Title: A meta-analysis of technology-based interventions on the phonological skills of children with dyslexia

Author(s): Thomas W. T. Sim and Zachary M. Walker

Source: Asia Pacific Journal of Developmental Differences, 1(2), 190-201

Published by: Dyslexia Association of Singapore

This document may be used for private study or research purpose only. This document or any part of it may not be duplicated and/or distributed without permission of the copyright owner.

The Singapore Copyright Act applies to the use of this document.

© 2014 Dyslexia Association of Singapore

Archived with permission.
A Meta-Analysis of Technology-Based Interventions on the Phonological Skills of Children with Dyslexia

Thomas W. T. Sim and Zachary M. Walker

DAS Academy, Singapore
National Institute of Education, Singapore

There is a growing awareness of the need to understand how technology can help in education, especially in the area of special educational needs. The purpose of this meta-analysis is to synthesise findings from independent studies gathered by a systematic review of the literature on the effectiveness of technology-based interventions on the phonological skills of children diagnosed with dyslexia in English. Keywords for the literature search were selected that best represented the research area: technology, computer, elearning, mobile learning, ICT; intervention, instruction, remediation, therapy; phonology, phonological skills, spelling; and dyslexia. These key terms were used for the computerised search of five databases: Academic Search Premier, Education Research Complete, ERIC, PsycARTICLES and PsycINFO. The studies that met the inclusion criteria were further meta-analysed for effect sizes with a fixed effects approach weighted by sample sizes. The inclusion criteria were that the studies must involve a technology-based intervention, participants of the studies must be formally diagnosed with dyslexia in English, outcome measures used must include at least one measure of phonological skills in reading, and studies must utilise a pre-test-post-test experimental design and include means, standard deviations, and sample sizes. There were a total of four studies that met all criteria and these four studies employed six different technology-based interventions. All four studies had significant results showing that technology-based interventions positively influenced phonological skills. A grand total of 157 participants across these four studies returned a significant result for weighted pooled estimates of overall effect size on non-word decoding (a measure of phonological skills) to be $d = 0.56$ (ranging from $d = 0.17$ to $1.38$), which is a medium effect size of the technology-based intervention. Thus, technology-based interventions is an effective method of remediating phonological skills of children with dyslexia.

Keywords: technology, intervention, therapy, dyslexia, phonological, reading in the identification of pre-school children “at risk” of dyslexia, albeit with some adaptations for use in the local context.

* Correspondence to:
Dr Thomas Sim, DAS Academy Ltd, 73 Bukit Timah Road, #05-01 Rex House, Singapore 229832
Email: Thomas@dasacademy.edu.sg
Dyslexia is a type of specific learning difficulty identifiable as a developmental difficulty of language learning and cognition (US Department of Education, 2006). Rose (2009) identified the characteristic feature of dyslexia to be primarily in phonological awareness and the National Institute of Child Health and Human Development (2000) has identified that appropriate literacy programmes should include components of phonemic awareness and phonics training. Thus, phonological interventions have become part and parcel of the majority of dyslexia remediation programmes worldwide, including in Singapore. More than 14 years ago, it was estimated that there were more than 10,000 instructional software programmes on the market (Zhang, 2000), this number is an underestimate now. In 2008, a review of off-the-shelf software packages that assist in writing and spelling reviewed 22 such software (Peterson-Karlan, Hourcade, & Parette, 2008), suggesting that there is a growing awareness of the need for understanding how technology can help education, especially in the area of special educational needs.

Zhang (2000) used a qualitative case study approach to investigate the writing skills of five fifth-grade students with learning disabilities after using a writing programme (ROBO-Writer). The findings indicated that all five students improved their writing skills. One student was previously labelled as lazy and ill-behaved. After using ROBO-Writer, the student wrote a 350-word essay with very few spelling errors, the "...longest written work he had ever produced..." (p.473). Another student was previously extremely reluctant to engage in self-expression but after using the programme, she "...became increasing willing to talk to her mentor as a method of identifying exposition topics, (thus improving) socialization..." (p.473). This indicated that technology could also improve aspects other than just learning per se and indeed could be used as a tool to draw out shy and withdrawn students.

Hetzroni and Shrieber (2004) used a single-subject ABAB research design with three participants aged 12-13 years old with diagnosed learning disabilities and average IQ to examine the effectiveness of using a word processor (Microsoft Word 2000) to aid writing, spelling, and reading. The results showed that all three participants achieved fewer spelling mistakes (from baselines of 11-17% spelling mistakes to final results of 1-3%) and fewer reading errors (from baselines of 5-12% reading errors to final results of 0-1%). In addition, teachers assessing the written quality of the work produced indicated that the essays produced were organised better than before. This study shows that a simple use of a widely available software package can improve writing, reading, and spelling ability.

Cullen, Richards, and Lawless-Frank (2008) used a case study approach with a modified multiple baseline with seven participants aged 10-11 years old with diagnosed learning disabilities to examine the effectiveness of a talking word processor (Write: Outloud) and word prediction program (Co: Writer) on their writing. There were three phases in the study - baseline, Write: Outloud intervention alone (Phase 2), and Write: Outloud intervention with Co: Writer...
Results indicated that mean spelling accuracy increased from baseline of 87% to 95% (phase 2) to 96% (phase 3). Using a standard marking rubric (maximum mark of 20), participants average rubric score also increased from a baseline of 9.4 to 9.9 (phase 2) (both failing scores) to a score of 11.3 (a passing mark). Each of the participants essays were marked by three teachers and averaged to produce a final mean score. This study shows evidence for the effectiveness of technology in improving spelling and writing skills.

All the three studies reviewed (Cullen et al., 2008; Hetzroni & Shrieber, 2004; Zhang, 2000) used non-experimental research methods. Thus, even though the results point to the effectiveness of using technology-based interventions with children with special needs, these results are difficult to generalise. There is thus a need for more experimental research into this area so as to properly inform intervention methods for children with special needs and generalise the results. Several experimental studies have already been conducted, however, these studies were often conducted with participants diagnosed with dyslexia in languages other than English. Although it is known that dyslexia in different languages can manifest in different ways and hence differ in their responsiveness to phonological remediation (Rose, 2009), it is still useful to review if technology-based interventions helped these groups of participants.

Ecalle, Magnan, Bouchafa, and Gombert (2008) investigated whether computer based training in phonemic awareness can improve reading in children with dyslexia in French. A total of 26 children diagnosed with dyslexia and with IQ higher than 70 participated in the study. A pre-test, intervention, post-test design was used. The participants were randomly assigned into an experimental group ($n=13$) who underwent ortho-phonological phonemic audio-visual computer-based training or a control group ($n=13$) who underwent a computer-based training that only showed text on screen for the participants to read. At post-test, the experimental group was significantly better in pseudo word reading, regular word reading, and irregular word reading, indicating an increased ortho-phonological ability. This showed that computer based training using ortho-phonological units can improve reading ability, showing the effectiveness of assistive technology.

Kast, Meyer, Vogeli, Gross, and Jancke (2007) investigated the effectiveness of a multisensory training software programme (Dybuster) on 43 children with dyslexia in German and 37 age-matched controls using an experimental pre-post design. The group of dyslexia children was further divided into two groups (those with intervention in the first three months and those without intervention in the first three months). Children with dyslexia without software intervention in the first three months as well as the control group showed reading improvements of only 0 to 9%. Children with dyslexia with software intervention in the first three months had reading improvement of 19 to 35%. Due to obvious ethical issues, the Children with dyslexia without software intervention during the first three months underwent the intervention and
subsequently showed an improvement of 27 to 35%. Although one problem with the research was that the experimental and control groups differed significantly on baseline measures of IQ, the results do indicate that the reading software improved reading of people with dyslexia in German.

A meta-analysis of current experimental research with people with dyslexia in English would provide an amalgamation of available information and provide a review of the impact of using technology-based interventions and thus inform practitioners in this area. The purpose of this meta-analysis is thus to synthesise findings from experimental research studies gathered by a systematic review of the literature on the effectiveness of technology-based interventions on the phonological skills of children diagnosed with dyslexia in English.

**Method**

**Procedure**

The meta-analysis employed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) flowchart for search strategy (see Figure 1) and adhered to the APA Meta-Analysis Reporting Standards (MARS) (APA, 2011).

**Search Strategy for Study Identification**

A computer literature search up to December 2013 was performed in the following electronic databases: Academic Search Premier, Education Research Complete, ERIC, PsycARTICLES, & PsycINFO. The following key words were used:

- technology, computer, elearning, mobile learning, ICT
- intervention, instruction, remediation, therapy
- phonology, phonological skills, spelling
- dyslexia

No other sources for records were searched. The inclusion criteria were that the studies must involve a technology-based intervention, participants of the studies must be formally diagnosed with dyslexia or at risk for reading difficulties in English, outcome measures used must include at least one measure of phonological skills, and studies must utilise a pre-test-post-test experimental design and include means, standard deviations, and sample sizes.

**Effect Size Analyses**

An effect size is a measure of the strength of a phenomenon (Kelley & Preacher, 2012). The larger the effect size of an intervention, the more effective that intervention. Individual effect sizes for each eligible study with reported means and standard deviations were calculated based on Cohen’s $d$, defined as the difference between two means (mean change) divided by the standard deviation for the control group. A Cohen’s $d$ score from 0.2 to 0.49 is considered a small effect, a score from 0.5 to 0.79 is considered a medium effect, and a score of 0.8 and above is considered a large effect.

The weighted mean differences method was chosen to obtain the pooled estimates of overall effect sizes for common outcome measures (Wolf, 1986).
Weights were based proportionately on the number of participants in each study (Hedges & Olkin, 1985).

Results

Thirty five articles were initially identified via database searching. There were 14 articles after removing duplicates. A total of 10 articles were excluded, one as it did not have pre-test post-test measures, another as it used morphological outcome measures, and eight others conducted in languages other than English. Thus, a total of four articles were included in the qualitative synthesis (see summary of studies in Table 1).
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>N</th>
<th>Technology Used</th>
<th>Intervention Length</th>
<th>Grades /Ages</th>
<th>Outcome Measures (refer to qualitative synthesis for more information)</th>
<th>Effect Sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blythe (2006)</td>
<td>10</td>
<td>Phonics Alive2: The Sound Blender</td>
<td>10 weeks (=11.5 hours)</td>
<td>Grades 1 to 6</td>
<td>1. Reading Composite Index (2+3+4) 2. Word Reading 3. Reading Comprehension 4. Pseudoword Decoding</td>
<td>1.0.62 2.0.27 3.0.40 4.0.58</td>
</tr>
<tr>
<td>28</td>
<td>Automaticity Programme</td>
<td>50 minutes, 3 times a week over 17 weeks (=42.5 hours)</td>
<td></td>
<td>1. Word Recognition 2. Spelling 3. Comprehension 4. Phonological elision 5. Rapid letter naming 6. Non-word reading 7. Sight word reading</td>
<td>1.0.10 2.0.09 3.0.24 4.0.26 5.0.19 6.0.17 7.0.18</td>
<td></td>
</tr>
<tr>
<td>Torgesen et al. (2010)</td>
<td>34</td>
<td>Read Write and Type (RWT)</td>
<td>4 50-minute sessions per week over a school year (=80.4 hours, with 44.6 hours on computers)</td>
<td>Grade 1</td>
<td>1. Word Identification 2. Word Efficiency 3. Word attack 4. Non-word efficiency 5. Blending Words 6. Phonological elision 7. Segmenting Words 8. Rapid Naming Digits 9. Rapid Naming Letters</td>
<td>1.0.41 2.0.22 3.0.59 4.0.26 5.0.70 6.0.28 7.0.64 8.0.67 9.0.33</td>
</tr>
<tr>
<td>35</td>
<td>Lindamood Phoneme Sequencing Program for Reading, Spelling, and Speech (LIPS)</td>
<td>4 50-minute sessions per week over a school year (=84.3 hours, with 35.6 hours on computers)</td>
<td></td>
<td></td>
<td>1. Word Identification 2. Word Efficiency 3. Word attack 4. Non-word efficiency 5. Blending Words 6. Phonological elision 7. Segmenting Words 8. Rapid Naming Digits 9. Rapid Naming Letters</td>
<td>1.0.64 2.0.52 3.0.95 4.0.81 5.0.44 6.0.61 7.0.87 8.0.33 9.0.00</td>
</tr>
</tbody>
</table>
Qualitative Synthesis

Blythe’s (2006) pilot study on the effectiveness of computer-based phonological skills training was conducted with 20 primary school students diagnosed with dyslexia using a pre-test post-test control group research design. The study involved using a commercial off-the-shelf computer software called “Phonics Alive 2: The Sound Blender®”. This programme consists of 12 modules which build phoneme awareness skills, phoneme-grapheme correspondence, sound and letter blending, and processing speed. Each module takes an average of 15 minutes to complete. The 10 participants in the computer intervention group were instructed to repeat each module until a mastery level of 90% correct responses was achieved. Parent reports indicated that all children were compliant throughout the 10-week training period. Eight out of ten children completed all twelve modules, with the two youngest completing through to module 10. The other 10 participants were in the control group. The outcome measures were items from the Wechsler Individual Achievement Test, second edition (WIAT-II) (Wechsler, 2002). The Reading Comprehension Index provided a measure of general reading ability. It is produced by combining the standardised scores of each of the following reading subtests: The Word Reading subtest that provided a measure of sight word reading, with participants reading aloud from a graded word-list; the Reading Comprehension subtest that provided a measure of textual comprehension by reading narrative passages (either aloud or silently) then answering comprehension questions; and the Pseudoword Decoding subtest that provided a measure of the student’s ability to apply phonetic decoding skills by having the students read aloud from a list of graded nonsense words designed to mimic the phonetic structure of words in the English language. There were significant interactions between treatment group and time for the overall Reading Composite Index (RCI) \(F(1,18) = 29.08, p < .001\) with the treatment group showing a greater increase in RCI and all subjects, effects sizes ranged from small to medium, \(d = 0.27\) to 0.62 (for specific effect sizes, see Table 1). Although the study only had 20 participants: 10 in the treatment group and 10 in the control group, participants were randomly assigned and compliance to the intervention programme was monitored. In spite of the small number of participants, results were still significant. The results show that computer-based training was effective in improving phonological skills of children with dyslexia even with only approximately 11.5 hours of training over 10 weeks.

Gaab, Gabrieli, Deutsch, Tallal, and Temple (2007) conducted an fMRI study investigating neural correlates of rapid auditory processing in children diagnosed with dyslexia but also included behavioural measures suitable for this analysis. Twenty-two children with dyslexia participated in a pre-test post-test control group research design with 23 matched controls. The technology-based intervention used was the Fast ForWord Language programme. The remediation consists of five 20-minute training sessions per day, five days a
Meta-analysis of Technology-based Interventions

week over a period of eight weeks for a total of approximately 13.3 hours. The outcome measures included Word, Non-Word decoding, and Passage Comprehension (subtests from the Woodcock-Johnson Reading Mastery Test Revised (WJ-RMT-R)); Listening Comprehension (subtest from the Woodcock-Johnson-Revised (WJ-R) Test of Achievement); Phonological Awareness, Phonological Memory, and Rapid Naming (subtests from the Comprehensive Test of Phonological Processing (CTOPP)). Results showed significant improvements in all these seven measures with effect sizes from medium to large, $d = 0.44$ to $1.38$ (for specific effect sizes, see Table 1). Of particular interest is the result that the effect sizes for the improvement in word reading and non-word decoding was over 1.0, which is a huge effect. This research was very controlled, had good participant numbers, and add to the evidence that computer based training is effective in improving phonological skills of children with dyslexia.

Higgins and Raskind (2004) investigated the effectiveness of Speech Recognition Based Programmes (SRBP) (auditory and visual representations of words) and Automaticity Programmes (AP) (Multisensory with more interactivity) on a total of 42 children with learning disabilities. The 42 children were divided into two groups, 28 children on assistive technology and 16 children on classroom teaching and the research design was an experimental pre-post control group intervention design. To control order effects, part of the 28 students did the AP intervention then SRBP, and part did SRBP then AP. The outcome measures were: Word Recognition, Spelling (subtests from the Wide Range Achievement Test-3); Comprehension (subtest from the Formal Reading Inventory); Phonological elision. Rapid letter naming (subtests from CTOPP); Non-word reading, and Sight word reading (subtests from the Test of Word Reading Efficiency (TOWRE). The SRBP group improved significantly more on word recognition, comprehension, phonological elision, and non-word reading, with effect sizes ranging from small to medium, $d = 0.17$ to $0.35$ (for specific effect sizes, see Table 1). The AP group improved significantly more on all the above measures and spelling, rapid letter naming, and sight word reading, however, effect sizes were small, $d = 0.09$ to $0.26$ (for specific effect sizes, see Table 1). This is despite the fact that the intervention for SRBP was for only 14.2 hours, far lesser than on AP that was for 42.5 hours. The small effect sizes for the AP intervention show that more hours of training does not necessary translate to a greater effect. The researchers also noted that one major limitation of their study was that their participants ranged in ages from 8 to 18, this wide range could have affected the results as a number of reading intervention studies report different treatment effects across ages (e.g. Wise, Ring, & Olson, 1999, 2000 cited in Higgins & Raskind, 2004).

Torgesen, Wagner, Rashotte, Herron, and Lindamood (2010) investigated the effectiveness of two computer-assisted instructional programmes on children at risk for dyslexia (but too young for formal diagnosis of dyslexia). The study employed a pre-test post-test experimental control group research.
design. Thirty four children went through the Read Write Type (RWT) programme (based on the premise of directly teaching students the spellings of phonemes) for 80.4 hours that had a computer-based training component of 44.6 hours. Thirty five children went through the Lindamood Phoneme Sequencing Program for Reading, Spelling, and Speech (LIPS) programme (based on the idea of early knowledge of the oral motor awareness of phonemes for decoding and encoding) for 84.3 hours with a computer-based training component of 35.6 hours. As noted, both programmes were not fully computer based programmes and were blended programmes that included one-to-one teacher instruction. The outcome measures were Phonological elision, Blending Words, Segmenting Words, Rapid Naming (subtests from CTOPP); Word Identification, Word Efficiency, Word attack (subtests from WJRM'T-R); and Non-word efficiency (subtest from TOWRE). The effect sizes of the RWT intervention ranged from $d = 0.28$ to $0.70$ (for specific effect sizes, see Table 1). The effect sizes of the LIPS programme ranged from no effect to large effects with effect sizes ranging from $d = 0.00$ to $0.95$ (for specific effect sizes, see Table 1). However, these results must be qualified, the researchers noted that the computer based instruction in this study was a supplement rather than a replacement for teacher-led instruction. There should have been an additional intervention group with just computer based intervention without any teacher-led instruction. In spite of this, the results do show that computer based interventions (whether solely or as a supplement) positively impacts the phonological skills of children with dyslexia. These results are to be interpreted in line with follow up tests after two years that indicated that although results were still significant, effects sizes for outcome measures have dropped to $d = 0.33$ to $0.43$, which were still medium effect sizes even after two years. This suggested that intervention outcomes could fade out over time.

The four studies used a variety of computer-based interventions. The computer programmes used by Higgins and Raskind (2004) were fairly primitive. The SRBP programme was based on a Microsoft PowerPoint 1997 platform with recorded speech and suffered from recognition errors common in all speech recognition type software. The AP programme employed a speech synthesiser (a 1998 model) that had a fairly artificial robotic speech. On the other hand, the other three studies used commercial off-the-shelf computer programmes that had high levels of interactivity and good graphics. It was thus not surprising that the Higgins and Raskind (2004) study had the lowest effect sizes among the four studies.

**Quantitative Analysis**

There was only one common outcome measure among the four studies, Non-word Decoding (called Pseudoword Decoding in Blythe (2006), Non-word reading in Higgins & Raskind (2004), and Non-Word efficiency in Torgesen et al. (2009)). Using this common outcome measure as a basis of comparison among the four studies, all four studies showed significant results for the technology-based interventions on the outcome measure of Non-word decoding. A grand total of 157 participants with
Meta-analysis of Technology-based Interventions

dyslexia had significant improvements in phonological skills after some form of technology-based intervention with effect sizes for non-word decoding ranging from $d = 0.17$ to $d = 1.38$ (see Table 2). Using the weighted mean differences method proportionately based on sample sizes, the pooled estimate of overall effect size for non-word decoding is $d = 0.56$, which is a medium effect size.

This provides evidence for the positive impact of technology-based interventions on the phonological skills of children with dyslexia. If we remove the Torgesen et al. (2010) study from the pooled estimate of overall effect size and account for technology-based intervention without the confounded issue of blended instruction, the results still showed that a total of 88 participants with dyslexia had significant improvements in phonological skills after technology-based intervention alone with a pooled estimate of overall effect size of $d = 0.58$, which is similar to the overall effect size of $d = 0.56$. This gives strong evidence for inclusion of technology-based intervention in dyslexia remediation.

Discussion and Conclusion

Four studies were selected in this meta-analysis. All four studies employed a pre-test post-test experimental research design with control groups. Under both qualitative synthesis and quantitative analysis, all four studies showed significant positive results in using technology-based interventions to improve the phonological skills of children with dyslexia. There were no

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>N</th>
<th>Technology Used</th>
<th>Outcome Measures</th>
<th>Effect Sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaab et al. (2007)</td>
<td>22</td>
<td>Fast ForWord</td>
<td>Non-Word decoding</td>
<td>1.38</td>
</tr>
<tr>
<td>Torgesen et al. (2010)</td>
<td>35</td>
<td>LIPS</td>
<td>Non-word efficiency</td>
<td>0.81</td>
</tr>
<tr>
<td>Blythe (2006)</td>
<td>10</td>
<td>Phonics Alive2: The Sound Blender</td>
<td>Pseudoword Decoding</td>
<td>0.58</td>
</tr>
<tr>
<td>Higgins &amp; Raskind (2004)</td>
<td>28</td>
<td>Speech Recognition Programme</td>
<td>Non-word Reading</td>
<td>0.35</td>
</tr>
<tr>
<td>Torgesen et al. (2010)</td>
<td>34</td>
<td>RWT</td>
<td>Non-word efficiency</td>
<td>0.26</td>
</tr>
<tr>
<td>Higgins &amp; Raskind (2004)</td>
<td>28</td>
<td>Automaticity Programme</td>
<td>Non-word reading</td>
<td>0.17</td>
</tr>
</tbody>
</table>
major methodological concerns with any of the four studies. However, Torgensen et al. (2010) study used a blended technology-based approach that included a teacher-led component, an example of good practice that may nevertheless have confounded the effects of a purely technology-based intervention.

The technology-based intervention that had the greatest effect was Fast ForWord (Gaab et al., 2007). It is worth noting that this intervention only occurred over five 20-minute training sessions per day, five days a week over a period of eight weeks for a total of approximately 13.3 hours. This seems to support the conclusion that repeated practice is one of the most useful approaches. The two interventions that had the smallest effect was Higgins & Raskind (2004) use of the Automaticity Programme (Cohen’s $d = 0.17$) that had an intervention period of 42.5 hours and Torgesen et al. (2010) use of RWT (Cohen’s $d = 0.26$) that had an intervention period of 80.4 hours (44.6 hours on computers). Thus, the amount of time spent on the intervention was not directly related to how effective it was. This conclusion is limited by the small number of studies that was examined, more research would have to be conducted to determine if the length of intervention was correlated to outcome.

All four studies used a measure of Non-word decoding as one of the outcome measures. It is thus suggested that Non-word decoding can be seen as a de facto standard for measuring phonological skills, especially for dyslexia remediation. Also, the presence of this common outcome measure allowed for the amalgamation of the results and a combined weighted overall effect size.

Based on consolidation of the evidence from these four studies that include the use of six different technology-based interventions, there is evidence to support the use of such interventions. However, this conclusion should be taken with caution as only four articles up to December 2013 met the criteria for inclusion (which also shows the lack of research in this area). Overall, more research with larger sample sizes should be conducted to better understand the effect of technology-based intervention and future research should include an outcome measure of non-word decoding to allow for continued future consolidation of research knowledge. Technology-based intervention should be considered an important element in dyslexia remediation of phonological skills.

References


