Australian Journal of Teacher Education

Volume 45 | Issue 12 Article 1

2020

The Efficacy of a Computer Program for Increasing Phonemic Awareness and Decoding Skills in a Primary School Setting for Children with Reading Difficulties

Tiffany Winn
Flinders University

Julia Miller
The University of Adelaide

Willem van Steenbrugge
Flinders University, South Australia

Recommended Citation

Winn, T., Miller, J., & van Steenbrugge, W. (2020). The Efficacy of a Computer Program for Increasing Phonemic Awareness and Decoding Skills in a Primary School Setting for Children with Reading Difficulties. *Australian Journal of Teacher Education, 45*(12).
http://dx.doi.org/10.14221/ajte.202v45n12.1

The Efficacy Of A Computer Program For Increasing Phonemic Awareness And Decoding Skills In A Primary School Setting For Children With Reading Difficulties

Tiffany Winn
Flinders University
Julia Miller
The University of Adelaide
Willem van Steenbrugge
Flinders University

Abstract: This paper addresses a gap in research regarding the efficacy of software programs to help children with reading difficulties. Forty-two children aged 5-13 years identified as poor readers participated in a study over twelve weeks using ReadingDoctor, a software program targeting phonemic awareness, orthographic-phonemic mappings, decoding ability and sight word recognition. Measures were taken using the Sutherland Phonological Awareness Test - Revised (SPAT-R), the Test of Word Reading Efficiency (TOWRE), and the graphemes and decoding subtests of the Phonological Awareness Test 2 (PAT-2). A quasi-experimental one group study with three multiple baseline measures was used. The dependent variables/measures were assessed seven times over a period of 32 weeks, allowing the research to be completed in the school-allocated timeframe. Significant improvements were found on all three measures of phonological/phonemic awareness and word-reading efficiency. These improvements were maintained when assessed three months later, during which time the software program was not used.

Introduction

Research suggests that phonological processing skills are critical for reading acquisition (Carson et al., 2014; McArthur & Castles, 2013; Suggate, 2010), and that explicit teaching of phonics is especially critical for children with a variety of conditions that make it harder to learn to read (Savage & Carless, 2008; Shaywitz et al., 2008; Vadasy, 2010). One prevalent condition, dyslexia, is characterised by “persistent literacy learning difficulties, especially difficulties in word recognition, spelling, and phonological recoding, where phonological recoding is the ability to convert letters and letter patterns into phonological forms” (Tunmer & Greaney, 2010, p. 231). Dyslexia affects at least 10% of the worldwide population (Vellutino et al., 2004). Other children, while not formally diagnosed with dyslexia, may also show poor reading skills. Research suggests that all poor readers, whether dyslexic or not, have significant difficulties with phonological processing (Foorman & Torgesen, 2001; Kuppen & Goswami, 2016; Tanaka et al., 2011), so approaches that target phonological awareness skills—particularly phonemic awareness (Ehri et al., 2001)—are relevant to all poor readers, regardless of the reason they are struggling.

While proficient readers are skilled at using letter/sound relationships to decode unfamiliar words and to spell correctly (Bosse et al., 2013), children with poor reading skills, including those with dyslexia, often find decoding difficult and may therefore struggle to read unfamiliar words. Phonological awareness has been found to be an important prerequisite for reading acquisition, and consists of a "broad awareness of the sound structures of speech, including awareness of syllables, phonemes and rhyme" (Chera & Wood, 2003, p. 37). Phonemic awareness is a subset of phonological awareness, and has been shown to be especially important for reading acquisition (Ehri et al., 2001). One key phonemic awareness skill is blending, which is the combining of phonemes to form words. Another key skill is segmentation, which refers to separating words into distinct sounds. Research in the area of neuroplasticity suggests that functional connectivity can be strengthened in the brain with appropriate phonics-based intervention (Richards & Berninger, 2008). Phonological awareness, in particular phonemic awareness, is therefore an important component of many programs designed to help poor readers improve their reading ability (Ehri et al., 2001).

Computer-assisted instruction (CAI) to promote reading skills and phonological awareness has progressed from basic programs for reading text aloud (Chera & Wood, 2003) to more sophisticated, interactive programs that incorporate a video gaming element (Kast et al., 2007). An early study by Chera and Wood (2003) demonstrated the efficacy of computer software in improving the phonological awareness of children aged between 3 and 6 years. The study involved "animated multimedia talking books" based on a Longman phonics-based reading scheme, which presented oral versions of a text along with activities targeting "visual and auditory letter sound awareness" (Chera & Wood, 2003, p. 41). However, while the results showed an increase in phonological awareness, there was no increase in word recognition. Moreover, since the children had not been identified as 'at risk' with respect to their reading skills, it is not clear whether the findings would also have applied to children experiencing reading difficulty. Similarly, Wild (2009) investigated the phonological awareness skills of 127 children aged 5 to 6 years who were starting to read. The children were randomly assigned to one of three groups: an experimental group of 44 children who used a computer program for *Rhyme and Analogy* during one session per week over a period of 10 – 12 weeks; a control group of 43 children who did a paper-based version of the computer program; and an additional control group of 40 children who underwent a computer-based practical mathematics program. Phonological awareness improved in all three groups, but the children in the experimental group who used the computer program specifically targeting phonological skills showed the greatest improvement, evidenced by a modest statistical improvement when compared to the two control groups. Smaller improvements were found in a study of kindergarten students (Macaruso & Walker, 2008), where modest gains in reading performance and phonological awareness were observed in both CAI and non-CAI groups. The students receiving supplementary CAI showed slightly greater gains. However, significant group differences, especially regarding improved phonological awareness, were only shown between the low performing students of the two groups, i.e. students with lower initial reading performance. In a similar study of first-grade students, significant group differences were again only found when analysis was restricted to struggling students (Macaruso et al., 2006).

The overall finding of the above studies, namely a small effect of (supplementary) CAI on reading ability with a larger effect seen in poor or struggling readers, was confirmed in a more recent review of twenty studies (Cheung & Slavin, 2013). Cheung and Slavin also noted a variation across different types of computer programs, with greater effect sizes for small-group interactive programs, modest effect sizes for supplementary software programs and small effect sizes for comprehensive software programs. Most of the research reviewed by Cheung and Slavin either focuses on younger children or shows either non-significant treatment effects or an effect size close to zero (less than 0.1). There were two exceptions. Bass et al. (1986) evaluated the use of a variety of software programs for students in grades 5 to 6 (expected ages 10 to 12 years) and found a benefit, albeit with a small effect size (+0.22 and +0.13 respectively on the two measures used). The other exception was a study on READ 180 (Dynarski et al., 2007), a comprehensive literacy intervention program which combines both computer and non-computer intervention. The study evaluated vocabulary and reading comprehension with children in grades 4 to 6 (expected ages 9 to 12 years) and obtained effect sizes of +0.23 and +0.31 respectively. Generally, however, there is a gap in research on the use of educational software with older children, aged 9 and above, and the limited research that exists does not evaluate phonological skills and does not demonstrate a benefit for the use of educational software with older children.

Despite the lack of studies independently evaluating particular tools, the literature offers advice on what constitutes well-designed educational software from a theoretical point of view. A review of the literature indicates three key points to consider: mode of use; the pedagogical basis of the software design; and practical principles for designing and working with technology.

Warschauer (2007) argued that the way a tool is used is perhaps even more important than the tool itself, and that a software program is most effective when used in conjunction with face-to-face teaching. Savage et al. (2013) further emphasise the role of the teacher in implementing a software program effectively, making two key points. First, they claim that outcomes from reading intervention software programs are significantly impacted by the way teachers manage the use of the software. Secondly, they highlight that key issues impacting on the way teachers implement such software include both teacher competence with technology and the availability of 'just-in-time' training.

A second key factor influencing the effectiveness of educational software is its pedagogical basis. Bishop and Santoro (2005) provide a framework for evaluating beginning reading software aimed at at-risk readers. From a pedagogical point of view, the two teaching areas which Bishop and Santoro argue are most critical for the software to include are phonological awareness and alphabetic understanding (phonemic awareness). Reflecting more broadly from an instructional design perspective on pedagogical issues in beginning reading software, Bishop and Santoro highlight the importance of a number of software features. For example, they argue that software needs to be "systematic" (p. 61), facilitating progression through increasingly more difficult tasks. Software should also be "instructionally supportive" (p. 62), providing enough information for a student to be able to work out how to use the program effectively. In a broader context, Wood (2008, p. 91) points out that "a software tool may be successful because it forces a certain style of learning that is effective, not because it is software". The focus when designing effective software for teaching has largely been on developing models, frameworks and principles to steer the implementation of relevant pedagogy into online learning (Clark, 2002; Conole et al., 2004). The most successful computer programs may be those that use both visual and auditory prompts in a multisensory approach (Kast et al., 2007), as such approaches have been shown to enhance learning (Kast et al., 2011).

Technological aspects of educational software represent a third key factor impacting its effectiveness. Focusing on beginning reading software for at-risk readers, Bishop and Santoro (2005) identify three key features of interface design: aesthetics, relating to the look and feel of the interface; operational support, relating to the support available for navigating the software that is available within the software itself; and interactivity, meaning the capacity of the software to engage the user and provide them with a sense of driving the program through their interaction with the software. In a broader context, Winn and Clark (2011) note that perhaps the most significant principles for the users and designers of software are Clark and Mayer's (2003) principles for multimedia learning. These principles are based on research into how the human brain learns most effectively and are formulated as transparent, practical principles that users and designers of online tools can make use of in their work. Among the most well-recognised of Clark and Mayer's principles are the Multimedia Principle, which suggests using words and graphics together rather than just words to encourage learning; the Contiguity Principle, which suggests aligning words with corresponding graphics; and the Modality Principle, which suggests using audio rather than on-screen text to describe graphics.

One of the issues in evaluating existing CAI programs that focus on phonological awareness is that there is, overall, a lack of both analysis of the pedagogical and technological design of the programs, and of peer-reviewed studies evaluating their effectiveness. For example, one of the most popular programs focusing on phonological awareness in Australia is called ABC Reading Eggs. However, the authors cannot find any study examining the effectiveness of this program. Another well-known approach is Sounds-Write, which uses a linguistic phonics approach and includes some computer-based elements. While the developers of Sounds-Write have evaluated its efficacy in a large study of students using the program over six years (Case et al., 2009), the authors cannot find any relevant peer-reviewed study.

One exception to this lack of peer-reviewed programs is ABRACADABRA (ABRA for short). This is a software program for improving literacy that was developed in Canada and has evaluations of its effectiveness reported in several peer-reviewed journal articles (e.g. Abrami et al., 2014; Savage et al., 2013; Wolgemuth et al., 2013). One large (n = 1067) cluster randomized control trial (RCT) of kindergarten, grade 1 and grade 2 students in Canada (Savage et al. 2013) found that the use of ABRA had significant positive effects on phonological blending and letter sound knowledge, and increased the speed at which students were able to do phoneme segmentation when compared with wholly manual instruction. However, no detail is provided about the type of instruction provided to students in the control group, so it is difficult to ascertain whether the improvements in the experimental group were a result of the use of ABRA per se or arose because of the type of instruction provided. For example, students in the experimental group were taught using a phonics-based approach whereas students in the control group may not have been taught phonics. A study using ABRA with Indigenous and non-Indigenous students in northern Australia (Wolgemuth et al., 2013) also found positive effects on phonological awareness ($d = .37$) and phoneme-grapheme knowledge ($d = .37$). This study did provide some information about the type of instruction provided to students in the control group, and control group instruction did include a phonics component, although whether the amount of time given to phonics instruction in both control and experimental groups is comparable is not clear.

From a teacher education point of view, it is important that software is 'on the radar' of teachers as being a useful tool for children struggling with reading. There are many children who struggle with reading but do not qualify for much, if any, School Support Officer (SSO) time. Yet the evidence is clear that the gap between good and poor readers tends to persist or increase over time without quality instruction (Ferrer et al., 2015; Stanovich, 1986). High quality, evidence-based software may therefore help fill this gap. However, in 2020 there were over 500,000 education apps available in the Apple store alone (Educational App Store, 2021). The variety can be overwhelming, and it is not known how many of these apps have been independently tested for efficacy. Such independent evaluation is vital so that teachers know which tools to trust and plays a key role in clarifying some of the obfuscation that emerges from marketing or lobbying by tech companies.

A further issue giving impetus to this study is the relatively high percentage of children struggling with reading. For example, the Progress in International Reading Literacy Study (PIRLS) (Thomson et al., 2017) demonstrates that 19% of Australian 10-year-old children do not meet the 'intermediate' benchmark in reading ability. In South Australia, the state where this study was conducted, that number is approximately 25%, and the most recent results of the phonics screening check (Government of South Australia, 2018) showed that 57% of students did not meet the expected achievement level (i.e. 28 words or more out of 40). In advantaged communities, 45% of students did not meet the expected achievement level; in schools serving low SES communities, the figure was a massive 82%. Overall, 566 students (4%) did not answer any of the words correctly in the test. This indicates that there are many children who need high quality, evidence-based intervention to support them to reach key reading benchmarks.

One of the challenges in providing high quality, evidence-based intervention in a classroom context is the differing knowledge and skill bases of individual students. There is some evidence, however, that differentiated instruction can increase learning (e.g. Haelermans et al., 2014), and software that provides an individually tailored, scaffolded approach that allows learning content to be differentiated for each individual student has the greatest potential. In addition, if children need additional reading support outside the classroom, a program is needed that allows an opportunity for the teacher to guide parents in how to support learning from home, using the software.

The current study arose from both a growing interest in the effectiveness of appropriate remedial computer-based intervention for students with a range of reading difficulties, and an interest in testing a particular multisensory tool (ReadingDoctor[1]) that focuses on developing those areas of phonics shown to be most critical for reading acquisition. The design of ReadingDoctor (RD) draws together many insights from current research in the area of reading acquisition and dyslexia, providing a software program that aims to increase phonemic awareness and grapheme to phoneme decoding, and to strengthen visual word recognition skills using a multisensory (auditory-visual) approach. ReadingDoctor was developed in Australia and is widely used in an Australian context. The main aim of the study was to investigate whether consistent use of computer-assisted instruction (in this case RD software used for twenty minutes, five days per week, over twelve weeks) significantly improved phonological awareness, especially phonemic awareness, and (word) reading skills in primary school children who had been identified by their teachers as having difficulty reading. Analysis of the results occurred in two age groups (5-8 years and 9-13 years), in the expectation that the younger children would show lower performance/ability in reading or phonological processing than the older children, and in order to separately identify the effect of the software on younger and older children. This particular software had not previously been tested in this way.

[1] This article includes a word which is or is asserted to be a proprietary term or trade mark. Its inclusion does not imply it has acquired for legal purposes a non-proprietary or general significance, nor is any other judgement implied concerning its legal status.

ReadingDoctor

ReadingDoctor (RD) comprises a number of games or exercises to improve phonological and phonemic awareness and visual word recognition, with an emphasis on phonological-orthographic processing. The activities include games designed to strengthen letter sound knowledge, blending, segmentation and sight word recognition.

ReadingDoctor's multisensory approach is designed to help students to make meaningful associations between letters and the phonemes and sounds they typically represent (see Figure 1).

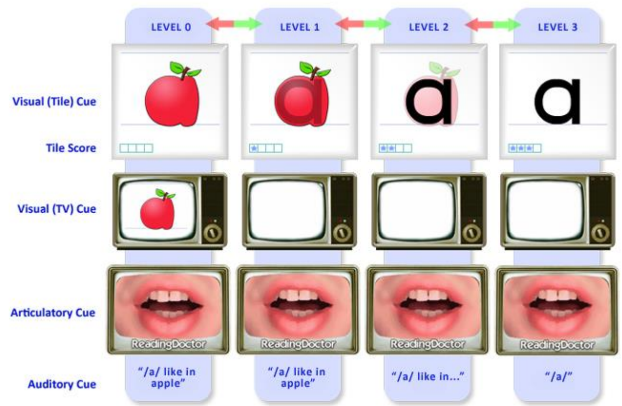

Figure 1: ReadingDoctor scaffolding provided by visual, articulatory and auditory cues, prompting the student to associate visual cues, written symbols and sounds.
At Level 0, cues are most explicit, becoming gradually less explicit as the student progresses to Level 3.

The program uses visual, auditory and articulatory mnemonics to establish letter/phoneme associations and strengthen decoding skills. In Figure 1, the vertical column below each level header (LEVEL 0, LEVEL 1, LEVEL 2, LEVEL 3) illustrates the cues that appear at each level. The levels are progressive, so that, for example, the more times a student clicks correctly on the apple picture, in response to an auditory cue, the more progressively the apple picture fades and the letter *a* comes to the fore. Both visual and auditory cues become less explicit with increasing levels, until they are not presented at all, although they reappear after an incorrect response.

Figure 2 shows the game for letter/sound correspondence. Pictures that closely resemble letter shapes, such as the ladder and tap, are superimposed with the letter representing the picture's initial sound. The umbrellas indicate that a particular letter/sound correspondence has been consecutively correctly identified enough times to be removed from the activity.

Figure 2: Screen from ReadingDoctor showing letter sounds using 25 symbols.
Symbols that have been correctly recognised up to level 3 have been covered by umbrella pictures; other symbols are at various levels, with the particular level indicated by the number of stars in the lower left corner of the tile box.

Points are awarded for correct identification of letter/sound correspondence. The bottom right corner of the screen shows a visual prompt which is the mouth positioning of the production of the sound; the mouth moves appropriately each time the RD program says the sound.

As well as facilitating the learning of letter/sound relationships, RD targets phonemic awareness skills, including blending, which is a crucial skill underpinning decoding ability (Gambrell et al., 2007). A screenshot from the blending game is shown in Figure 3.

Figure 3: A screenshot from the blending game.
The blending game requires a word, in this case "frog", to be matched to a corresponding picture tile.

In the blending game, the student moves the mouse over the tiles forming a word, and visual, verbal and articulatory cues assist the student with the blending task, which is to match the word with the appropriate picture tile. As with all games in ReadingDoctor, visual and auditory cues fade as the level increases. Initially, the program sounds out the letters in the word and blends the word for the student. If the student provides correct responses, the support is gradually removed until the student is required to decode the word independently. The number of words to be blended can be varied from 2 to 15. The cueing level, the number of responses required to complete a target and the difficulty of vocabulary used can also be adjusted.

Another key reading skill is recognition of high-frequency 'sight words', many of which are irregularly spelled and therefore need to be learned using an element of visual memorisation rather than decoding skills alone. The sight words that are presented in RD were selected from the Medical Research Council (MRC) psycholinguistic database, based on frequency of occurrence in the language, with the addition of words that were not picked up by the MRC database but recurred in several other high frequency word lists. Preference was given to irregularly spelled sight words and regularly spelled sight words that contain graphemes that are typically not taught until later in the synthetic phonics sequence (e.g. *the*, *by*).

RD uses a multisensory approach to facilitate sight word recognition, as shown in Figure 4.

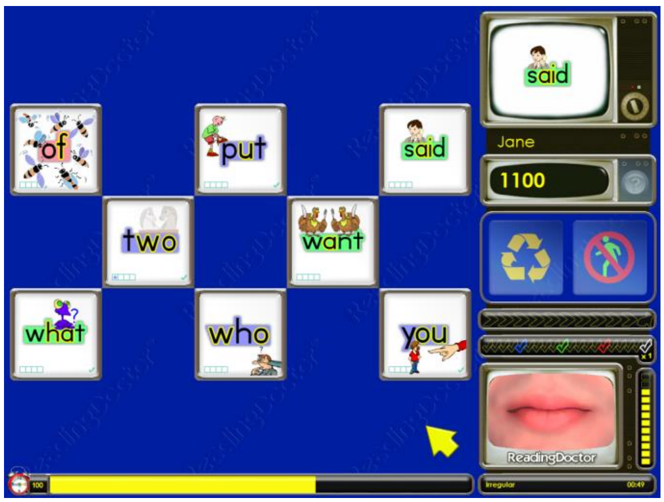

Figure 4: A sight word recognition game in ReadingDoctor, where words are not blended but recognised by emphasising the shape of the word.

Sight words are not blended or segmented, but instead the shape of the word is emphasised with coloured highlighting; in Figure 4 the words are highlighted in purple, orange and green. The difference between regular and irregular components of the word is emphasised with yellow highlighting (for irregular components) around the particular letters; for example, in Figure 4 in the word "two", the "t" is not highlighted because it is pronounced as expected, phonologically, whereas the "wo" is highlighted in yellow because its pronunciation is irregular. The current word that the student needs to click on (the yellow arrow is the cursor) is shown in the upper-right corner of the television screen. The mouth in the lower-right corner produces the word for the child to do auditory word-picture matching. For each word, the more times a student correctly clicks on the particular word, the more the visual and auditory cues fade. The student can start from any level and the number of words can be varied from 2 to 25.

Why ReadingDoctor?

A key reason for choosing ReadingDoctor (RD) for this study is that it was developed locally and is widely used in South Australian schools. The other key reason is that when the researchers evaluated RD against criteria suggested in the literature as being important for educational software, both from a pedagogical and technological point of view, RD stands up well.

From a pedagogical point of view, a key feature of RD is that it uses a multisensory approach to teach children phonemic awareness and to develop visual word recognition skills. There are four important reasons for the approach taken by RD. First, multisensory approaches to teaching reading—approaches involving more than one of visual, auditory, kinaesthetic and tactile elements—have been found to be more effective than single-sense approaches (Rose, 2009). A multisensory approach is particularly critical for poor readers (Finn et al., 2014; Rose, 2009). Secondly, a growing body of research shows differences in the way the brains of good and poor readers function when reading (Shaywitz & Shaywitz, 2005). Poor readers tend to show less activity in areas of the brain that link language to visual cues (Finn et al., 2014), and the region of the brain responsible for multisensory integration of letters and speech sounds has been shown to be underactive in both children (Blau et al., 2010) and adults (Blau et al., 2009) with dyslexia. Thirdly, a number of studies have now shown that with appropriate, phonics-based intervention, the brain functioning of poor readers begins to resemble that of more proficient readers (Richards & Berninger, 2008; Shaywitz & Shaywitz, 2005). Fourthly, the use of technology to provide multisensory learning enables RD to deliver a level of multisensory learning that would be very difficult for a classroom teacher to provide without the use of technology, and allows the multisensory learning to be differentiated for individual students; this may increase learning (e.g. Haelermans et al., 2014).

Another key, pedagogical principle that is well implemented in the software is the importance of timely, constructive feedback. Wiliam (2007) provides evidence of a positive relationship between constructive feedback and improved student results. When a student using RD gives a wrong answer, they are immediately provided with feedback that their answer is wrong and a succinct explanation as to why. This level of consistent feedback would be very difficult for a classroom teacher to provide to individual students.

ReadingDoctor (RD) meets all of Bishop and Santoro's (2005) criteria for beginning reading software. For example, from a pedagogical point of view, Bishop and Santoro argue that beginning reading software must focus on phonological and phonemic awareness, and RD includes an intensive focus on phonological and phonemic awareness, as described in the previous section. In addition, Bishop and Santoro argue that software must be "systematic" (p. 61) and "instructionally supportive" (p. 62).

RD is systematic because it provides numerous ways of scaffolding the program so that it can be tailored to particular student needs, and students can advance to new levels of attainment. For example, as a student correctly identifies a letter/sound correspondence more and more times, both visual and auditory cues decrease. This fading of prompts matches Woolley's (2010) advice that teachers should scaffold student vocabulary learning so that students "develop effective metacognitive strategies to monitor and regulate their own meaning-making activity" (p. 128) without constant teacher support (p. 121), and Wild's (2009) suggestion that scaffolding should become less explicit as the learner's skills progress. In contrast to the visual and auditory cues, the articulatory cue—the mouth that makes the appropriate movements to produce the particular sound under consideration—does not fade as the learner progresses. This is because the purpose of the articulatory cue is to demonstrate explicitly the orthographic representation of speech sounds in language. This is important, because there is an established link between articulatory awareness and reading ability (Adair et al., 1999; Griffiths & Frith, 2002; Montgomery, 1981).

RD is also instructionally supportive, providing clear audio instructions on mouse click, as appropriate, to explain to students what their task is and how to use the software to complete the task. For example, students hear the message "click on the green man to start the game", and are able to click on a repeat symbol to re-hear particular instructions during the game.

Shifting to an instructional design point of view, Bishop and Santoro (2005) argue that software needs to be "aesthetic" (p. 59), "operationally supportive" and "interactive" (p. 60). RD provides a consistent look and feel for all games throughout the software, thus meeting the "aesthetic" criterion. It is also operationally supportive, because it provides considerable, multisensory support for navigating around the program, and it is highly interactive. Overall, RD thus meets all of Bishop and Santoro's (2005) characteristics for beginning reading software.

ReadingDoctor also accords with Clark and Mayer's principles (2003), of which the three most significant are discussed here in more detail. Both letters and words, and pictures representing them, are regularly used to facilitate learning (the Multimedia Principle). Text is clearly aligned to corresponding graphics in a variety of ways so that both text and picture can be seen at the same time at the appropriate levels of presentation (the Contiguity Principle). Students hear a sound or word spoken and are required to match what they hear with the corresponding graphic (the Modality Principle).

In summary, ReadingDoctor was chosen for this study because it is locally designed, widely used in the local context, and designed in a way that is consistent with current research on how to use software to teach literacy skills.

Method

Ethics

This study received ethics approval from the Flinders University Social and Behavioural Research Ethics Committee and the University of Adelaide Human Research Ethics Committee. Parents gave their consent for their children to participate in the study.

Participants

The participants in this study were 45 primary school students, of whom 42 completed the study. The children were aged between 5 and 13 years, and were identified by their teachers as being poor readers, although only four had been given a formal diagnosis of dyslexia. The wide age range reflected the inclination of the participating school that students of all ages should be able to participate. It also allowed comparison of the efficacy of the intervention in a school setting across different age groups.

The participants were divided into two age groups in the analysis of the results—24 children between the ages of 5 and 8 years and 18 children aged between 9 and 13 years—in the expectation that the younger children would show lower performance/ability in reading or phonological processing than the older children.

The Study

All participants undertook set exercises from the ReadingDoctor software program over two six-week periods, with a two-week school holiday break in between. Each student was given exercises designed to strengthen certain aspects of their phonological processing skills. Specific exercises included were: letter/sound recognition; blending (i.e. joining sounds together to make words); segmentation (i.e. separating words into distinct sounds); and sight word recognition (i.e. quick recognition of high frequency words, many of which were irregular). Each student was given the same exercises in each of these areas, and all students were allocated approximately 75% of their work on phonemic awareness and decoding tasks, and 25% of their work on sight word tasks, reflecting the relatively higher importance of decoding ability in learning to read. Each student participated in a learning process involving gradually more challenging levels within games or tasks, according to their competency. Overall, each student was allocated tasks requiring twenty minutes of work each day. Approximately every two weeks, student progress was checked by reviewing the tasks they had completed and the speed with which they were able to complete them, and the set exercises were adjusted accordingly.

In this study, given Warschauer's (2007) advice that the way a tool is used is at least as important as the tool itself, and that a software program is most effective when used together with face-to-face teaching, the use of RD was carefully scaffolded and monitored. For example, each student had several RD games to play during each twenty-minute session. In order to make sure that students were able to use the software correctly, instruction cards with visual cues were prepared for each student to highlight the games they should play. A routine was established where, after completing a game, students drew the attention of the instructor in the room to choose a sticker to display on their card. This allowed the instructor to keep track of whether the software was being used correctly, and also to be present when the student was setting up the subsequent game, and the instructor could then correct any errors in the game selection or setup before the game started. Further, the software automatically provided usage data which was analysed on a regular basis to make sure that students were using the software appropriately; the vast majority of students were doing so, with only a few exceptions where students missed games they should have played.

Research Design

Originally, the study was designed as a single-subject design in anticipation of a small sample size. However, as more students with reading difficulty were identified and more computers became available, the participating school requested that up to 45 students participate in the study. Consequently, the current study was a quasi-experimental, one-group study without a control group, but with three multiple baseline measures to control for repeated testing.

A total of seven tests were conducted throughout the study to evaluate the efficacy of RD. The same test measures for phonemic awareness and reading efficiency at word level were taken each time. Three baseline measures or tests were taken before the intervention, with the third baseline measure scheduled immediately prior to the start of the intervention. The fourth test was conducted following six weeks of using RD. There was then a two-week school holiday break with no exposure to RD. The fifth test was conducted at the end of these two weeks. The sixth test was conducted at the end of the second period of six-week exposure to RD. The seventh and final test was conducted three months later, after a six-week summer holiday break and the start of the new school year. The final test was conducted to demonstrate whether or not any improvement in phonemic awareness and/or reading ability was maintained over the three-month period of non-exposure.

During each of the seven testing sessions, each student was tested using three different measures of reading ability: the Sutherland Phonological Awareness Test - Revised (SPAT-R) (Neilson, 2009), the Test of Word Reading Efficiency (TOWRE) (Torgeson et al., 1999), and the graphemes and decoding subtests of the Phonological Awareness Test 2 (PAT-2) (Robertson & Salter, 2007). These measures were used rather than some form of connected text because they enabled detailed testing to be completed in an appropriate timeframe and because of the wide age range of children (5 to 13 years) in the study. The SPAT-R provides a focus on early phonological skills, such as rhyming, and is normed on Australian children; its inclusion allowed pre-reading skills to be tested in younger children and highlighted the areas in which some older children were struggling. The TOWRE is a timed test and measures the speed of (word and nonword) reading accuracy. It was particularly useful with some older children who demonstrated a very different capability with respect to reading nonwords compared to words. The PAT-2 was included because it tests graphemes of increasing difficulty, and also includes a nonword component (the decoding subtest), enabling a more precise evaluation of decoding difficulty than either of the other two tests. For example, the PAT-2 includes nonwords in the following groups: Vowel-Consonant (VC), Consonant-Vowel-Consonant (CVC), consonant digraphs, consonant blends, vowel digraphs, R-controlled vowels, CVC words ending in *e*, and diphthongs. The TOWRE and the SPAT-R each have an A or B version of the test; the A version was used each time for consistency in the measurement. Seven different testers were trained in the administration of the tests and participated in the testing process, each following the guidelines for administration provided by each of the tests.

Results

The mean group performances on the two assessments for phonemic awareness (SPAT-R and PAT-2) and the assessment for word reading proficiency (TOWRE) were analysed as follows. First, three two-way Analyses of Variance (ANOVA) with repeated measures (3) were conducted for each assessment across the three baseline measures to test for any significant differences in the overall mean group performances before the start of the first six weeks of computer-assisted intervention. Secondly, a two-way ANOVA with repeated measures (5) was performed on the mean group performance in each of the three tests across five points in time to determine the overall significance between the mean performances on each of the three assessments. Hence, the within-subjects variable comprised testing at baseline 3; after the first six-week block of computer-assisted intervention; after the two-week break; after the second six-week block of computer-assisted intervention; and 3 months after the intervention. The between-group variable was Age, comprising predicted differences in the performances of the 5-8 year olds and 9-13 year olds.

Apriori pairwise comparisons of specific interest were the comparisons of the mean test performances between baseline 3 and performance after the first six-week block of computer-assisted intervention, between baseline 3 and performance after the second six-week block of computer-assisted intervention, and between baseline 3 and performance at 3 months after the entire intervention. All pairwise comparisons were conducted with the Bonferroni adjustment.

Baseline Measures

The mean performances and standard deviations across the three baseline measures are displayed in Table 1. The three two-way ANOVAs with repeated measures showed significant differences across the three baseline measures for each assessment as follows: SPAT-R ($F(2, 80) = 18.44$, $p < 0.001$); PAT-2 ($F(2, 80) = 7.75$, $p < 0.001$); TOWRE ($F(2, 80) = 19.51$, $p < 0.001$). Significant Group differences were also found for each assessment [SPAT-R ($F(1, 80) = 20.94$, $p < 0.001$); PAT-2 ($F(1, 80) = 25.69$, $p < 0.001$); TOWRE ($F(1, 80) = 44.76$, $p < 0.001$)], whereas none of the three interaction effects (Baseline measure x Group) was significant [SPAT-R ($F(2, 80) = 0.38$, $p = 0.68$); PAT-2 ($F(2, 80) = 0.52$, $p = 0.60$); TOWRE ($F(2, 80) = 1.10$, $p = 0.34$)]. The significant group differences combined with the lack of significant interaction effects indicate a parallel improvement in performance across the two age groups, i.e. a test-retest or learning effect in both age groups.

Post hoc pairwise comparisons using the Bonferroni correction revealed a significant test-retest or learning effect between baseline measure 1 and baseline measures 2 and 3, but not between baseline measures 2 and 3 on all three assessments (see Table 1).

The total mean score across the two age groups went up on the SPAT-R by an average of 3.46 ($p < 0.001$) between baseline measures 1 and 2, and by an average of 3.97 ($p < 0.001$) between baseline measures 1 and 3. The average difference of 0.51 between baseline measures 2 and 3 was non-significant.

The total mean score on the PAT-2 increased by an average of 6.35 ($p < 0.001$) between baseline measures 1 and 2, and by an average of 6.88 ($p < 0.001$) between baseline measures 1 and 3. The average difference of 0.54 between baseline measures 2 and 3 was non-significant.

The total mean score on the TOWRE increased by an average of 4.62 ($p < 0.001$) between baseline measures 1 and 2, and by an average of 5.54 ($p < 0.001$) between baseline measures 1 and 3, whereas the average difference of 0.92 between baseline measures 2 and 3 was again non-significant.

	SPAT-R		PAT-2		TOWRE	
	Mean	StDev	Mean	StDev	Mean	StDev
Baseline 1 overall	**37.90**	**12.79**	**73.40**	**29.77**	**51.31**	**31.26**
5 to 8 years	31.88	13.26	58.29	26.81	31.79	25.84
9 to 13 years	37.90	12.79	93.56	20.32	77.33	14.11
Baseline 2 overall	**41.31**	**12.54**	**79.50**	**31.49**	**55.74**	**33.30**
5 to 8 years	34.96	12.82	62.88	29.82	35.08	28.55
9 to 13 years	49.76	4.91	101.67	16.67	83.28	12.89
Baseline 3 overall	**41.90**	**11.77**	**80.05**	**31.46**	**56.69**	**33.44**
5 to 8 years	36.08	11.74	63.50	29.81	36.25	28.31
9 to 13 years	49.67	6.03	102.11	16.82	83.94	15.38

	SPAT-R		PAT-2		TOWRE	
	Mean	StDev	Mean	StDev	Mean	StDev
Comp program 1	**43.95**	**11.50**	**85.26**	**29.07**	**61.19**	**34.05**
5 to 8 years	38.46	11.86	70.83	27.74	41.17	30.04
9 to 13 years	51.28	5.41	104.50	17.66	87.89	15.94
Break overall	**43.79**	**11.61**	**85.29**	**29.33**	**62.76**	**33.53**
5 to 8 years	38.29	11.74	69.25	26.61	41.46	27.04
9 to 13 years	51.11	6.26	106.67	16.39	91.17	15.06
Comp program 2	**45.48**	**11.39**	**88.64**	**29.77**	**65.05**	**34.21**
5 to 8 years	39.92	12.079	74.29	30.14	43.88	29.33
9 to 13 years	52.89	3.71	107.18	15.13	93.28	13.52
Maintenance overall	**46.36**	**11.05**	**91.88**	**31.20**	**66.14**	**34.52**
5 to 8 years	41.08	11.58	78.08	33.57	45.67	30.73
9 to 13 years	53.39	4.63	110.28	14.00	93.44	14.68

Table 1: Mean group performances and standard deviations on the two assessments of phonological awareness (SPAT-R and PAT-2) and word reading efficiency (TOWRE).

Forty-two students participated in the study; results were analysed overall (n = 42) as well as in two groups, one of 5 to 8 year olds (n = 24) and another of 9 to 13 year olds (n = 18).

Consequently, baseline measure 3 was used in determining the overall efficacy of the computer-assisted remedial reading program.

Overall Efficacy of the Computer-Assisted Remedial Reading Program (ReadingDoctor)

The three two-way ANOVAs with repeated measures showed significant differences across the five assessment periods during the intervention (baseline 3, first six-week block, break, second six-week block, and maintenance after 3 months): SPAT-R ($F(4, 160) = 12.05$, $p < 0.001$); PAT-2 ($F(4, 160) = 12.06$, $p < 0.001$); TOWRE ($F(4, 160) = 7.71$, $p < 0.001$). Significant Group differences were also found for each assessment [SPAT-R ($F(1, 40) = 20.33$, $p < 0.001$); PAT-2 ($F(1, 40) = 22.33$, $p < 0.001$); TOWRE ($F(1, 40) = 45.87$, $p < 0.001$)], whereas none of the three interaction effects (Baseline measure x Group) was significant [SPAT-R ($F(4, 160) = 0.22$, $p = 0.93$); PAT-2 ($F(4, 160) = 1.29$, $p = 0.28$); TOWRE ($F(4, 160) = 0.22$, $p = 0.89$)].

The overall TOWRE score is based on the combined scores for reading real words or sight word efficiency and reading nonwords or phonemic coding efficiency. Performance of both subcomponents improved significantly over the duration of the study: sight word efficiency ($F(4, 160) = 5.87$, $p < 0.001$), and phonemic decoding efficiency ($F(4, 160) = 6.37$, $p < 0.001$. As expected, the younger age group showed significantly lower performance: sight word efficiency ($F(1, 40) = 51.98$, $p < 0.001$), phonemic decoding efficiency ($F(1, 40) = 22.82$, $p < 0.001$) (see Table 2). The interaction effects were again not significant.

	Sight Word Efficiency		Phonemic Coding Efficiency		TOWRE	
	Mean	**StDev**	**Mean**	**StDev**	**Mean**	**StDev**
Baseline 1 overall	**38.14**	**23.02**	**13.17**	**9.89**	**51.31**	**31.26**
5 to 8 years	23.17	19.12	8.63	7.31	31.79	25.84
9 to 13 years	58.11	6.71	19.22	9.78	77.33	14.11
Baseline 2 overall	**40.07**	**23.15**	**15.67**	**11.15**	**55.74**	**33.30**
5 to 8 years	25.00	19.33	10.08	9.80	35.08	28.55
9 to 13 years	60.17	6.23	23.11	8.20	83.28	12.89
Baseline 3 overall	**41.69**	**24.08**	**15.00**	**11.37**	**56.69**	**33.44**
5 to 8 years	26.25	20.73	10.00	8.49	36.25	28.31
9 to 13 years	62.28	5.63	21.67	11.49	83.94	15.38
Comp program 1	**43.31**	**23.28**	**17.88**	**12.59**	**61.19**	**34.05**
5 to 8 years	28.63	20.46	12.54	10.61	41.17	30.04
9 to 13 years	62.89	5.56	25.00	11.67	87.89	15.94
Break overall	**44.07**	**22.76**	**18.69**	**12.28**	**62.76**	**33.53**
5 to 8 years	29.54	19.61	11.92	8.28	41.46	27.04
9 to 13 years	63.44	5.74	27.72	10.94	91.17	15.06
Comp program 2	**46.21**	**23.85**	**18.83**	**11.97**	**65.05**	**34.21**
5 to 8 years	31.25	21.15	12.63	9.41	43.88	29.33
9 to 13 years	66.17	5.34	27.11	9.94	93.28	13.52
Maintenance overall	**46.52**	**22.88**	**19.62**	**13.28**	**66.14**	**34.52**
5 to 8 years	32.46	20.82	13.21	11.09	45.67	30.73
9 to 13 years	65.28	4.98	28.17	11.13	93.44	14.68

Table 2: Mean group performances and standard deviations on the two subcomponents (reading real words or sight word efficiency and reading nonwords or phonemic coding efficiency) of the assessment of word reading efficiency (TOWRE), shown together with the combined TOWRE score.

Forty-two students participated in the study; results were analysed overall (n = 42) as well as in two groups, one of 5 to 8 year olds (n = 24) and another of 9 to 13 year olds (n = 18).

In summary, there were significant overall differences in the average performances across the intervention period, despite the fact that, overall, younger participants showed a significantly lower mean performance than the older participants across the study.

Using the Bonferroni correction, post hoc comparisons between baseline measure 3 and the assessment after the two blocks of the computer-assisted reading program revealed significant improvements in the performances, i.e. the overall efficacy of RD in improving phonemic awareness, and word reading efficiency. In terms of phonemic awareness, the total mean score on the SPAT-R went up by an average of 3.53 ($p < 0.001$), and the total mean score on the PAT-2 increased by an average of 8.23 ($p < 0.001$). In terms of word reading efficiency, the total mean score on the TOWRE increased by an average of 8.48 ($p < 0.001$).

Post hoc analysis also revealed that improved phonemic awareness was sustained 3 months after the computer-assisted reading program; the mean score difference between baseline measure 3 and maintenance was 4.36 ($p < 0.001$) on the SPAT-R and 11.38 ($p < 0.001$) on the PAT-2. The significant improvement was also found with respect to reading efficiency with an average improvement on the TOWRE of 9.46 ($p < 0.001$) between baseline measure 3 and mean performance 3 months after the intervention.

On the other hand, post hoc testing between baseline measure 3 and the assessments after the first six-week block of computer-assisted reading did not reveal significant improvements in phonemic awareness given the smaller improvements in average performance of 1.89 (p=0.03) on the SPAT-R and of 4.86 (p= 0.06) on the PAT-2. The same was found regarding reading efficiency with an average increase of 1.49 (p= 1.0) on the TOWRE.

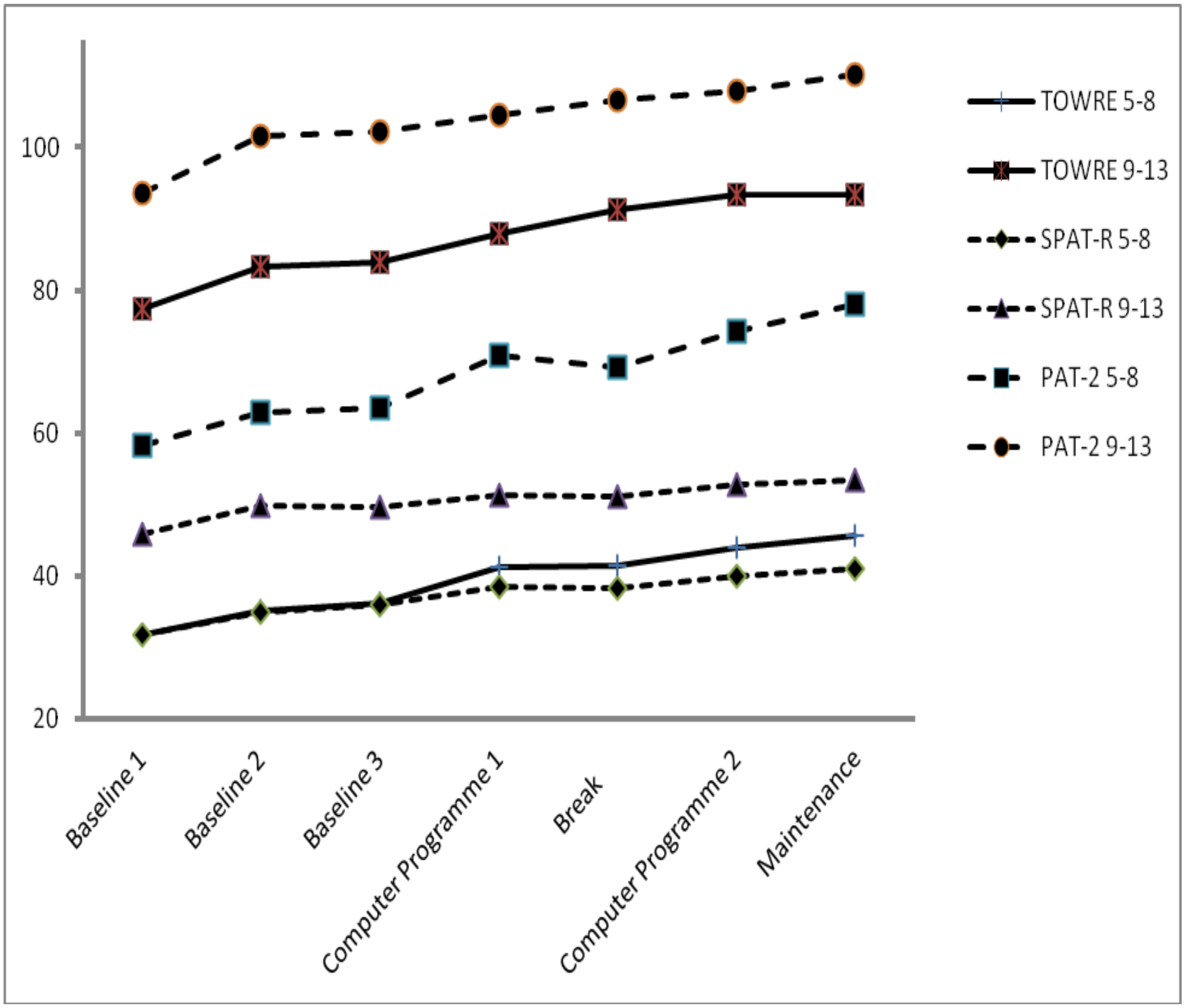

Figure 5: Mean group performance on assessments of phonemic awareness and word reading efficiency divided by age (5-8 years (n = 24) vs. 9-13 years (n = 18)).

Discussion

Overall, the results indicate that the use of RD made a significant difference to both phonemic awareness and word reading efficiency after two six-week blocks of the program, regardless of age. With respect to phonemic awareness, the results from the SPAT-R and PAT-2 showed significant improvements when comparing the third baseline measure with the results after the second block but not the first block. With respect to word reading efficiency, the TOWRE test also showed significant improvement, again after the second but not the first block.

Both subcomponents of the TOWRE—sight word reading and phonemic decoding of nonwords—showed a significant improvement across the duration of the study.

All improvements were significant in both age groups, although, as expected, the performances of the younger age group were significantly lower; younger children would be expected to have less well-developed knowledge of phonological and phonemic awareness. All improvements after the second block of RD were also long lasting, in that they were upheld three months after the second, final block.

Overall, the findings show the efficacy of (supplementary) computer-assisted instruction with RD in children with reading difficulties, improving phonological and, more specifically, phonemic awareness as well as basic word reading skills in terms of reading (decoding) nonwords and the efficiency of sight word reading.

The results are particularly interesting because of the wide age range of children involved in the study (5-13 years). Suggate (2010) argued that phonics-based interventions are more effective for younger children (before grade one) and Shaywitz et al. (2008) argued that remedial approaches in general are less effective post-second grade. Torgesen et al. (2010), however, found that phonological intervention was also effective for 8 to 10 year old children with very low word reading ability. The current study suggests that CAI that focuses on phonemic awareness and decoding skills is effective, and may be potentially helpful for older as well as younger children, since fourteen children were aged between 10 and 13 years at the time of the first baseline test, and results of younger (5 – 9 years) and older (10 – 13 years) children were analysed in two separate groups. Further work should involve a larger study comparing children in two age groups. The aim would be to see whether the reading gains experienced in each group are similar, and whether a benefit in increasing phonemic awareness can be demonstrated for older children. Also, a study that includes a ‘business as usual’ group and a group being delivered the same content by a teacher would be beneficial.

Another interesting aspect of the results of this study is that the use of RD led to gains in both phonemic awareness and sight reading. In the large Canadian study where sight word reading, fluency and comprehension were evaluated, the use of ABRA did not lead to gains in these areas (Savage et al., 2013). It may be that the combined, intense focus on phonemic and phonological awareness, together with sight word practice, led to this result. However, given the limitations of this study with respect to the lack of a control group, more research is needed to draw definite conclusions in this area.

Phonemic and single-word reading measures, rather than some form of connected text, were used in this study because they enabled detailed testing to be completed in an appropriate timeframe and because of the wide age range of children (5 to 13 years) in the study. The SPAT-R test allowed pre-reading skills to be tested in younger children and highlighted areas in which some older children were struggling. The TOWRE tested the speed of both word and nonword reading skills and this was particularly useful with some older children who demonstrated a very different capability when it came to the speed and accuracy with which they could read words and nonwords. The PAT-2 subtests tested detailed grapheme knowledge. Interestingly, there is evidence (Kjeldsen et al., 2014) that phonological awareness levels in kindergarten predict reading comprehension as late as year 9, so tests such as those used in this study may increasingly be used as measures to predict future reading ability.

An important consequence of not having a control group in a study is that the findings in a single group study could be attributed to a test-retest practice effect. However, while a test-retest effect cannot be completely ruled out―and this is a limitation of the study―the improvements in phonemic awareness and word reading are unlikely to be the result of a test-retest practice effect, since a practice effect was present for all three assessments between the first and second pretests, and then for none of the assessments between the second and third pretests. Any practice effect is likely to be largest between the first and second attempts at the test (test 1 and test 2) and will wear off subsequently. Performances post intervention (after the first and second intervention period of six weeks) were compared with the performance on the third baseline measure, by which time the (test) practice effect was absent. [2] Future work should be focussed on whether or not these results can be replicated in a study with a control group, preferably with random allocation of the participants to the experimental and control groups. Another limitation was the decision to only use Form A of the TOWRE and SPAT-R tests. Forms A and B should have been alternated to reduce the potential influence of test-retest effects. However, the time between the final pre-test and the first post-test was six to seven weeks. While this pair of tests (test 3 and test 4) showed a significant improvement in skills tested, there was no such improvement between the first post-test (test 4) and subsequent pre-test (test 5) before the second period of intervention. Since testing periods 4 and 5 served the purpose of re-testing after a school break where students had not used the software, and were only two to three weeks apart, it would seem to suggest that any test-retest effect during intervention periods was minimal. However, it would have been useful to retest during normal instruction rather than over a holiday break—in other words, to stop the program while normal classroom instruction continued and retest following this pause—in order to completely rule out a test-retest effect.

A further limitation of the study is that the inclusion criteria for the study were for students to be identified by their teachers as having difficulty reading. A future study ought to determine inclusion by testing students using recognised assessment measures.

The findings of the current study support the growing body of research suggesting that phonological awareness, especially phonemic awareness, is critical for reading acquisition (Ehri et al., 2001; McArthur & Castles, 2013). During this study, participants did their normal reading curriculum activities in the classroom, but they did no extra reading practice in the classroom or as part of the study. Participants simply undertook RD exercises that focussed on phonemic awareness (75% of exercises) and sight word recognition (25% of exercises). This use resulted in a clear improvement in their word and non-word reading ability, assessed using the TOWRE. Further, Whiteley et al. (2007) found letter knowledge to be a key factor limiting improvement in response to phonologically-based intervention. Blau et al. (2009) and Blau et al. (2010) concluded that both children and adults with dyslexia have inadequate multisensory representations of letter/sound pairs in the brain, and that these inadequate representations are linked to poor reading. All participants in this study explicitly practised letter/sound recognition skills in the multisensory environment provided by RD as part of their phonological awareness exercises; this may be part of the reason for the improvements gained.

The findings of the study are consistent with earlier research suggesting that it is possible to use a well-designed software program to increase phonological awareness and reading skills (Torgesen et al., 2010; Wild, 2009). While Wild found that the use of a software program provided increased improvement compared to a non-CAI approach, Torgesen et al. found that a CAI approach was indistinguishable from a non-CAI approach in improvement gained. Nevertheless, both Wild and Torgesen et al. indicated improvement in skills due to CAI. The current study did not compare a CAI approach to a non-CAI approach, but it would be useful to

---

[2] There was a practice effect between the first and second baseline measures; hence all comparisons were based on the performance on the third baseline measure.

do so in an effort to better establish which kinds of computer software are most effective for supporting reading skills. Further, it may be that the way in which the software is used is important (Warschauer, 2007).

With respect to poor readers, the results in this study are consistent with a review of 20 studies by Cheung and Slavin (2013), who found a larger effect of CAI on reading ability for poor or struggling readers. All participants in the current study had been identified by their teachers as poor readers, and results on all three testing measures showed significant improvement due to the use of RD. However, it needs to be noted that while the review by Cheung and Slavin focuses on the effect of educational technology on reading outcomes for struggling readers, the term ‘educational technology’, as used by Cheung and Slavin, encompasses a broad range of programs, including comprehensive literacy intervention models. For example, one of the programs reviewed in two studies, READ 180, is a comprehensive literacy intervention program which combines both computer and non-computer intervention. In addition, a number of the programs reviewed by Cheung and Slavin do not teach phonological or phonemic awareness, but instead focus on other areas, such as reading comprehension. There is still a pressing need for independent studies of individual software programs, especially programs focusing on phonological and phonemic awareness, in order to determine their efficacy in a variety of classroom settings.

The use of computer software may be a more time-efficient way of gaining improvement in reading for struggling readers. For instance, in this study, two adults supervised 45 children for twenty minutes each session. Without the use of the software, ensuring 45 children completed as many repetitions of the assigned activities would likely have entailed many hours of work, involving, for example, flashcards, children working in a potentially noisy classroom without headphones, and teachers having to attempt to provide scaffolding that the computer program provides automatically and that is very difficult to provide manually. It must be noted, however, that there was a significant effort required to keep updating the program for 45 students individually. Students needed variety in their exercises so that they did not get bored, and also to challenge them further as they progressed. Monitoring their progress and updating their program took approximately four hours for the group every ten days. Nevertheless, while there was some significant associated work, the use of RD was deemed to be a time-efficient way of delivering phonemic awareness and decoding skill practice when compared to manual instruction.

In summary, and consistent with previous research, the findings of the present study show that regular use of a computer software program designed to increase phonemic awareness, grapheme naming and decoding skills significantly improves reading skills in primary school children with reading difficulties. This study adds to previous research by demonstrating that this particular software program, ReadingDoctor, is an effective tool for improving phonemic awareness and reading skills, including sight word reading. In addition, it would be worth investigating whether the benefits for older children in phonemic awareness and reading skills demonstrated in this study can be replicated in a larger study.

References

Abrami, P., Wade, C., Lysenko, L., Marsh, J., & Gioko, A. (2016). Using educational technology to develop early literacy skills in Sub-Saharan Africa. *Education and Information Technologies*, *21*(4), 945-964. https://doi.org/10.1007/s10639-014-9362-4

Adair, J. C., Schwartz, R. L., Williamson, D. J. G., Raymer, A. M., & Heilman, K. M. (1999). Articulatory processes and phonologic dyslexia. *Neuropsychiatry, Neuropsychology and Behavioral Neurology*, *12*(2), 121–127.

Bass, G., Ries, R., & Sharpe, W. (1986). Teaching basic skills through microcomputer assisted instruction. *Journal of Educational Computing Research*, *2*(2), 207–219. https://doi.org/10.2190/KEAN-RWUX-7BL2-FP3V

Bishop, M. J., & Santoro, L. E. (2005). Evaluating beginning reading software for at-risk learners. *Psychology in the Schools*, *43*(1), 57-70. https://doi.org/10.1002/pits.20129

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology*, *19*(6), 503–508. https://doi.org/10.1016/j.cub.2009.01.065

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: A functional magnetic resonance imaging study of dyslexic children. *Brain*, *133*(3), 868–879. https://doi.org/10.1093/brain/awp308

Bond, J., Coltheart, M., Connell, T., Firth, N., Hardy, M., Nayton, M., Shaw, J., & Weeks, A. (2010). *Helping people with dyslexia: A national action agenda*. www.dyslexiaaustralia.com.au/DYSWP.pdf

Bosse, M.-l., Chaves, N., Largy, P., & Valdois, S. (2013). Orthographic learning during reading: The role of whole-word visual processing. *Journal of Research in Reading* (online edition). https://doi.org/10.1111/j.1467-9817.2012.01551.x

Carson, K., Boustead, T., & Gillon, G. (2014). Predicting reading outcomes in the classroom using a computer-based phonological awareness screening and monitoring assessment (Com-PASMA). *International Journal of Speech-Language Pathology*, *16*(6), 552-561. https://doi.org/10.3109/17549507.2013.855261

Case, S. Philpot, D., & Walker, J. (2009). *Sounds-Write report to schools: Longitudinal study of literacy development from 2003-2009, following 1607 pupils through Key Stage 1*. http://www.sounds-write.co.uk/sites/soundswrite/uploads/files/42-sounds_write_research_report_2009.pdf.

Chera, P., & Wood, C. (2003). Animated multimedia "talking books" can promote phonological awareness in children beginning to read. *Learning and Instruction, 13*(1), 33-52. https://doi.org/10.1016/S0959-4752(01)00035-4

Cheung, A., & Slavin, R. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly, 48* (3), 277-299. https://doi.org/10.1002/rrq.50

Clark, R. (2002, September 10). Six principles of effective e-learning: What works and why. *Learning Solutions e-Magazine*. http://www.eLearningGuild.com

Clark, R. C., & Mayer, R. E. (2003). *e-Learning and the science of instruction* (2nd ed.). Wiley & Sons, Inc.

Conole, G., Dyke, M., Oliver, M., & Seale, J. (2004). Mapping pedagogy and tools for effective learning design. *Computers and Education*, *43*(1-2), 17-33. https://doi.org/10.1016/j.compedu.2003.12.018

Dynarski, M., Agodini, R., Heaviside, S.N., Novak, T., Carey, N., Campuzano, L., Means, B., Murphy, R., Penuel, W., Javitz, H., Emery, D., & Sussex, W. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort* (NCEE 2007-4005). Federal Reserve Bank of St Louis.

Educational App Store. (2021). https://www.educationalappstore.com/app-lists/apps-for-education.

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. https://doi.org/10.1598/RRQ.36.3.2

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. E., Michaels, R., & Shaywitz, S. E. (2015). Achievement gap in reading is present as early as first grade and persists through adolescence. *The Journal of Pediatrics*, *167*(5). https://doi.org/10.1016/j.jpeds.2015.07.045

Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., Shaywitz, S.E., Shaywitz, B.A., & Constable, R. T. (2014). Disruption of functional networks in dyslexia: A whole-brain, data-driven analysis of connectivity. *Biological Psychiatry*, *76*(5), 397-404. https://doi.org/10.1016/j.biopsych.2013.08.031

Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities: Research and Practice*, *16*(4), 203-212. https://doi.org/10.1111/0938-8982.00020

Gambrell, L. B., Morrow, L. M., & Pressley, M. (Eds.) (2007). *Best practices in literacy instruction* (3rd ed.). Guilford Publications, Inc.

Government of South Australia, Department of Education. (2018). *2018 Phonics screening check.* https://www.education.sa.gov.au/sites/default/files/2018-phonics-screening-check-fact-sheet.pdf

Griffiths, S., & Frith, U. (2002). Evidence for an articulatory awareness deficit in adult dyslexics. *Dyslexia*, *8*(1), 14–21. https://doi.org/10.1002/dys.201

Haelermans, C., Ghysels, J., & Prince, F. (2014). Increasing performance by differentiated teaching? Experimental evidence of the student benefits of digital differentiation. *British Journal of Educational Technology*, *46*(6), 1161–1174. https://doi.org/10.1111/bjet.12209

Kast, M., Meyer, M., Vögeli, C., Gross, M., & Jäncke, L. (2007). Computer-based multisensory learning in children with developmental dyslexia. *Restorative Neurology & Neuroscience*, *25*(3/4), 355-369. www.metapress.com

Kast, M., Baschera, G.-M., Gross, M., Jäncke, L., & Meyer, M. (2011). Computer-based learning of spelling skills in children with and without dyslexia. *Annals of Dyslexia*, *61*(2), 177-200. https://doi.org/10.1007/s11881-011-0052-2

Kuppen, S. E. A., & Boswami, U. (2016). Developmental trajectories for children with dyslexia and low IQ poor readers. *Developmental Psychology*, *52*(5), 717-734. https://doi.org/10.1037/a0040207

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. https://doi.org/10.1111/j.1467-9817.2006.00282.x

Macaruso, P., & Walker, A. (2008). The efficacy of computer-assisted instruction for advancing literacy skills in kindergarten children. *Reading Psychology, 29*, 266-287. https://doi.org/10.1080/02702710801982019

McArthur, G., & Castles, A. (2013). Phonological processing deficits in specific reading disability and specific language impairment: Same or different? *Journal of Research in Reading*, *36*(3), 280-302. https://doi.org/10.1111/j.1467-9817.2011.01503.x

Montgomery, D. (1981). Do dyslexics have difficulty accessing articulatory information? *Psychological Research*, *43*, 235–243. https://doi.org/10.1007/BF00309832

Neilson, R. (2009). *Sutherland Phonological Awareness Test-Revised* [Measurement instrument]. Dr. Roslyn Neilson.

Ramus, F., Marshall, C. R., Rosen, S., & van der Lely, H. K. J. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain*, *136*(2), 630-645. https://doi.org/10.1093/brain/aws356

Richards, T. L., & Berninger, V. W. (2008). Abnormal fMRI connectivity in children with dyslexia during a phoneme task: Before but not after treatment. *Journal of Neurolinguistics*, *21*(4), 294-304. https://doi.org/10.1016/j.jneuroling.2007.07.002

Robertson, C., & Salter, W., (2007). *Phonological Awareness Test - 2* [Measurement instrument]. Pro-ed.

Rose, J. (2009). *Identifying and teaching children and young people with dyslexia and literacy difficulties.* Department for Children, School, and Families.

Savage, R., & Carless, S. (2008). The impact of early reading interventions delivered by classroom assistants on attainment at the end of Year 2. *British Educational Research Journal*, *34*(3), 363-385. https://doi.org/10.1080/01411920701609315

Savage, R., Abrami, P. C., Piquette, N., Wood, E., Deleveaux, G., Sanghera-Sidhu, S., & Burgos, G. (2013). A (Pan-Canadian) cluster randomized control effectiveness trial of the ABRACADABRA web-based literacy program. *Journal of Educational Psychology*, *105*(2), 310-328. https://doi.org/10.1037/a0031025

Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological Psychiatry*, *57*(11), 1301-1309 https://doi.org/10.1016/j.biopsych.2005.01.043

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology*, *59*, 451-475. https://doi.org/10.1146/annurev.psych.59.103006.093633

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/rrq.21.4.1 https://doi.org/10.1598/RRQ.21.4.1

Suggate, S. P. (2010). Why what we teach depends on when. *Developmental Psychology*, *46*(6), 1556-1579. https://doi.org/10.1037/a0020612

Tanaka, H., Black, J. M., Hulme, C., Stanley, L. M., Kesler, S. R., Whitfield-Gabrieli, S., Reiss, A. L., Gabrieli, J. D., & Hoeft, F. (2011). The brain basis of the phonological deficit in dyslexia is independent of IQ. *Psychological Science*, *22*(11), 1442-1451. https://doi.org/10.1177/0956797611419521

Thomson, S., Hillman, K., Schmid, M., Rodrigues, S., & Fullarton, J. (2017). *Highlights from PIRLS 2016: Australia's perspective* [PDF file]. Australian Council for Educational Research. https://research.acer.edu.au/cgi/viewcontent.cgi?article=1001&context=pirls

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A., (1999). *Test of Word Reading Efficiency* [Measurement instrument]. Pro-ed.

Torgesen, J. 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. https://doi.org/10.1007/s11881-009-0032-y

Tunmer, W., & Greaney, K. (2010). Defining dyslexia. *Journal of Learning Disabilities*, *43*(3), 229-243. https://doi.org/10.1177/0022219409345009

Vadasy, P.F. (2010). Efficacy of supplemental phonics-based instruction for low-skilled kindergarteners in the context of language minority status and classroom phonics instruction. *Journal of Educational Psychology*, *102*(4), 786-803. https://doi.org/10.1037/a0019639

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *The Journal of Child Psychology and Psychiatry*, *45*(1), 2-40. https://doi.org/10.1046/j.0021-9630.2003.00305.x

Warschauer, M. (2007). The paradoxical future of digital learning. *Learning Inquiry*, *1*(1), 41-49. https://doi.org/10.1007/s11519-007-0001-5

Whiteley, H. E., Smith, C. D., & Connors, L. (2007). Young children at risk of literacy difficulties: Factors predicting recovery from risk following phonologically based intervention. *Journal of Research in Reading*, *30*(3), 249-269. https://doi.org/10.1111/j.1467-9817.2007.00342.x

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. (pp. 1053– 1098). Information Age Publishing.

Wild, M. (2009). Using computer-aided instruction to support the systematic practice of phonological skills in beginning readers. *Journal of Research in Reading*, *32*(4), 413-432. https://doi.org/10.1111/j.1467-9817.2009.01405.x

Winn, T., & Clark, J. (2011). Tertiary students learning mathematics online: Guidelines for selecting effective software. *International Journal of Instructional Technology and Distance Learning*, *8*, 45-58.

Wolgemuth, J. R., Savage, R., Helmer, J., Harper, H., Lea, T., Abrami, P. C., Kirby, A., Chalkiti, K., Morris, J., Carapetis, J., & Louden, W. (2013). ABRACADABRA aids Indigenous and non-Indigenous early literacy in Australia: Evidence from a multisite randomized controlled trial. *Computers & Education*, *67*, 250-264. https://doi.org/10.1016/j.compedu.2013.04.002

Wood, L. (2008). University learners of mathematics. In H. Forgasz, B. Anastasios, A. Bishop, B. Clarke, S. Keast, W-T. Seah, & P. Sullivan (Eds.), *Research in mathematics education in Australasia* (pp. 73-98). Sense Publishers.

Wood, E., Gottardo, A., Grant, A., Evans, M. A., Phillips, L., & Savage, R. (2012). Developing tools for assessing and using commercially available reading software programs to promote the development of early reading skills in children. *NHSA Dialog*, *15*(4), 350-354. https://doi.org/10.1080/15240754.2012.725489

Woolley, G. (2010). A multiple strategy framework supporting vocabulary development for students with reading comprehension deficits. *Australasian Journal of Special Education*, *34*(2): 119-132. https://doi.org/10.1375/ajse.34.2.119