digraph increased following intervention at session four, with response accuracy for the second targeted remaining low, and the accuracy for the untreated digraph decreasing. Following intervention for the second digraph at session eight, response accuracy for the two targeted as well as the untreated digraph was high.

The Tau-U analysis of the responses on the WordDriver-2 AxNW Lists for P4 is shown in Table 7.

These results show that for the two treated digraphs, there was no trend in the baseline phase and significant gains following intervention. There was no trend for the untreated digraph with Tau value close to zero and $p < .05$.

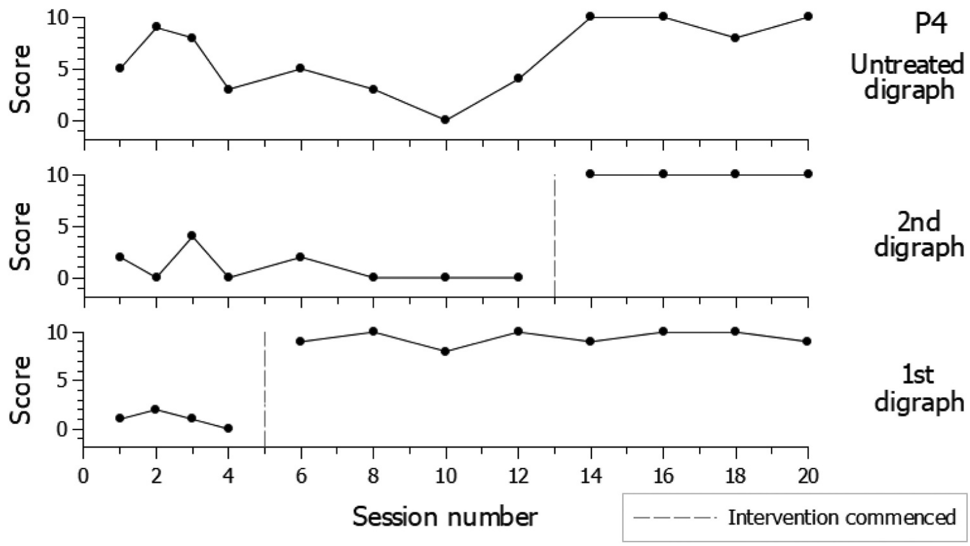

Figure 7. P4 Graphed responses WordDriver-2 AxNW lists.

Table 7. P4 Tau-U repeated measures WordDriver-2 AxNW lists.

Phase	s score	z score	*p* value	Tau	90%CI
Target 1					
Baseline trend	−3	−1.02	0.31	−0.5	−1,0.31
Intervention	32	2.72	0.007	1	0.40,1
Target 2					
Baseline trend	−9	−1.11	0.27	−0.32	−0.80,0.15
Intervention	32	2.72	0.007	1	0.40,1
Untreated					
Baseline	14	0.40	0.34	0.21	−0.15,0.58

Participant 5

The graphed responses on WordDriver-2 AxNW Lists for P5 are shown in Figure 8.

This graph shows that during the baseline phase, accuracy for the two treated digraphs was low, with higher but inconsistent scores for the untreated digraph. Response accuracy for the first targeted digraph increased following intervention at session four, with response accuracy for the second targeted remaining low, and inconsistency for the untreated digraph continuing. Following intervention for the second digraph at session eight, response accuracy for the second targeted digraph increased, the first targeted digraph remained high, and the untreated digraphs remained inconsistent.

The Tau-U analysis of the responses on the WordDriver-2 AxNW Lists for P5 is shown in Table 8.

These results show that for the two treated digraphs, there was no trend in the baseline phase and significant gains following intervention. There was no trend for the untreated digraph.

Taken together, the results suggest that for all participants, following intervention, significant gains in accurate decoding on researcher-developed nonword lists were made on each of the two targeted vowel digraphs. Three participants (P1, P2, P3) showed low scores on the untreated vowel digraph during the baseline and intervention phases, and

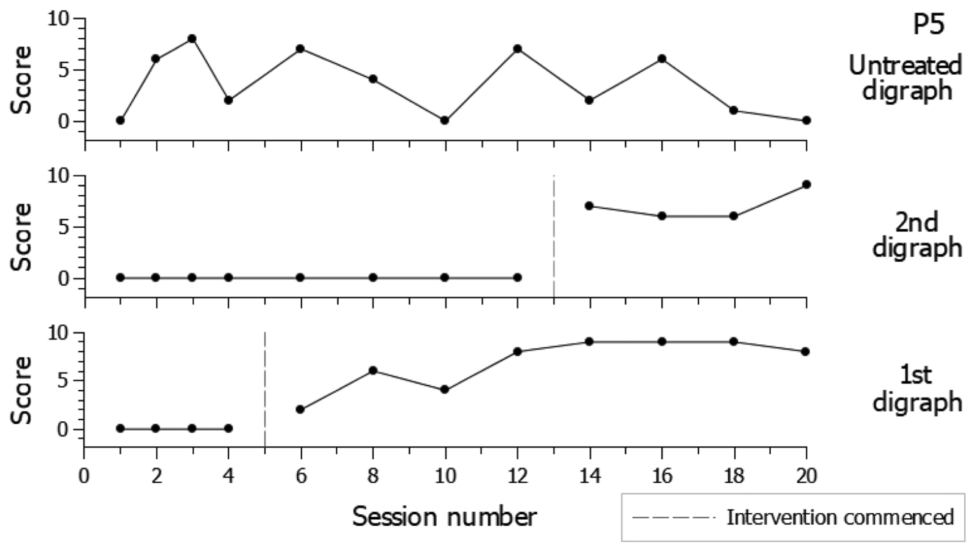

Figure 8. P5 Graphed responses WordDriver-2 AxNW lists.

Table 8. P5 Tau-U repeated measures WordDriver-2 AxNW lists.

Phase	s score	z score	*p* value	Tau	90%CI
Target 1					
Baseline trend	0	0	1	0	−0.81,0.81
Intervention	32	2.72	0.007	1	0.40.1
Target 2					
Baseline trend	0	0	0.34	0	−0.48,0.48
Intervention	32	2.72	0.007	1	0.40,1
Untreated					
Baseline	−14	−0.40	0.34	−0.21	0.58,0.15

two participants (P4, P5) showed inconsistent responses on the untreated vowel digraph through baseline and intervention phases.

Research question 3

The decoding subtests of the Phonological Awareness Test-2 (PhAT2) were used to examine if gains in decoding accuracy were reflected in a standardised measure of nonword reading. Table 9 shows the standard scores for all participants on the seven subtests prior to intervention (Study 1 Baseline), following WordDriver-1 (Study 2 Baseline), and following WordDriver-2 (post intervention). An asterisk is used to highlight clinically significant changes in standard scores between any two consecutive data collection points. A clinically significant change is considered to occur when the standard score crosses a clinical boundary as defined in the PhAT2 test manual (e.g. from below average to average). The subtests that are expected to detect changes following WordDriver-1 (items with 1:1 letter-sound correspondence) are Consonant-Vowel-Consonant and Consonant Blends, while those expected to detect changes following

Table 9. Standard scores on PhAT2 subtests for all participants.

Subtest	Time	P1	P2	P3	P4	P5
CVC	Study 1 B	97	70	86	<66	<66
	Study 2 B	108	105*	101	95*	102*
	Post Ix	101	105	95	102	95
C Bl	Study 1 B	90	94	81	84	65
	Study 2 B	102	90	104*	91*	97*
	Post Ix	104	96	89	91	97
C Dig	Study 1 B	78	72	91	83	83
	Study 2 B	104*	108*	89	109*	92*
	Post Ix	112	98	89	109	75*
V Dig	Study 1 B	<78	<74	<78	<64	76
	Study 2 B	<78	<66	<74	<64	64
	Post Ix	97*	77*	79	88*	70
Diph	Study 1 B	82	78	87	62	68
	Study 2 B	87*	<74	<78	73	<62
	Post Ix	83*	84	83	68*	85*
CVCe	Study 1 B	<80	75	<80	<64	70
	Study 2 B	<80	<71	<75	<64	<64
	Post Ix	80	71	<75	<64	<64
RV	Study 1 B	<81	<75	<81	<63	63
	Study 2 B	<81	<69	<75	92*	69
	Post Ix	84	69	<75	63*	75*

Note: B = Baseline: Ix = intervention; CVC = Consonant-vowel-consonant; C Dig = Consonant digraph; C Bl = Consonant blends; V Dig = Vowel digraphs; RV – R vowels; CVCe = "e" rule; Diph = Diphthong vowels; * = clinically significant change by crossing a clinical boundary between two consecutive data collection points [standard scores between 86–115 = normal range; 71–85 = below average; <70 = severely below average]

WordDriver-2 are Vowel Digraphs (which included "oa" and "ai") and Diphthongs (which included "ou").

These results show that the expected pattern occurred following WordDriver-1: three participants (P2, P4, P5) made clinically significant gains in Consonant-Vowel-Consonants (P1 and P3 were in the normal range at Study 1 baseline); and three (P3, P4, P5) made clinically significant gains in Consonant Blends (P1 and P2 were in the normal range at Study 1 baseline). Additionally, four participants (P1, P2, P4, P5) made clinically significant gains in Consonant Digraphs (P3 was in the normal range at Study 1 baseline) suggesting generalisation to items with consonant digraphs, and scores on Consonant-Vowel-Consonant and Consonant Blends remained in the normal range post-intervention for all participants. Among all participants, there were minimal changes in the subtests measuring vowel digraphs following WordDriver-1, apart from P1 who made a slight gain in the Diphthong subtest.

Following WordDriver-2 intervention, the expected trend was observed for most participants on the two subtests that assessed targeted vowel digraphs: on the Vowel Digraphs subtest, three participants (P1, P2, P4) made clinically significant gains; and on Diphthongs, P5 made a clinically significant gain, and P2 and P3 showed strong positive trend. However, P1 and P4 showed a clinically significant negative change on Diphthongs. On the subtests assessing nontargeted digraphs (R-Vowels and Consonant-Vowel-Consonant+e), minimal clinically significant changes were observed: P5 made a clinically significant gain and P4 showed a negative change in R-Vowels.

Table 10. Scores on WordDriver-2 AxNW spelling lists.

	Raw score out of 20					
	Pre WordDriver-2 Intervention			Post WordDriver-2 Intervention		
	Target 1	Target 2	Untreated	Target 1	Target 2	Untreated
P1	3	2	0	12	11	8
P2	1	0	0	8	7	0
P3	0	4	1	15	17	0
P4	1	1	13	2	4	11
P5	13	1	14	9	0	12

Research question 4

The final research question examined whether gains in orthographic learning of vowel digraphs were reflected in encoding skills measured by researcher-developed nonword spelling lists (WordDriver-2 AxNW Spelling Lists). Table 10 shows the raw scores of the targeted and untreated vowel digraphs before and after WordDriver-2.

These results suggest that the expected pattern occurred for three participants following WordDriver-2 intervention: P1's scores increased on both targeted and untreated digraphs, while P2 and P3 scores increased on only the targeted digraphs. The spelling accuracy on targeted and untargeted digraphs for P4 and P5 remained unchanged.

Discussion

Our research programme aimed to examine the impact of an intervention, the WordDriver web apps, that specifically target accurate phonological recoding to support orthographic learning. WordDriver is based on evidence showing that most children with reading disorders have impaired word reading skills; that phonological recoding is a requirement for orthographic learning – a key skill to develop fluent word reading; and that while multicomponent interventions are necessary to support the two main areas required for skilled reading (oral language and accurate word reading), tasks that target accurate decoding and orthographic learning have been shown to be an essential intervention element. Additionally, other researchers have concluded that further research is needed to determine if increased intensity and "over training" in phonics and decoding is beneficial for children with severe and persistent word reading impairment.

As this population typically has severe difficulty with all aspects of decoding, this preliminary research involved two studies. The first study was a replication of our previous research. It examined whether WordDriver-1, shown to be effective in face-to-face delivery, would increase nonword reading accuracy of items with known grapheme-phoneme correspondences when delivered via teletherapy (RQ1). Study 2 explored whether use of WordDriver-2 targeting specific grapheme-phoneme correspondences would result in orthographic learning as measured by researcher-developed nonword lists (RQ2), standardised assessment of nonword reading (RQ3), and nonword spelling using researcher-developed nonword spelling lists (RQ4).

The results of Study 1 suggest that most participants made significant gains in decoding for items with 1:1 grapheme-phoneme correspondences on researcher-developed nonword lists, and all achieved scores in the normal range (which were maintained at post-intervention testing) on standardised tests of nonword reading. Consistent with our previous study in which WordDriver-1 was delivered face-to-face, these results showed that when delivered via teletherapy, significant gains in accurate phonological recoding occurred following WordDriver-1 for students who, despite having mastered grapheme-phoneme knowledge for short vowels and single consonants, and (in this study) had received Tier 2 intervention using systematic synthetic phonics, remained below average in using this knowledge when decoding. Further, this targeted approach was shown to be efficient: the gains occurred following a total of four hours of intervention (15 × 15 minute sessions), and use of teletherapy decreased travel demands and interruption during the school day.

Study 2 aimed to determine if, once participants had mastered accurate use of phonological recoding, use of WordDriver-2 would result in orthographic learning of vowel digraphs. The results suggest that, when decoding accuracy was measured by researcher-developed nonword lists (RQ2), all participants made statistically significant gains following intervention on the two targeted vowel digraphs with no gains observed on the untreated digraph. Further, the observed gains in decoding accuracy for treated digraphs were generalised to a standardised measure of nonword decoding (RQ3): most participants made clinically significant gains on subtests that contained targeted vowel digraphs, with no gains observed on subtests that contained non-targeted vowel digraphs.

These results suggest, first, that children with severe word reading impairment (requiring Tier 3 intervention) may benefit from an intervention that provides many repetitions to consolidate grapheme-phoneme knowledge for unknown vowel digraphs. Though the participants had previously received Tier 2 intervention and continued to participate in classroom-based intervention using systematic synthetic phonics throughout the research study, significant gains in orthographic learning of targeted vowel digraphs occurred following WordDriver-2. Furthermore, it was observed that response accuracy on the untreated vowel digraph remained unchanged: the three younger participants continued to score close to zero and the two older participants' scores remained inconsistent. This observation suggests that children requiring Tier 3 decoding intervention may have difficulty with generalisation and may need decoding interventions that provide many repetitions of targets that closely match their orthographic learning needs. Second, these gains occurred irrespective of pre-intervention oral language and phonological processing profile: two participants had below average receptive vocabulary, and all had significant weaknesses in one or more areas of phonological processing. The observed improvements suggest that this targeted approach may be a useful component within reading intervention for children with a range of risk factors for reading impairment.

While gains in decoding accuracy for targeted vowel digraphs occurred for all participants, responses on the encoding assessment using researcher-developed nonword spelling lists (RQ4) suggest mixed results: only three participants made gains on raw scores for targeted digraphs. Two observations may be drawn from these results. First, it is likely that this population would benefit from interventions that include both decoding and encoding to optimise orthographic learning, as was found in the Biname, Danzio, and Poncelet (2015) study. The second observation is that all of the younger students made gains on the encoding tasks, but neither of the older students. While further research is required, this observation suggests that early intervention increases intervention efficiency and that the intervention process may take longer if commencement of targeted decoding intervention is delayed.

Turning to the theoretical considerations, the results are consistent with Ehri's phase model and the phonological recoding theory. These participants were stuck at the second stage of word reading development – the partial alphabetic stage: though they had mastered grapheme-phoneme knowledge of short vowels and single consonants, their decoding attempts were inaccurate at the consonant-vowel-consonant level, and all had severe delays in grapheme-phoneme knowledge for most of the vowel spelling patterns. Following WordDriver-1, mastery of the decoding process was achieved, which provided them with the necessary skills (accurate decoding) to allow progression to the third phase – the full alphabetic phase. According to Share's phonological recoding theory, accurate decoding is a self-teaching process supporting further orthographic learning. Within this short intervention targeting two vowel digraphs, the repetition and corrective feedback provided during WordDriver-2 was successful in supporting orthographic learning for targeted vowel digraphs. However, as has been demonstrated in previous studies (Biname, Danzio, & Poncelet, 2015), children with severe reading delays require increased intensity with more repetitions of key skills, combined with other intervention components, such as encoding, to ensure orthographic learning is consolidated.

There are some limitations to this research. First, these preliminary studies were completed by the authors of the web app, on a small number of participants, using a single case experimental design. Future research replicating the efficacy of this intervention by other research groups, across larger numbers using group designs, would result in more robust findings and provide information about generalisability to other environments, for example, when delivered by teachers and parents. Second, future single-case design investigations could include flexibility in the number of baseline sessions as it was noted that in Study 1, the final baseline points for two of the participants demonstrated a rise just prior to treatment. Third, the outcome measure assessed decoding at the single word level. Future research could assess the impact of improved decoding on text reading fluency and reading comprehension. Fourth, this programme of research investigated the efficacy of one component that previous research has suggested was an essential element. Future research that included other elements, such as a spelling task, may result in an equally efficient intervention that produced more substantial gains in orthographic learning. And finally, though telehealth has the potential to provide intervention to a larger number of students, technical issues and the possibility that some students may not cope with telehealth are limitations.

In conclusion, the results of this preliminary research suggest that this intervention may be a useful component within reading interventions for learners with severe and persistent word reading impairment. Consistent with the evidence regarding word reading development, it provides an efficient method to ensure first., that accuracy in the decoding process has been established, and second, that the learner is provided with many decoding repetitions specifically targeted to their orthographic learning needs. Further, the telehealth delivery process may facilitate broad access to this intervention component, which previous research suggests is a key element within interventions for reading impairment. And finally, access to the WordDriver web apps is available freely for clinicians, teachers, and researchers at <worddriver.com> and <languageandliteracyinyoungpeople.com>, and the Assessment Nonword Lists used in this research are at <worddriver.com/docs/WordDriverAxNWLists.pdf>.

Acknowledgements

The authors thank the parents and children who participated in the study, and the school staff who were so enthusiastic about involvement in research that aims to help children with reading difficulty. They extend special acknowledgement to Dr Rob Seiler for his computer programming expertise.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Apel, K., Thomas-Tate, S., Wilson-Fowler, E. B., & Brimo, D. (2012). Acquisition of initial mental graphemic representations by children at risk for literacy development. *Applied Psycholinguistics*, *33*(2), 365–391. doi:10.1017/s0142716411000403

Austin, C. R., Vaughn, S., & McClelland, A. M. (2017). Intensive Reading interventions for inadequate responders in grades K–3: A synthesis. *Learning Disability Quarterly*, *40*(4), 191–210. doi:10.1177/0731948717714446

Berninger, V. W., Vermeulen, K., Abbott, R. D., & McCutchen, D. (2003). Comparison of three approaches to supplementary reading instruction for low-achieving second-grade readers. *Language, Speech & Hearing Services in Schools*, *34*(2), 101–116. doi:10.1044/0161-1461(2003/009)

Biname, F., Danzio, S., & Poncelet, M. (2015). Relative ease in creating detailed orthographic representations contrasted with severe difficulties to maintain them in long-term memory among dyslexic children. *Dyslexia*, *24*(4), 361–370. doi:10.1002/dys.1506

Biname, F., & Poncelet, M. (2016). The development of the abilities to acquire novel detailed orthographic representations and maintain them in long-term memory. *Journal of Experimental Child Psychology*, *143*, 14–33. doi:10.1016/j.jecp.2015.10.010

Bowen, C. (1996). The quick screener.

Bowey, J. A., & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading: A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *92*(3), 203–219. doi:10.1016/j.jecp.2005.06.005

Bus, A. G., & van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, *91*(3), 403–414. doi:10.1037/0022-0663.91.3.403

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language & Hearing Research*, *49*(2), 278–293. doi:10.1044/1092-4388(2006/023)

Claessen, M., Leitao, S., & Barrett, N. (2010). Investigating children's ability to reflect on stored phonological representations: The silent deletion of phonemes task. *International Journal of Language & Communication Disorders*, *45*(4), 411–423. doi:10.3109/13682820903111945

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology*, *33A*(4), 497–505. doi:10.1080/14640748108400805

Coltheart, M. (2006). Dual route and connectionist models of reading: An overview. *London Review of Education*, *4*(1), 5–17. doi:10.1080/13603110600574322

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology*, *82*(3), 185–199. doi:10.1016/S0022-0965(02)00008-5

Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013). Effects of tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *Journal of Educational Psychology*, *105*(3), 633–648. doi:10.1037/a0032581

Department of Education, S. A. T. (2005). *Teaching reading: Report and recommendations*. Melbourne: Australian Government.

DIBELS. (2018). DIBELS®: Dynamic indicators of basic early literacy skills. Retrieved from https://dibels.uoregon.edu/materials/dibels

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167–188. doi:10.1207/s1532799xssr0902_4

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*(1), 74–111. doi:10.3102/0034654313499616

Gillon, G. T. (2002). Follow-up study investigating the benefits of phonological awareness intervention for children with spoken language impairment. *International Journal of Language & Communication Disorders*, *37*(4), 381–400. doi:10.1080/1368282021000007776

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. doi:10.1177/074193258600700104

Hatcher, P. J., Hulme, C., Miles, J. N., Carroll, J. M., Hatcher, J., Gibbs, S., Snowling, M. J. & Snowling, M. J. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: A randomised controlled trial. *Journal of Child Psychology and Psychiatry*, *47*(8), 820–827. doi:10.1111/j.1469-7610.2005.01559.x

Hempenstall, K. (2012). Response to intervention: Accountability in action. *Australian Journal of Learning Difficulties*, *17*(2), 101–131. doi:10.1080/19404158.2012.704879

Herrmann, J. A., Matyas, T., & Pratt, C. (2006). Meta-analysis of the nonword reading deficit in specific reading disorder. *Dyslexia*, *12*(3), 195–221. doi:10.1002/dys.324

Hudson, R. F., Torgesen, J. K., Lane, H. B., & Turner, S. J. (2012). Relations among reading skills and sub-skills and text-level reading proficiency in developing readers. *Reading and Writing*, *25*(2), 483–507. doi:10.1007/s11145-010-9283-6

Kratochwill, T. R. P., Hitchcock, J. H. P., Horner, R. H. P., Levin, J. R. P., Odom, S. L. P., Rindskopf, D. M. P., & Shadish, W. R. P. (2013). Single-Case Intervention Research Design Standards: RASE. *Remedial and Special Education*, *34*(1), 26. doi:10.1177/0741932512452794

Kyte, C. S., & Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology*, *93*(2), 166–185. doi:10.1016/j.jecp.2005.09.003

Lane, H. B., Pullen, P. C., Hudson, R. F., & Konold, T. R. (2009). Identifying essential instructional components of literacy tutoring for struggling beginning readers. *Literacy Research and Instruction*, *48*(4), 277–297. doi:10.1080/19388070902875173

Machin, S., McNally, S., & Viarengo, M. (2018). Changing How literacy is taught: evidence on synthetic phonics. *American Economic Journal: Economic Policy*, *10*(2), 217–241. doi:10.1257/pol.20160514

Martin-Chang, S., Ouellette, G., & Bond, L. (2017). Differential effects of context and feedback on orthographic learning: How good is good enough? *Scientific Studies of Reading*, *21*(1), 17–30. doi:10.1080/10888438.2016.1263993

McCandliss, B., Beck, I. L., Sandak, R., & Perfetti, C. (2003). Focusing attention on decoding for children with poor reading skills: Design and preliminary tests of the word building intervention. *Scientific Studies of Reading*, *7*(1), 75–104. doi:10.1207/S1532799XSSR0701_05

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*. Retrieved from Washington, D.C: https://www.nichd.nih.gov/publications/pubs/nrp/smallbook

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, *23*(2), 189–208. doi:10.1007/s11145-008-9159-1

Pullen, P. C., & Lane, H. B. (2014). Teacher-directed decoding practice with manipulative letters and word reading skill development of struggling first grade students. *Exceptionality*, *22*(1), 1–16. doi:10.1080/09362835.2014.865952

Rastle, K., Harrington, J., & Coltheart, M. (2002). Nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology*, *358*(4), 534. doi:10.1080/02724980244000099

Robertson, C., & Salter, W. (2007). *The phonological awareness test 2*. East Moline, IL: LinguiSystems, Inc.

Seiler, A., Leitão, S., & Blosfelds, M. (2013). The effectiveness of a computer-supported intervention targeting orthographic processing and phonological recoding for children with impaired word identification. *Journal of Clinical Practice in Speech-Language Pathology*, *15*(1), 13–18.

Seiler, A., Leitão, S., & Blosfelds, M. (2018). WordDriver-1: Evaluating the efficacy of an app-supported decoding intervention for children with reading impairment. *International Journal of Language & Communication Disorders*, *54*(2), 155–305. doi:10.1111/1460-6984.12388

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218. doi:10.1016/0010-0277(94)00645-2

Sparks, R., Patton, J., & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: Replication of '10 years later'. *Reading and Writing*, *27*(1), 189–211. doi:10.1007/s11145-013-9439-2

Torgerson, C. J., Brooks, G., & Hall, J. (2006). A systematic review of the research literature on the use of phonics in the teaching of reading and spelling (*report no. 711*). Nottingham: DfES Publications.

Torgesen, J. K. (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. doi:10.1177/002221940103400104

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of word reading efficiency* (*Second Edition* ed.). Austin, TS: Pro-Ed.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive test of phonological processing* (*Second Edition* ed.). Austin, Texas: Pro-Ed.

Wales, D., Skinner, L., & Hayman, M. (2017). The efficacy of telehealth-delivered speech and language intervention for primary school-age children: A Systematic Review. *International Journal of Telerehabilitation*, *9*(1), 55–70. doi:10.5195/ijt.2017.6219

Wechsler, D. (2016). *WIAT-III A&NZ*. Sydney NSW: Pearson Clinical and Talent Assessment.

Wheldall, K., & Beaman, R. (1999). *An evaluation of MULTILIT - Making up lost time in literacy. executive summary*. Retrieved from Sydney, Australia: Macquarie University.