Asia Pacific Journal of Developmental Differences
Vol. 10, No. 2, July 2023 pp. 187—210
DOI: 10.3850/S2345734123000880

Response and Non-response to Intervention for Reading Difficulties: What Role do Cognitive Correlates Play?

Beth A O'Brien[1*], Tan Chee Soon[2] and Malikka Habib[1]

1. Centre for Research in Child Development , National Institute of Education, Nanyang Technological University, Singapore
2. Psychology and Child & Human Development, National Institute of Education, Nanyang Technological University, Singapore

Abstract

Within the field of learning disabilities many intervention studies find that treatment resisters remain despite gains in our understanding of best practices and effective treatment for reading development and disability. In this study we examine good vs. poor responders in an intervention study with 147 early primary grade students in a learning support programme. Students were assessed for reading accuracy and fluency after completion of a tablet-based reading intervention, and classified as responders vs. non-responders based on criterion-referenced scores for word reading accuracy and fluency. Differences between the two groups were evaluated for the rate of growth on literacy measures over the intervention phase, their cognitive attributes at pre-intervention, and their in-lesson performance on the tablet-based intervention activities. Findings show the responder group had initial superior performance on decoding and spelling measures, as well as broad abilities related to nonverbal reasoning, working memory, phonological awareness and rapid symbol naming. Further, the gap in performance on decoding and spelling measures increased over time, with the non-responder group showing some improvement in these skills, but to a significantly smaller degree than the responder group. Different approaches to phonics intervention in the study resulted in the same proportion of non-responders. Further, children's confusions with specific sound-symbol associations over the course of the interventions suggest potential challenges that teachers may highlight.

Keywords: reading disorder, learning disorder, treatment non-responders

* Correspondence to:
Dr Beth O'Brien, Deputy Centre Director, Centre for Research in Child Development, OER Centre for Research in Child Development Principal Research Scientist, National Institute of Education - Office of Education Research, Nanyang Technological University. Email: beth.obrien@nie.edu.sg

With population rates of about 5-15%, learning disorders related to reading account for large numbers of school-going children and a high degree of educational resources for supplemental support services (Wagner et al., 2021). As a result of this pressing demand, considerable research has been conducted over the past decades focusing on the characteristics, etiology and potential learning mechanisms involved in reading disorders (Peterson & Pennington, 2015; Snowling & Hulme, 2021). Although there has been much progress in reading science, contributing to improved approaches to intervention (National Reading Panel, 2000; Abadiano & Turner, 2003), a persistent finding across studies is the existence of treatment resisters, or non-responders. It has been estimated that as many as 30% of students at risk for reading difficulties may not benefit from generally early literacy interventions (Blachman, 1994; Juel, 1994; Torgesen et al., 1992). Students who do not respond to multiple tiers of research-based interventions require further specialized support (Austin et al., 2017).

This study focuses on groups of responders and non-responders from a previous investigation comparing phonics interventions targeting different grainsizes of phonology (phonemes vs. rimes vs. whole words) (O'Brien, Habib, & Onnis, 2019). While that study reported results in terms of the group performance across treatments, here we zero in on the non-responding children across the groups. We aim to find out what characterizes these children, whether certain treatment approaches may be more beneficial for them, and what their online performance looks like during the phonics intervention activities. The treatments utilized a technology-based platform, allowing a unique closer examination of their performance during the intervention phase, in addition to pre- and post-intervention assessments.

Characteristics of Non-responders

Most students can learn to read adequately from quality classroom reading instruction (Torgesen, 2000) or from supplemental small-group or individual reading interventions (Austin et al., 2017; Denton et al., 2006; Denton et al., 2013). However, there is a group of students with intractable reading difficulties. Children who are difficult to remediate scored below typical readers and children who were readily remediated on tests which evaluated phonological skills, suggesting that their persistent reading problems may be caused by phonological deficits (Vellutino et al., 1996).

Researchers have been interested in studying the characteristics of students who do not respond to effective literacy because identifying these characteristics could improve screening measures and selection of the most appropriate students for intensive interventions (Al Otaiba & Fuchs, 2002). Results of a longitudinal study that spanned over two years suggest that non-responsiveness to generally effective interventions may be an indicator of a long-term reading disability (Al Otaiba and Fuchs, 2006). This study found that almost 92% of the students who were non-responsive to reading interventions in kindergarten continued to be non-responsive in first grade. All except one of the non-responders required an Individualized Education Plan (IEP) with reading goals when they were in third grade.

A review of 23 studies on children who did not benefit from early literacy interventions reported that poor phonological awareness characterized the majority of unresponsive students (Al Otaiba & Fuchs, 2002). Of the 21 studies reviewed that explored phonological awareness, 70% of them found that phonological awareness was related with non-responders. The review also reported other child characteristics which were associated with non-responsiveness to intervention, including phonological memory, rapid naming, general intelligence (IQ), attention, behaviour problems, and orthographic processing. Nelson, Brenner and Gonzalez (2003) extended the review by including an additional seven studies to those reviewed by Al Otaiba and Fuchs in their meta-analytic review. Based on 30 studies, they found that, in order of magnitude, rapid naming, phonological awareness, problem behaviour, alphabetic principle, memory and IQ appeared to predict responsiveness to early literacy interventions. Consistent with the earlier review (Al Otaiba & Fuchs, 2002), Nelson and colleagues found that rapid naming and phonological awareness were more strongly related to responsiveness to interventions than alphabetic principle, memory and IQ. Further converging evidence was provided by a study by Fletcher et al. (2011) which also found that phonological awareness was most related to inadequate response to reading intervention. Rapid naming, syntactic comprehension, working memory and listening comprehension made smaller contributions to the group comparison between adequate and inadequate responders. In a study on the neurocognitive predictors of response to intervention to GraphoGame, a computer- assisted reading intervention, phonological awareness was found to be the strongest predictor of response to the intervention (Wilson et al., 2021). Given the important role phonological awareness appears to play in treatment response, we consider whether it also may be a vital prerequisite for intensive practice with phonics activities in the current study.

Treatment Approaches Using Technology Assisted Reading Instruction

Students with reading difficulties require intensive and explicit instruction in phonological awareness and the alphabetic principle to develop word reading skills (Vellutino, 1991). These students benefit from one-to-one tutoring that can adapt to their individual needs (Cheung & Slavin, 2013). However, it is challenging for schools to provide such individualized practice due to a lack of instructional resources and adequately trained teachers (Torgesen et al., 2010). It has been proposed that computer aided instruction could be part of the solution to helping these students because of its ability to provide highly specialized instruction (Anderson-Inman & Horney, 2007; Stetter & Hughes, 2010) and to allow individualized repetition (Saine et al., 2011).

The findings of studies that investigated the efficacy of computer-assisted reading interventions have been mixed. A large-scale evaluation of five reading software products found no statistically significant improvement in reading scores from computer-based instruction (Dynarski et al., 2007). Comprehensive models that use computer-assisted instruction did not produce significant positive effects (Cheung & Slavin, 2013). However, when computer-assisted instruction was provided as a supplement, instead of a replacement for teacher-led instruction, students performed significantly better in phonemic awareness, phonemic decoding, reading accuracy, rapid automatic naming and reading comprehension than control-group students (Torgesen et al., 2010). According to Torgesen et al. (2010), computer-assisted instruction that was tightly coordinated to extend knowledge and skills that had initially been taught by teachers is an effective way to provide reading instruction to young readers at risk for learning disabilities.

GraphoGame (later known as GraphoLearn), a computer-assisted reading intervention which uses systematic phonics to train the connections between spoken and written language, was first devised in the Finnish language. It has now been adapted for over 20 languages (Ahmed et al., 2020) but a recent meta-analysis of 19 GraphoGame studies in a range of languages measuring GraphoGame's impact on word-reading outcomes concluded that it was only effective in certain educational contexts, with effect sizes for word reading ranging from - 1.07 to 1.58 (McTigue et al., 2019). A study with Finnish-speaking beginning readers showed that those who received computer-assisted reading intervention using the GraphoGame program made significant gains in letter knowledge, decoding, accuracy, reading fluency and spelling (Saine et al., 2011). This could be because it is programmed to adapt task difficulty according to student performance. Also, students may find the game-like learning environment provided by GraphoGame to be a fun way to learn and may be more willing to practice (Saine et al., 2011). GraphoGame was subsequently extended to German and different versions were used with Austrian second and fourth graders and prereading students in Switzerland (Brem et al., 2010; Huemer et al., 2008). The Austrian students showed improvement in reading accuracy (Huemer et al., 2008) and the students in Switzerland made significant improvement in letter knowledge (Brem et al., 2010). An English version of GraphoGame was subsequently developed and its efficacy was demonstrated as a supplementary program for second graders who were identified as poor readers in the UK (Kyle et al., 2013). The experimental group made significant improvement in reading, spelling and phonological skills in comparison to the control group. Subsequently, a large-scale randomized controlled trial study of GraphoGame Rime was conducted in the UK (Worth et al., 2018). The study did not yield evidence that the intervention resulted in improved reading outcomes over business-as-usual instruction. However, the sample was reanalysed by Ahmed et al. (2020) for the "top half " of GraphoGame Rime players only (designated adequate responders). When this group of students were compared with the full sample of control students, significant gains were made on a nonword measure by those who received the GraphoGame intervention compared to those in the business-as-usual group.

GraphoGame has also been found to be effective in improving foundational literacy among English Language Learners (ELLs). When the English version of GraphoGame was used with first and second grade English ELLs, in India (Patel et al., 2021), students in the experimental group showed significantly greater and faster development on in-game assessments of letter-sound knowledge, rime unit recognition, and word recognition compared to those in the control group.

While educational computer games can greatly increase the amount of exposure and practice that struggling learners have with the content to be learned (Aravena et al., 2013) (in this case grapheme-to-phoneme correspondences), the use of such technology-based interventions have not been examined from the perspective of treatment responders and non-responders. The current study aims to fill this gap, by conducting an examination of students who responded to a pull-out learning support program coupled with experimental technology-based applications that supplemented the program. Contrasting these responders to students who did not respond will provide valuable information about how to predict which children may require different types or degrees of intervention. Also, findings may help educators to develop interventions that are targeted to specific needs of individual students.

Thus, the objectives of the current study are:

1. To compare the rate of growth on literacy measures between responders and non-responders
2. To identify the role played by cognitive correlates in student response to reading interventions
3. To investigate whether a specific intervention approach yields more responders
4. To identify problematic areas of learning revealed by performance on the technology-based activities

METHOD

Participants

The participants include 148 students (M = 6.75 years old) in Singapore public schools (136 in primary grade 1 and 12 in primary grade 2) who were identified as at risk for reading difficulties at primary school entry and who were enrolled in a learning support programme within their school. They further received supplementary experimental interventions in a randomized-controlled design, contrasting three approaches, described below and in O'Brien, Habib, & Onnis, (2019). Random assignment to treatment groups was conducted with matched sets of students based on baseline measures of reading ability, to ensure that the groups did not differ in overall ability.

Responder/Non-responder Classification

Subsequently for this study, participants were categorized as treatment responders and non-responders based on their pretest and post-test scores on two measures: (1) word reading accuracy (Woodcock Johnson Test of Achievement 3rd Edition, WJIII) and (2) word reading fluency (Test of Word Reading Efficiency, TOWRE-2). WJIII word reading scores were converted to grade equivalents based on the published manual. A non-responder was defined as a child whose pretest and post-test scores were below their grade level; a responder was defined as a child who scored below grade level at pretest, but who achieved a score at or above their grade level on the post-test. TOWRE-2 sight word reading scores were converted to standardized scores (z) based on locally developed norms (Chen et al., 2016). Children who scored 1 standard deviation (z=-1) or more below the mean for their age-group on both pretest and post-test assessments were defined as non-responders. Children who scored 1 standard deviation or more below the mean at pretest, but within 0.5 standard deviation of the mean (z=-0.5 or above) at post-test were defined as responders. The final classification was determined as such: Responders included children who were defined as a responder with either (1) or (2); if the child did not respond to either (1) or (2) then they were included in the non-responders group. From this process we found 77 responders and 66 non-responders in this study, 133 in grade 1, and 10 in grade 2. (71 responders, 45 non-responders according to reading accuracy criterion (1); and 33 responders, 76 non-responders according to reading fluency criterion (2)).

Measures

Baseline Assessments

Reading ability was assessed with the British Ability Scales – Third Edition (BAS-III, Elliot & Smith, 2012) word reading task. Words are presented in order of increasing difficulty, and test administration is discontinued after a given number of errors (8 out of 10).
Nonverbal cognitive ability was assessed with the Matrices and Quantitative reasoning subtests of the BAS-III (Elliot & Smith, 2012). Matrices involves choosing from a set of 4-6 pictures the correct one that completes a matrix of figures (e.g., shapes). Administration is discontinued after 3 consecutive trials with errors. Quantitative reasoning involves completing a sequence of numbers according to the relationship between a given pair of numbers shown. Test administration is discontinued after 3 consecutive trials with errors.

Working memory was assessed with the Memory for Digits subtest of the Comprehensive Test of Phonological Processing (CTOPP-2; Wagner et al., 2013). For this task, one repeats a series of digits in reverse order. The span of the series increases over trials, and administration is discontinued after three consecutive errors.

Implicit learning was assessed with a visual statistical learning task adapted from Arciuli &

Simpson (2011) and Raviv & Arnon (2018). First, during a training phase, a series of cartoon aliens are presented (via a tablet) with an embedded sequence of 4 triplets. During this phase, the cover task is to pay attention and press a button when an alien is repeated. After training, a 2 alternative-forced choice task is presented where one identifies which triplet of aliens was seen together previously, and total correct response out of 32 trials is recorded.

Vocabulary was assessed as a measure of language proficiency with a receptive task (BLAB; Rickard Liow & Sze, 2009) where the child chooses which image out of 4 matches the meaning of an aurally presented word. All 80 items are administered.

Rapid symbol naming was assessed with the CTOPP-2 RAN Letters subtest (Wagner et al., 2013). This task involves naming all the symbols presented on a printed card in 4 rows by 8 columns. The time to complete all items is recorded, along with errors.

Phonological awareness was assessed with the CTOPP-2 Elision subtest (Wagner et al., 2013). For this task, an aurally presented word is repeated, then the child is asked to say the word after deleting a designated part (e.g., syllable, onset, phoneme). Corrective feedback is given for the first ten items, and administration is discontinued after 3 consecutive errors.

Orthographic awareness was assessed with measures of orthographic choice and wordlikeness choice (Cunningham et al., 2001). Orthographic choice involves selecting the correctly spelled word in a pair of similar sounding word-pseudowords (e.g., rain-rane). All 23 trials are administered. The wordlikeness task involves selecting which of two pseudowords looks more like a real word (e.g., beff-ffeb). All 19 trials were administered.

Pretest and Post-test Assessments for Responder Classification

As noted above, classification into responder and non-responder groups was made with scores on the following tests.

Reading Accuracy was evaluated with the Woodcock-Johnson III word reading subtest (WJIII, Woodcock et al., 2007). The student reads aloud printed words, and administration is discontinued after six incorrect responses. The total number correct is tallied, and these scores are converted to grade equivalents based on the published norms (Woodcock et al., 2007).

Reading Fluency was assessed with the Test of Word Reading Efficiency (TOWRE-2, Torgesen et al., 2012) sight word subtest. After practice trials, the student read as many words as possible within a 45 sec time limit. The total number correct within the time limit was converted to z-scores based on a local normative sample (Chen et al., 2016).

Repeated Measures Literacy Assessments

Children were evaluated at four time points: before the intervention (pretest), at the midpoint during the intervention (after the first 7 weeks – mid-test), immediately after the intervention (post-test), and three months after the commencement of the intervention (follow-up).

Decoding accuracy was assessed with the WJIII (Woodcock et al., 2007) Word attack subtest. The student reads pseudowords aloud, and administration is discontinued after six incorrect responses. The total number correct per is tallied.

Decoding fluency was assessed with the TOWRE-2 (Torgesen et al., 2012) Phonemic decoding subtest. After the practice trial, the student reads as many pseudowords as possible within a 45 second time limit. The total number correct is scored within the time limit.

Spelling was assessed with a word list based on the BAS-III (Elliot & Smith, 2012), with additional items including morphologically and orthographically challenging words. Words were dictated, read in a sentence, and repeated. The first ten items were administered, then administration was discontinued after eight errors in the second block of trials. Total correctly spelled words were taken as the final score.

Experimental Intervention Conditions

Children took part in supplementary interventions during their learning support classes. The interventions involved gamelike activities rendered on iPad tablets, and children worked individually on their own iPad while listening to instructions and feedback through earphones. The intervention lessons involved 10 minutes of play 5 times per week over two 7 week periods during the second and third school terms (February to October). Matched sets of 3 children with regard to their baseline reading ability (BAS-III) were randomly assigned to one of the three intervention conditions: Phonics-phoneme, Phonics-rime, or Word-level approaches. Each condition focused attention in the activities to a different grainsize for learning sound-symbol relationships for reading English – from individual phonemes, to rime patterns, to whole words (refer to O'Brien, Habib, & Onnis, 2019).

The ***Phonics-phoneme condition*** involved 7 weeks of activities with Seeword reading (Seward et al., 2014; O'Brien, Seward, & Zhang, 2022a), and 7 weeks with GraphoLearn-Phoneme activities (Lyytinen et al., 2004; Saine et al., 2011).

The ***Phonics-rime condition*** involved 7 weeks of activities with Seeword reading (Seward et al., 2014; O'Brien, Seward, & Zhang, 2022a), and 7 weeks with GraphoLearn-Rime activities (Lyytinen et al., 2004; Saine et al., 2011).

The ***Word-level intervention*** emphasized decoding whole words through a different activity/game each day in the week (O'Brien & Yeo, 2022b). Children were asked to tap pictures that match a printed word in meaning, including single or multiple matching of words to pictures, and flash cards, where they read aloud and recorded their responses. In later weeks, activities included a choice task to match phrases to pictured meanings or building sentences from a word bank.

DATA ANALYSIS

The groups of identified responders and non-responders, across all the intervention conditions, were compared on different assessments to address the research questions. First, to examine group differences in growth in literacy across the intervention, between group repeated measures MANOVAs are performed on dependent measures of decoding accuracy, decoding fluency, and spelling across four timepoints (pretest, mid–treatment, post-test, and follow-up). Due to missing data of selected tests (e.g., due to child absences) each dependent measure was run in a separate MANOVA to maximize sample size. Child age was covaried in each analysis.

Second, to examine group differences in cognitive attributes, a between-group profile analysis is conducted with cognitive measures (nonverbal reasoning, working memory, implicit learning), language (vocabulary), and literacy related measures (phonological awareness, rapid naming, orthographic awareness). To identify the attribute or attributes that best predict responder/non-responder status, logistic regression analysis is also run with stepwise sets of the cognitive, language, and literacy predictors.

Third, to examine which intervention approach might yield more treatment responders, a chi-square test is conducted, comparing proportions of responders and non-responders across the three experimental intervention conditions (Phonics-phoneme, Phonics-rime, and Word-level conditions).

Finally, an examination of data collected during the in-lesson activities of the Grapholearn program for the phonics intervention conditions serves as a more fine-grained analysis of responder/non-responder differences. In this case, using confusion matrices of sound-symbol correspondences, we dissect particular points of confusion that may characterize the non-responder group.

RESULTS

Research Question 1

Data from the responder and non-responder groups (across all the intervention conditions) were submitted to 2 (between group) by 4 (repeated measures over time) mixed MANOVAs, with age covaried. First, the MANOVA for decoding accuracy (word attack) revealed a significant Group X Time interaction, $F_{(2.2, 282.7)} = 10.78$, $p < .001$, $\eta 2 = 0.079$. There were 55 non-responders and 74 responders with full data for this analysis. As shown in Figure 1a, the groups showed an initial difference in word attack scores at pre-test, and this gap widened over time, through to the follow-up point. Post-hoc pairwise comparisons revealed that the groups differed significantly at each time point (with Bonferroni adjustment, all p's < 0.001), with the responder group consistently outperforming the non-responder group. Further, while simple effects analysis showed that each group improved performance over time, the responder group showed a larger effect over time than the non-responder group ($F_{(3, 124)} = 82.77$, $p < 0.001$, $\eta 2 = 0.667$, $F_{(3, 124)} = 20.87$, $p < 0.001$, $\eta 2 = 0.336$, respectively).

The MANOVA for decoding fluency also revealed a Group X Time interaction that was significant, $F_{(2,231)} = 17.61$, $p < .001$, $\eta 2 = 0.13$. There were 49 non-responders and 70 responders with full data for this analysis. Similar to decoding accuracy, the groups showed an initial difference while the responder group showed stronger growth over time (Figure 1b). The responder group performed significantly better at each time point according to post-hoc pairwise comparisons (with Bonferroni adjustment, all p's < 0.001). Simple effects analysis also showed that responder group had a larger improvement over time than the non-responder group ($F_{(3, 114)} = 97.12$, $p < 0.001$, $\eta 2 = 0.719$, $F_{(3, 114)} = 16.92$, $p < 0.001$, $\eta 2 = 0.308$, respectively).

Finally, comparing the groups on spelling performance, the MANOVA for spelling scores showed a significant Group X Time interaction, $F_{(2.1, 227.4)} = 52.21$, $p < .001$, $\eta 2 = 0.320$. There were 47 non-responders and 67 responders with full data for this analysis. The initially small advantage of responders at pre-test was expanded over time, with bigger gains in spelling for the treatment responders (Figure 1c). Responders outperformed non-responders at each time point (all p's < 0.001 with Bonferroni adjustment). Simple effects analysis also revealed a larger improvement over time for the responder group compared to the non-responder group ($F_{(3, 109)} = 214.16$, $p < 0.001$, $\eta 2 = 0.855$, $F_{(3, 109)} = 27.74$, $p < 0.001$, $\eta 2 = 0.433$, respectively).

Thus, the responder group showed broad increased literacy gains across all three of the literacy measures, decoding accuracy, decoding fluency, and spelling, as compared to the non-responder group. In addition, it can be seen that this group started out at the pre -intervention test point, with better scores than the non-responder group. This initial gap between the groups increased over time for each literacy measure.

Research Question 2

Examining group differences on cognitive attributes between the responder and non-responder groups (54 non-responders, 62 responders across all the intervention conditions), a profile analysis showed that the groups differed in the pattern of their scores across these variables, $F_{(1, 114)} = 17.35$, $p < .001$. Univariate between-group effects

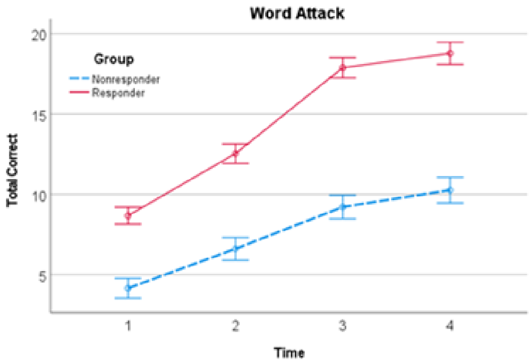

Figure 1a.

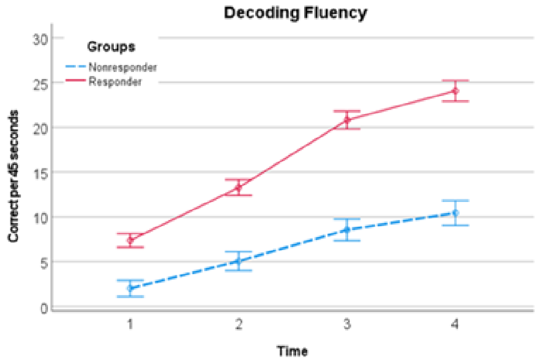

Figure 1b.

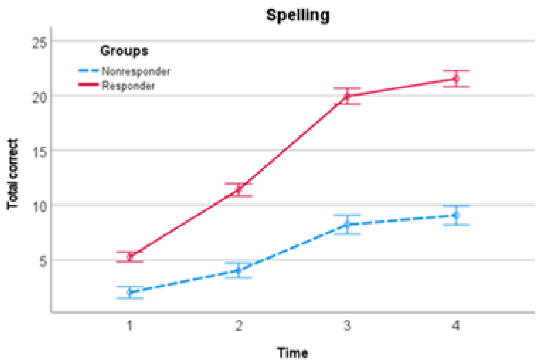

Figure 1c.

Figure 1. Performance by Responder group (solid line) and Non-responder group (dashed line) over time (1- pretest, 2- mid-test, 3- post-test, 4- follow-up) for decoding accuracy (a), decoding fluency (b), and spelling (c).

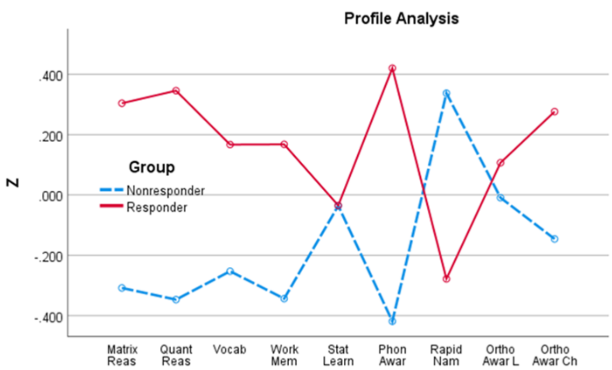

Figure 2. Mean performance (z-scores) per Responder (solid line) and Non-responder groups (dashed line) across cognitive, language and literacy measures. Nonverbal measures (BAS-III Matrix Reasoning, Quantitative Reasoning), vocabulary (BLAB), verbal memory (CTOPP-2 Memory for Digits), statistical learning, phonological awareness (CTOPP-2, Elision), rapid naming (CTOPP-2 RAN letters), and orthographic awareness (wordlikeness, orthographic choice).

were examined (using $\alpha = 0.0056$ with Bonferroni correction), and revealed that there were significant between group differences for cognitive assessments of nonverbal matrix reasoning, $F_{(1, 114)} = 11.96$, $p < .001$, nonverbal quantitative reasoning, $F_{(1, 114)} = 14.88$, $p < .001$, and working memory, $F_{(1, 114)} = 8.46$, $p = .004$, while the groups did not differ for statistical learning. In terms of the literacy measures, the groups also differed significantly in phonological awareness, $F_{(1, 114)} = 23.60$, $p < .001$, and rapid naming, $F_{(1, 114)} = 9.81$, $p = .002$. Group differences did not reach significance for orthographic choice and word-likeness measures, nor for English receptive vocabulary (*p*'s > .005). As shown in Figure 2, in each case of the group differences, the responder group showed stronger skills at baseline.

To further examine the relative contribution of these sets of competencies to differences in responder and non-responder groups, a logistic regression was conducted next, with children's responder/non-responder status as the dependent variable. The predictors were entered in blocks as follows: (1) nonverbal reasoning (quantitative and matrix reasoning), (2) verbal proficiency (English vocabulary), (3) memory (working memory, implicit memory-statistical learning), and (4) pre-literacy (phonological awareness, rapid naming, orthographic awareness – word likeness, orthographic choice). In the first block, adding nonverbal reasoning, improved the fit over the null model with intercept-only, $\chi 2(8) = 18.84$, $p = 0.016$. In this model the sensitivity (true positive rate of correctly identifying non-responder status) was 55.6 percent, while the specificity (true negative rate of correctly identifying responder status) was 75.8 percent. Adding verbal proficiency (vocabulary) in the second block did not improve model fit, $\chi 2(8) = 14.99$, $p = 0.059$, nor did the addition of memory (working memory, statistical learning) in the third

Table 1. Logistic regression predicting Responder/Non-responder group status

Final Model Entered Variables	step	B	S.E.	Wald	df	Sig.	Odds ratio	CI Lower	CI Upper
Matrix Reas	1	-0.010	0.064	0.022	1	0.882	0.990	0.873	1.124
Quant Reas	1	0.098	0.082	1.448	1	0.229	1.103	0.940	1.295
Vocab	2	0.011	0.024	0.203	1	0.652	1.011	0.964	1.060
Work Mem	3	0.023	0.049	0.217	1	0.641	1.023	0.930	1.125
Stat Learn	3	0.008	0.063	0.017	1	0.896	1.008	0.891	1.141
Phon Awar	4	0.112	0.058	3.796	1	0.051	1.119	0.999	1.252
Rapid Nam	4	-0.056	0.024	5.370	1	0.020	0.946	0.902	0.991
Ortho Awar L	4	0.025	0.120	0.042	1	0.837	1.025	0.810	1.297
Ortho Awar Ch	4	0.060	0.067	0.819	1	0.365	1.062	0.932	1.210

block, $\chi^2(8) = 6.59$, $p = 0.582$. With all variables in the first three blocks, sensitivity improved (63%) but specificity declined (72.6%). The final block, including pre-literacy variables (phonological awareness, rapid naming, and orthographic awareness) marginally improved the model fit, $\chi^2(8) = 15.39$, $p = 0.52$, but this model improved both the sensitivity rate (64.1%) and the specificity rate (80.6%). Of all the variables in the full model, only rapid naming contributed significantly to the group prediction of responders vs. non-responders ($p = .020$), while there was a trend for phonological awareness ($p = .051$). Parameters for the final model are shown in Table 1.

Thus, in spite of the groups differing in a range of skills, including the cognitive, verbal, and memory skills, it is the pre-literacy skills that best predicted response to intervention in this sample of children with reading disorders – and specifically, the two best predictors that are widely recognized in the literature on reading disorders for English.

Research Question 3

To see whether any of the three interventions in this RCT study resulted in better response to intervention, a Pearson chi-square was conducted comparing responder to non-responder ratios across the three intervention conditions. There was no significant difference in incidence of responders relative to non-responders across the three experimental intervention conditions, $\chi^2(2) = 0.061$, $p = 0.970$ (see Figure 3). This means that there was no particular advantage for any of the interventions over the others in terms of addressing non-responders more effectively.

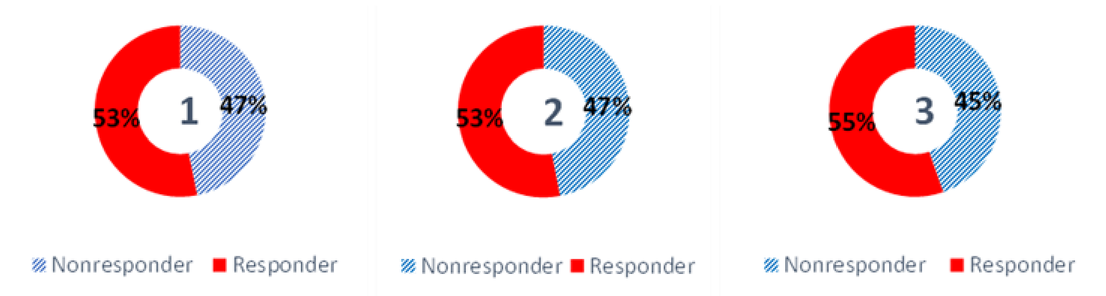

Figure 3. Percent Responders and Non-responders per intervention condition.
Condition 1 – Phonics-Phoneme, Condition 2 – Phonics-Rime, Condition 3 – Word-level interventions.
Responders (overall N= 77), Non-responders (overall N= 66).

Research Question 4

Examining differences of non-responders compared to responders at a finer timescale, the data from their online performance of the activities across the Grapholearn lessons was examined for the two phonics intervention groups. There were 26 responders and 23 non-responders in the Phonics-Rime condition, and 25 responders and 22 non-responders in the Phonics-Phoneme condition.

First, for each intervention group, letter-confusion matrices were generated over the completed set of activities that were derived across all the participants. These matrices are shown in Figures 4a and 4b, and display the level of confusions (percent) for each pairing of the correct letter representing an aurally presented phoneme ("target", y-axis) with "distractor letters" (x-axis). Shading represents the degree of confusion, with darker cells showing higher levels of confusion. Confusions exceeding 10% are marked with symbols.

Of interest, across both intervention conditions there are similar confusion types. As seen in Figures 4a and 4b, /ɛ/ ('e') is strongly confused with the letter 'a'.

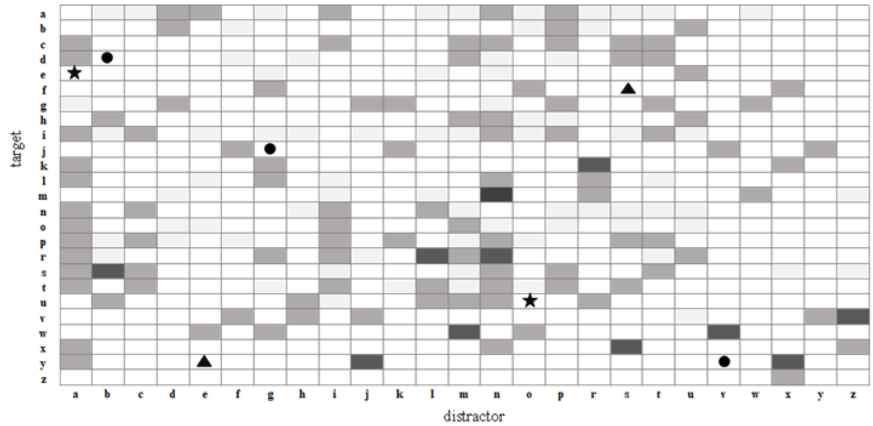

Figure 4a.

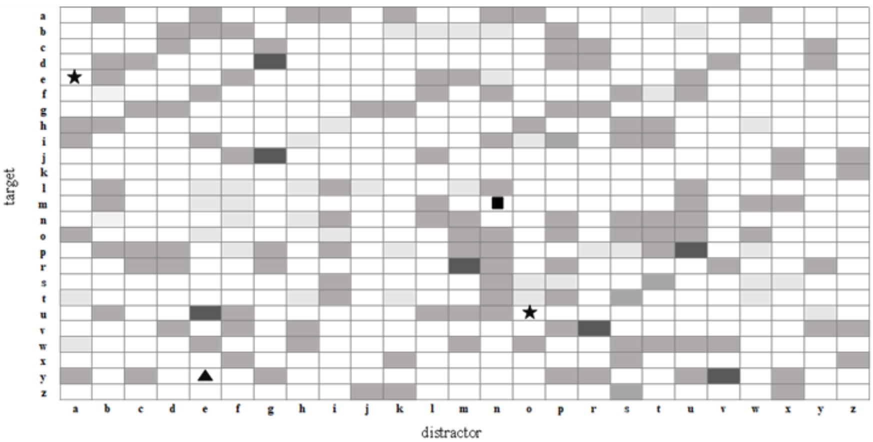

Figure 4b.

Figure 4. Confusion Matrix across all students in Phonics-rime condition (a) and in the Phonics-phoneme condition (b). Darker colours indicate more confusion. Confusions exceeding 10% are marked with symbols. There are two main categories of letters' similarity: phonetic similarity (marked with a "star") and visual similarity (marked with a "circle). It is also possible that both (marked with a "square") or neither categories are occurring (marked with a "triangle").

This can be concluded to be the most challenging sound-letter pair for the students, plausibly because both acoustic and visual similarity compromise building the connection. Also the vowel /ʊ/ ('u') is often mixed-up with the letter 'o'. The children also showed difficulties in mapping /dʒ/ with 'j' rather than the target 'g'. Other notable confusions involve the target for /f / ('f') confused with 's', and 'd'-'b' confusions in the Phonics-rime group and 'm'-'n' confusions in the Phonics-phoneme group.

Examining these high occurrence confusions further, we tallied the number of confusions made per responder/non-responder group within the Phonics intervention conditions (Rime, Phoneme). These breakdowns are shown in Figures 5 a, b and c for the letter confusions in student responses noted above. For e-a (Figure 5a) and u-o letter confusions (Figure 5b), trials involving matching isolated phonemes to single letters appear on the left bars, while trials involving matching the phoneme within rime patterns are shown in the middle bars.

Finally, errors related to incorrect vowel pairs are shown in the bars on the right. Most e-a confusions occurred on trials where the /ɛ/ is presented within rimes (middle bars) for both responders and non-responders, although a sizable number of these confusion errors also occurred when /ɛ/ was presented as an isolated phoneme to match to its letter (left bars).

For u-o confusions (Figure 5b) the isolated presentation of /ʊ/ presented the biggest challenge for the non-responders in both intervention conditions. These non-responders were less affected when /ʊ/ was presented in the context of a rime pattern (lesser confusions), and the non-responders in the Phonics-Rime condition were even less affected by the u-o confusion in this case (middle bar, black segment). The non-responder group was affected to a similar degree by the vowel pair representations of this phoneme (e.g., 'ou' as in a rime pattern), whereas the responder group was less so (right bars).

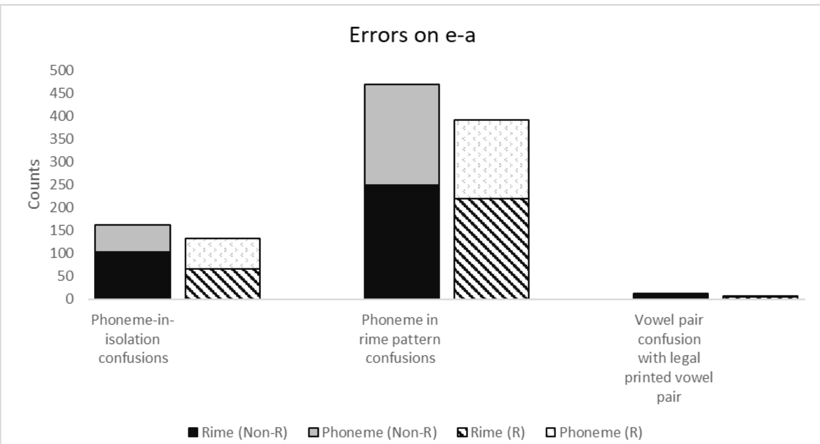

Figure 5a.

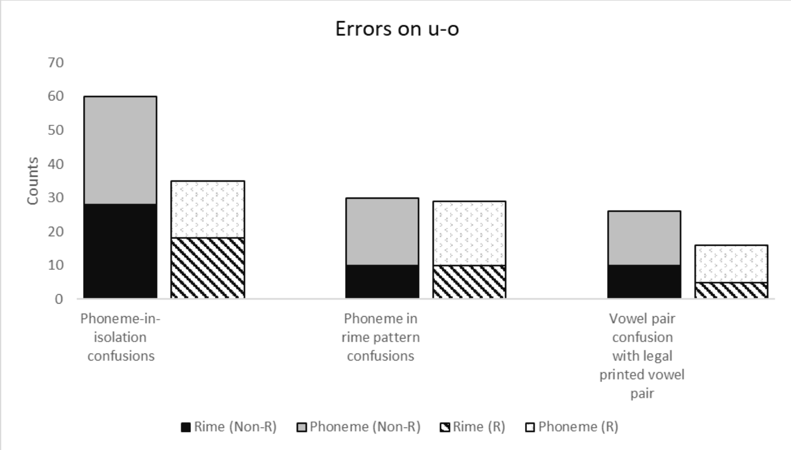

Figure 5b.

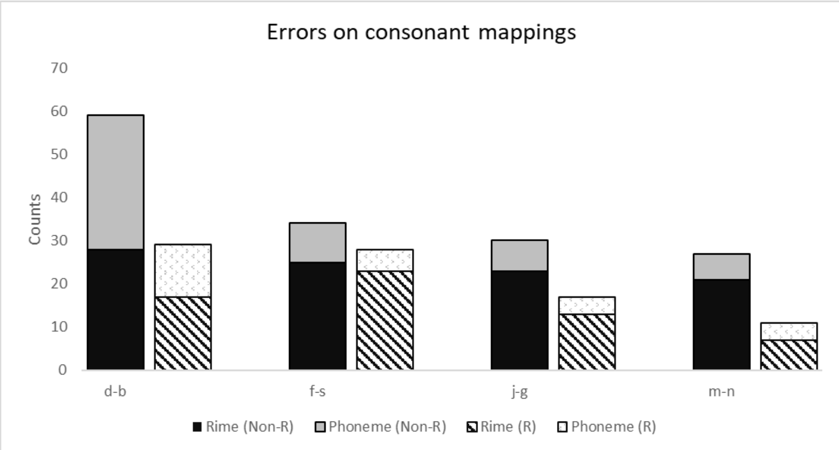

Figure 5c.

Figure 5. Total number of confusions by non-responders (solid bars) and responders (patterned bars), split by intervention conditions (Phonics-Rime – dark, Phonics-Phoneme – light). (a) Left bars show errors made on single sound-letter mapping, middle bars errors are on /u/ occurring within rime patterns (uck-ock, ug-og), and right bars are errors of target vowel pairs (ou) within rime patterns (oun -uon, ouse-uuse, ould-uold). (b) Left bars show errors made on single sound-letter mapping, middle bars errors are on /e/ occurring within rime patterns (et-at, ep-ap), and right bars are errors of target vowel pairs (ee) within rime patterns (eed-aad, een-aan). (c) Total number of confusions between consonant targets and distractors. Target sounds were presented in isolation.

Finally, for the confusions of consonants (Figure 5c), those in the Phonics-Rime condition had more difficulties in each case with mapping the correct letter to the aural sound when the phoneme was given in isolation. Compared to the responders, the non-responders made more of these errors, and this difference was most exaggerated in the d-b confusion (left bar), and secondly in the f-s confusion, followed by the m-n and j-g confusions (right bars).

DISCUSSION

This study examined the performance of students who received a pull-out learning support program that was supplemented with experimental technology-based applications. The interventions focused on either the phoneme, rime or whole word unit. It compared the rate of growth on literacy measures between good treatment responders and poor treatment responders and identified the difference on cognitive attributes between the two groups.

In the current study, the responder group initially started with a wide range of skills that were more advanced than the non-responder group. These included nonverbal reasoning, short term memory, phonological awareness and rapid naming. This group also started with relatively better decoding accuracy and fluency and spelling. These findings are consistent with earlier studies on responsiveness to intervention that reported that when compared with non-responders, responders were characterized by initial superior performance on rapid naming, phonological awareness, memory (Al Otaiba & Fuchs, 2002; Fletcher et al., 2011; Nelson et al., 2003). This is also in line with the study on GraphoGame by Wilson et al. (2021) which found that existing phoneme awareness skill is predictive of response to the computer-assisted reading intervention.

Another finding of this study was that the gap between the responders and the non-responders increased over time on measures on decoding accuracy and fluency and spelling. A review of various studies that investigated the effect of intervention on rate of growth in reading reported a mean effect size (ES) for reading measures ranging between .45 to .79 for treatment versus control conditions (Swanson & Hoskyn, 1998) and .82 to .95 for single-case design studies (Swanson & Sachse-Lee, 2000), in favour of children who were not at risk. In another synthesis, the magnitude of ES between responders and non-responders increased from 1.10 on the pretest to 1.28 on the post test on a word attack measure (Tran et al., 2011). These results provide support to the concept of a Matthew Effect which states that children with initially high levels of achievement should show a higher rate of progress in academic learning compared to children attending school with normal or low levels of initial achievement (Walberg & Tsai, 1983). The Matthew Effect was first applied to reading development by Stanovich (1986) who hypothesized that children who enter school with markedly underdeveloped phonological awareness have difficulty understanding the alphabetic principle and experience delays in breaking of the grapheme-to-phoneme code. Given the fact that pre-literacy skills (phonological awareness and rapid naming) skills are predictors of responsiveness to intervention in this study, it is important that children with reading difficulties be identified early and receive intervention with foundational skills even before they are exposed to formal reading instruction (Vellutino et al., 1996).

The different conditions of the supplemental technology-based interventions yielded the same proportion of responders, suggesting no advantage of focusing at a grainsize of phoneme, rime or whole word units in terms of reducing non-responder rates. Besides considering pre-to-postest assessments, in terms of responder status, we were also able to examine data downloaded during the Grapholearn interventions for students who played these activities (Phonics-phoneme, Phonics-rime groups). This revealed the types of confusions that students tended to make in terms of the phoneme-grapheme mappings they were learning. It appears some of these confusions may be sound-based (e.g., matching single phonemes-with-letters that represent similar sounds). Additional confusions appear to be more print-based (such as orthographic errors related to selecting illegal vowel pairs, e.g., oun-uon, ouse-uuse, ould-uold, eed-aad, een-aan).

The difficulties faced by students in both Phonics-phoneme and Phonics-rime conditions in confusing the vowel 'a' with 'e' is not surprising as the conflation of /e/ and /æ/ by speakers of Singapore English has been reported by a study by Deterding (2003). The lack of distinction between the vowel sounds /ɒ/ and /ʌ/ that was observed in both groups could be because these two vowels are often produced with almost the same quality and are differed mainly in terms of length (Deterding, 1997). In Singapore English, the lack of distinction between the long and short vowel pairs has been well-documented (Bao, 1998; Brown, 1988; Deterding & Poedjosoedarmo, 1998; Low & Brown, 2005). Another noteworthy confusion between commonly mixed letters 'f' and 's' may be related to letter f sometimes being pronounced phonetically as /fer/ in the Singapore context.

Further, the confusion between consonant letters (d-b, m-n, j-g) among students in the study provides evidence to support the letter-confusability hypothesis, which states that if the letter is visually confusable, students may have more difficulty identifying the lower-case letter correctly (Huang & Invernizzi, 2014). Some researchers have suggested that visual confusability of the letter is a major factor determining children's ability to recognize lower case letter successfully (Treiman & Kessler, 2003; Treiman et al., 2006).

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

It is important to consider several limitations in this study. Firstly, with regard to intervention intensity, the total duration of the intervention was limited. Each student spent less than 12 hours in total on the games (10 minutes of play 5 times per week over two 7-week periods). A study with a longer duration that allows students to complete all 25 streams in the game would increase the likelihood that the students would have learned the phonics "rules" of English, and in this way would more accurately assess the predictors of response to the intervention. In addition, the interventions in this study were given as a supplement to the pull-out reading instruction in the school, following earlier findings of successful computer-assisted instruction provided as a supplement to teacher-led instruction (e.g., Torgesen et al., 2010). However, the role of teachers in computer-assisted intervention needs further investigation. The students in this study played the tablet-based games relatively independently with minimal support from their teachers. Yet studies with high adult interaction produced an average positive effect size (g = 0.48)

(McTigue et al., 2019), suggesting higher levels of active engagement may be needed even with supplementary game-based learning. Indeed, one challenge with the adaptive Grapholearn game was that some students took more attempts to master the different levels. Calibrating task difficulty and cognitive load may be necessary to maintain student engagement, and teachers likely play a more effective role in this than the computer algorithm. Future research could consider ways teachers can monitor the data of their students' progress and how teachers would use this data to adapt an individual student's learning plan on such platforms as Grapholearn. Further, while computer-administered instruction allows children to recognize items presented aurally, there is little feedback to students regarding their pronunciation of the letter sounds or words. Teachers would play a very important role in providing this feedback.

Secondly, as a re-examination of treatment response to different intervention conditions, the treatment by responder-status subgroups here were too small to test for interaction effects. There did not appear to be any pattern suggesting that the confusion errors by Phonics-phoneme and Phonics-rime groups differ by responder status, as it was noted in the previous study with these children (O'Brien, Habib, & Onnis, 2019) that treatment group effects for confusion errors might be moderated by individual learner differences. Here there was only a main effect, whereby both responders and non-responders appear to show that the Phonics-rime intervention accounted for more of the confusion errors overall (the dark segments in the bar plots, Figures 5a-c), with a few exceptions where the Phonics-phoneme intervention showed equal or more error rates (u-o, d-b). Thus, further study with larger subsamples is needed to conclusively understand the relation of individual differences and treatment approach to confusion types. The current study has also shown that letter confusion may be dependent on the context as the errors made by students may be influenced by oral language used around them. Hence, it would be necessary for further research to investigate how a computer-assisted reading intervention can be adapted to different linguistic contexts to enhance its effectiveness.

IMPLICATIONS

This study found that vocabulary was not a predictor of whether a student would be able to progress through the Grapholearn game, consistent with the finding by Wilson et al. (2021). This suggests that students with limited vocabulary can benefit from this type of game-based intervention. However, phonological awareness and rapid naming were the most important predictors of response to the intervention. This reinforces the importance of these two early predictors of reading success, and also suggests that initial work on foundational phonological skills may be required to prepare children before game-based learning.

This study also points to an important role for teachers in the administration of supplemental, game-based intervention programs. In harnessing the potential of computer-based interventions for individualized learning, teachers are best positioned to effectively integrate such interventions into their instruction so as to ensure that there is alignment between what is learned in the computer-based intervention and what is instructed in the classroom (Muralidharan et al., 2019). Teachers can also learn how they can interpret the progress monitoring data generated by such intervention programs, and provide individualized instruction while giving feedback to struggling students. They may also use the data gathered by the programs to identify and address common problematic confusions faced by students.

CONCLUSION

Non-responder groups, as in the current study, present a puzzle in reading intervention research. Following previous findings, the current non-responder group showed initially poor literacy skills, phonological awareness and rapid naming, along with weaker skills in verbal memory. This may dampen their progress after intensive intervention. Beyond simply extended practice, these students may need closer guidance in the type of practice employed, including by technology-based literacy apps. Considering the weaker skills that the non-responder group had at baseline, they may require practice that involves a lesser cognitive load. For example, in the adaptive games employed in the Grapholearn app, children advanced to more challenging trials and content once they mastered the previous level (at 80% correct performance). More difficult trials included selecting amongst more distractor letters to map to the given phoneme. For non-responders it may be beneficial to use the simpler levels where they have to choose between two print alternatives rather than multiple responses to map to each sound. While adaptive games are heralded as one way to individualize instruction and practice, the use of an adaptive game format also may have to be recalibrated to maintain students' engagement while increasing the challenge level of the tasks.

Regarding the responder group in the present study, responder status included fluency gains, and many of the students (41%) achieved word reading fluency in line with typical developing peers. This is encouraging, considering that intervention studies show less tractable effects for fluency and speed of decoding words and reading text, in contrast to reported strong effects on decoding accuracy (Torgesen et al., 2001). The present results suggest the interactive apps can supplement students learning and fluency through extended practice.

REFERENCES

Abadiano, H. R., & Turner, J. (2003). The RAND report: Reading for understanding: Toward an R&D program in reading comprehension. *New England Reading Association Journal, 39*(2), 74.

Ahmed, H., Wilson, A., Mead, N., Noble, H., Richardson, U., Wolpert, M. A., & Goswami, U. (2020). An evaluation of the efficacy of GraphoGame rime for promoting English phonics knowledge in poor readers. *Frontiers in Education, 5*, 132. doi: 10.3389/feduc.2020.00132

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special education, 23*(5), 300-316.

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning disabilities, 39* (5), 414-431.

Anderson-Inman, L., & Horney, M. A. (2007). Supported etext: Assistive technology through text transformations. *Reading Research Quarterly, 42*(1), 153–160. doi:10.1598/RRQ.42.1.8

Aravena, S., Snellings, P., Tijms, J., & van der Molen, M. W. (2013). A lab-controlled simulation of a letter–speech sound binding deficit in dyslexia. *Journal of Experimental Child Psychology, 115*, 691–707. doi: 10.1016/j.jecp.2013.03.009

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: the role of age and speed of stimulus presentation. *Developmental Science, 14*, 464–473. doi: 10.1111/j.1467-7687.2009.00937.x

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

Bao, Z. (1998). "The sounds of Singapore English". In J. Foley, et al. (Ed.) *English in New Cultural Contexts: Reflections from Singapore*. Singapore: Oxford University Press, 152–74.

Blachman, B. A. (1994). What we have learned from longitudinal studies of phonological processing and reading, and some unanswered questions: A response to Torgesen, Wagner, and Rashotte. *Journal of learning disabilities, 27*(5), 287-291.

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., ... & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences, 107*(17), 7939-7944. doi:10.1073/pnas.0904402107

Brown, A. (1988). "Vowel differences between Received Pronunciation and the English of Malaysia and Singapore: Which ones really matter?". In J. Foley, (Ed.) *New Englishes: The Case of Singapore. Singapore: Singapore University* Press, 129–47.

Chen, S. H. A., Ng, B., Tan, S. C., Poon, K. K., Yeo, L. S., Aljunied, S. M., et al. (2016). Findings and Data from Grant Supported by the Singapore National Research Foundation, Science of Learning Planning Grant (NRF2015-SOL001-011). Singapore: National Research Foundation.

Cheung, A. C., & Slavin, R. E. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly, 48*(3), 277-299.

Cunningham, A. E., Perry, K. E., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading & Writing, 14*, 549–568.

Denton, C. A., Fletcher, J. M., Anthony, J. L., & Francis, D. J. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of learning disabilities, 39*(5), 447-466.

Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013). Effects of

Tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *Journal of educational psychology, 105*(3), 633.

Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the international Phonetic Association, 27*, 47-55.

Deterding, D., & Poedjosoedarmo, G. (1998). *The Sounds of English: Phonetics and Phonology for English Teachers in Southeast Asia*. Singapore: Prentice Hall.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., ... & Sussex, W. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort.* Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Elliot, C. D. & Smith, P. (2012). *British Ability Scales – Third Edition*. London: GL Assessment.

Fletcher, J. M., Stuebing, K. K., Barth, A. E., Denton, C. A., Cirino, P. T., Francis, D. J., & Vaughn, S. (2011). Cognitive correlates of inadequate response to reading intervention. School *Psychology Review, 40*(1), 3-22.

Huang, F. L., & Invernizzi, M. A. (2014). Factors associated with lowercase alphabet naming in kindergarteners. *Applied Psycholinguistics, 35*(6), 943-968.

Huemer, S., Landerl, K., Aro, M., & Lyytinen, H. (2008). Training reading fluency among poor readers of German: Many ways to the goal. *Annals of Dyslexia, 58*(2), 115–137. doi:10.1007/s11881-008-0017-2

Juel, C. (1994). At-risk university students tutoring at-risk elementary school children: What factors make it effective. Getting reading right from the start: *Effective Early Interventions*, 39-62.

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

Low, E. L., & Brown, A. (2005). *English in Singapore: An Introduction.* Singapore: McGraw-Hill.

Lyytinen, H., Aro, M., Eklund, K., Erskine, J., Guttorm, T., Laakso, M.-L., et al. (2004). The development of children at familial risk for dyslexia: Birth to early school age. *Annals of Dyslexia, 54*, 184–220.

McTigue, E. M., Solheim, O. J., Zimmer, W. K., & Uppstad, P. H. (2019). Critically reviewing GraphoGame across the world: recommendations and cautions for research and implementation of computer-assisted instruction for word-reading acquisition. *Reading Research Quarterly, 55*, 45–73. doi: 10.1002/rrq.256

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review, 109*(4), 1426–1460. https://doi.org/10.1257/aer.20171112

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* National Institute of Child Health and Human Development, National Institutes of Health.

Nelson, R. J., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research & Practice, 18*(4), 255-267.

O'Brien, B.A., Habib, M.B., Onnis, L. (2019). Technology-based tools for literacy: Examining factors of intervention grain size and individual differences. *Frontiers in Psychology*. doi: https://doi.org/10.3389/fpsyg.2019.02625

O'Brien, B.A., Seward, R. & Zhang, D. (2022a). Multisensory interactive digital text for English phonics instruction with bilingual beginning readers. *Education Sciences, 12* (11), 750, https://doi.org/10.3390/educsci12110750.

O'Brien, B.A. & Yeo, S.W.A. (2022b). *Technology-based tools for teaching early literacy skills – Word learning app*, Nanyang Technological University Ref: 2022-125.

Patel, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2022). Assessing the effectiveness of a game-based phonics intervention for first and second grade English language learners in India: A randomized controlled trial. *Journal of Computer Assisted Learning, 38*(1), 76-89.

Peterson, R.L. & Pennington, B.F. (2015). Developmental dyslexia. *Annual Review of Clinical Psychology, 11*,283-307.

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Developmental Science, 21*(4), e12593. https://doi.org/10.1111/desc.12593

Rickard Liow, S. J., & Sze, W. P. (2008). *Bilingual Language Assessment Battery (BLAB)*. Singapore: National University of Singapore.

Saine, N., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-assisted remedial reading intervention for school beginners at risk for reading disability. *Child Dev. 82*, 1013–1028.

Seward, R., O'Brien, B., Breit-smith, A. D., & Meyer, B. (2014). Linking design principles with educational research theories to teach sound to symbol reading correspondence with multisensory type. *Visible Language, 48*(3), 86-108.

Snowling, M. J. & Hulme, C. (2021). Annual Research Review: Reading disorders revisited – the critical importance of oral language. *Journal of Child Psychology and Psychiatry, 62*(5), 635 –653. doi:10.1111/jcpp.13324

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.

Stetter, M. E., & Hughes, M. T. (2010). Computer-assisted instruction to enhance the reading comprehension of struggling readers: A review of the literature. *Journal of Special Education Technology, 25*(4), 1–16.

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research, 68*, 277–321.

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of Learning Disabilities, 3*3(2), 114–136.

Torgesen, J. K., Morgan, S. T., & Davis, C. (1992). Effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational psychology, 84* (3), 364-370.

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning disabilities research & practice, 15*(1), 55-64.

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*, 40-56.

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency-2*. London: Pearson Clinical Assessment.

Tran, L., Sanchez, T., Arellano, B., & Lee Swanson, H. (2011). A meta-analysis of the RTI literature for children at risk for reading disabilities. *Journal of learning disabilities, 44*(3), 283-295.

Treiman, R., & Kessler, B. (2003). The role of letter names in the acquisition of literacy. In R. V. Kail

(Ed.), *Advances in child development and behavior* (Vol. 31, 105–135). San Diego, CA: Academic Press.

Treiman, R., Kessler, B., & Pollo, T. C. (2006). Learning about the letter name subset of the vocabulary: Evidence from US and Brazilian preschoolers. *Applied Psycholinguistics, 27*, 211–227.

Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction. *Journal of Educational Psychology, 83*(4), 437.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*(4), 601.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing-2*. Austin: Pro-Ed.

Wagner, R. K., Zirps, F. A., Edwards, A. A., Wood, R. G., Joyner, R. E., Becker, B. J., Liu, G., & Beal, B. (2020). The prevalence of dyslexia: A new approach to its estimation. *Journal of Learning Disabilities, 53*(5), 354–365. doi:10.1177/0022219420920377

Walberg, H. J., & Tsai, S. (1983). Matthew effects in education. *American Educational Research Journal, 20*, 359-373.

Wilson, A., Ahmed, H., Mead, N., Noble, H., Richardson, U., Wolpert, M. A., & Goswami, U. (2021). Neurocognitive predictors of response to intervention with GraphoGame rime. *Frontiers in Education,* 6:639294. doi: 10.3389/feduc.2021.639294

Woodcock, R. W., McGrew, K. S., & Mather, N. (2007). *Woodcock Johnson III Tests of Achievement.* Rolling Meadows, IL: Riverside Publishing.

Worth, J., Nelson, J., Harland, J., Bernardinelli, D., & Styles, B. (2018). GraphoGame rime: Evaluation report and executive summary. National Foundation for Educational Research. https://www.nfer.ac.uk/graphogame-rime-evaluation-report-and-executive-summary