Computer-assisted Language Learning In Reading Comprehension by Using Visual Mnemonics

Mojtaba Aghajani1*, Hanieh Amanzadeh2

1Khatam ol-Anbia University, Tehran, Iran
2Alzahra University, Tehran, Iran

Corresponding Author: Mojtaba Aghajani, E-mail: Mojtaba_Aghajani45@yahoo.com

ABSTRACT

A profitable method to address reading delays is to use computer-assisted learning, but these techniques are not always effective. In this research, the researchers evaluated a commercially available computer system, which uses visual mnemonics, in a randomised controlled trial with 78 English-speaking children (mean age 7 years) who their schools identified as needing reading support. School based individual tutorials usually took place 2-3 times/week. Only the experimental group received the intervention in the first 10 months, thereafter both the experimental and control groups received the intervention for 6 months. After 10 months, the experimental group had significantly higher standardised scores than the waiting list control group of decoding, phonological awareness, naming speed, phonological short-term memory and executive loaded working memory. The computer-assisted intervention was effective and this suggests that this medium can be used for reading interventions with English speaking children.

Key words: Computer-assisted Learning, Reading Comprehension, Visual Mnemonics

INTRODUCTION

INTRODUCE THE PROBLEM

Education and employment prospects are often impaired in children with delays in reading (Dugdale & Clark, 2008, p. 45). There is a range of interventions to assist these children (Carroll, Bowyer-Crane, Duff, Snowling & Hulme, 2011). Slavin, Lake, Davis & Madden (2011) have reviewed non-computer interventions designed to help struggling readers (see also Blanchman et al., 2014). They concluded that one-to-one tutoring, especially when this involves phonics (i.e., relations between letters and sounds), and involves teachers rather than paraprofessionals usually is very effective; they also concluded that classroom wide programs, particularly cooperative learning, can have very positive effects (see also Griffiths & Stuart, 2013). However, these interventions can be expensive in terms of teaching practitioner time. In contrast, computer-assisted interventions offer the possibility of ‘expert’ instruction at relatively low cost and high fidelity, with advantages such as increased motivation, immediate feedback, self-pacing, and consolidation of learning and non-judgemental feedback (Mathes, Torgesen & Allor, 2001; Lynch, Fawcett & Nicolson, 2000; Torgesen, Wagner, Rashotte, Herron, & Lindamood, 2010).

However, several reviews suggest that computer-based interventions are not particularly effective. Blok, Oostdam, Otter and Overmaat (2001) concluded that these interventions produced positive but small effects; their review was based on experimental studies, many of which were not RCTs. More recently, Slavin, Lake, Davis and Madden (2011) concluded that “computer-assisted instruction had very positive effects on reading” (p1) and Cheung and Slavin (2012) concluded that there usually were small gains and effect sizes in computer-based interventions, especially when compared to non-computer interventions (p.120). Cheung and Slavin (2013) reviewed 20 investigations that met reasonable research rigour (13 of the studies involved randomisation, but six did not appear to be peer reviewed). They found that there was only a small positive effect size from a combined analysis of all the studies. Archer, Savage, Sanghera-Sidhu, Wood and Gottardo (2014) in their tertiary meta-analysis found that overall effect sizes were positive, but small. They also suggested that training and support for practitioners was associated with greater effectiveness, but they were not able to identify particular methods of intervention that were more effective than others. Since the reviews were conducted, there have been a number of reports of positive effects from computer-based interventions (e.g. Hughes, Phillips & Reed, 2013; Kyle, Kujala, Richardson, Lyytinen & Goswami, 2013; Schneider et al., 2016; Tyler, Hughes, Beverly & Hastings, 2015; see also Karemacker, Pitchford & O’Malley, 2010), but it also should be acknowledged that other investigations that were not included in the reviews have failed to find significant positive effects (e.g. Brooks, Miles, Torgerson & Torgerson, 2006; Given, Wasserman, Chari, Beattie & Eden, 2008).
Explore Importance of the Problem

Thus, the weight of available evidence suggests that most computer-based interventions are not of substantial help to children with reading difficulties. This raises questions about whether the medium of computer-based education is inherently ineffective with these children. One explanation of these findings is that computer-assisted learning fails because the subject matter does not lend itself to computer support, a very different position to that of technological determinism where it is believed that technology will help education and learning (Oliver, 2010). The absence of strong positive effects also may be part of a common historical cycle, identified by Crook and Lewthwaite (2010), where initial enthusiasms for new educational technology often fail to be supported by empirical evaluations. Consequently, given the range of findings about computer-based interventions, it is important to investigate a diversity of computer-based interventions to establish whether they fail to provide substantial help to children or whether there are techniques and approaches that are effective; an issue that is of considerable importance to practitioners and policy makers.

Another issue concerning computer-assisted interventions is the rarity of rigorous investigations that demonstrate significant improvements with English speaking children (see Dynarski et al., 2007; Campuzano, Dynarski, Agodini, & Rall, 2009). Several of the recent reviews of computer based reading interventions have highlighted the very small evidence base especially when taking into account the potential importance of the topic, the rarity and desirability of randomised controlled trials (RCTs), and the small increases in literacy abilities in the meta-analyses. For example, Slavin et al., (2011) only identified 14 relevant studies (3 were RCTs). A more extensive review by Cheung and Slavin (2012) concerned 84 studies using computer-assisted applications to support reading in typically developing and struggling readers, but only 16 of these studies appeared to have been subject to peer review and published since 2000. Thus, there continues to be a need for rigorous evaluations such as those provided by RCTs.

Describe Relevant Scholarship

Although reviewers have concluded that computer-assisted interventions involving English-speaking children produce small positive effects, there are a number of reports of effective interventions with children who speak other languages. A review by Blok, Oostdam, Oerstra, and Overmaat in 2002 indicated that computer interventions involving English-speaking children were more effective than those with non-English speaking children (mostly Dutch, speakers). However, since then a number of computer interventions with non-English speakers have been found to be very effective. These include Spanish (Levine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2010; Heikkila, Mikko, Narhi, Westerholm, & Ahonen, 2013; Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007), Spanish (Jiménez et al., 2007), and French (Ecalle, Kleinz, & Magnan, 2013; Magnan & Ecalle, 2006). Because these non-English languages have more consistent relations between graphemes and phonemes, this may explain their effectiveness in contrast to computer interventions with English speaking children who face the challenge of inconsistent grapheme-phoneme relationships. This connects with extensive literature about the central role played by letter-sound connections in the development of reading (Hulme & Snowling, 2009). Relevant to this issue, Kyle et al. (2013) successfully implemented a computer intervention that had been found to be effective with Finnish children (Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007); the UK intervention involved a fine-grained mapping of graphemes to phonemes that may have addressed issues of lack of transparency in English.

State Hypotheses and their Correspondence to Research Design

Thus, our investigation is designed to contribute to the limited evidence base involving RCTs about the effects of computer-based interventions with children who have reading delays. The research questions were:

1. Can the decoding abilities of English speaking children be improved by a computer-assisted intervention that uses visual mnemonics? Moreover, is there evidence of an effect of intervention duration?
2. Do any positive effects extend beyond the targeted abilities (decoding and phonological awareness) to near, non-targeted reading related abilities (spelling and naming speed), and to far abilities involving the working memory system (PSTM and ELWM)?

LITERATURE REVIEW

A long-standing technique to facilitate letter-sound mappings is the use of visual mnemonics, particularly embedded picture mnemonics, as used in books for initial readers (e.g., the Letterland series, Wendon, 2007; Alphafriends, 2001). The mnemonics contain both an object whose name begins with the target letter (e.g., snake), and whose shape can be drawn so that it resembles the target letter (e.g., a snake looks like the letter S). Furthermore, the technique appears to help the learning of letter-sound relationships. Schmidman and Ehri (2010, p. 22) gave English speaking children unfamiliar Hebrew letters to compare the benefits of pairing the novel letter with an embedded picture of the letter, or a picture of an object whose first letter corresponded to the sound of the letter to be learnt (e.g., a picture of a key when k had to be learnt). The embedded pictures were more effective in helping children learn the new letter-sound relations.

The importance of learning letter-sound correspondences suggests that interventions should use mnemonic devices, as these are more helpful than the simple pairing of letters and sounds, although the technique does not appear to have been previously studied in computer based learning. Partly for this reason it is decided to investigate an innovative computer based system for children with reading delays using what is termed Trainer text. At the core of this system is the use of visual mnemonics that appear above the text to supply relevant clues about each phoneme, including long vowels, so
that a child can decode words without other help. The visual cues consist of characters that have amusing names, which often provide an extra mnemonic to help remember the phoneme. This key part of the computer intervention provides scaffolding as children can access the visual mnemonics when they have difficulties decoding a written word, rather than being provided with explicit feedback contingent on the correct answer (see Muis, Ranellucci, Trevors & Duffy; 2015). As children are able to decode words that they would be unable to read without the support of visual mnemonics they are operating in their zone of proximal development. Furthermore, in Trainer text, visual mnemonics are provided for letters which have both normal and irregular phonetic rules, so the word ‘gas’ has the visual mnemonic of the Ant in Pink Pants available above the ‘a’, whereas the word ‘was’ has the visual mnemonic of the Octopus who Knocked a Puss (or the Ook with a Book in some dialects). In this way if a word is a barrier to reading, the learner can look above the letter and see, in this case, that the middle sound is /o/.

Although Trainer text is a key process in the intervention we investigated, there are other activities, such as games, to help with motivation and help consolidate reading related abilities. In addition, if the system detects signs of a specific cognitive weakness, such as visual tracking a moving target, the teacher is alerted to this with coaching on how to help the child overcome that specific difficulty. Consequently, our investigation involves a question about whether the use of computer-assisted intervention that uses visual mnemonics as a key part of the tutorial process helps the development of reading abilities. In previous research, it has been rare for the effects of reading interventions to be evaluated beyond the key abilities targeted by an intervention (see Slavin et al., 2011). In other areas of research, there has been interest in near and far effects of interventions (Dunning, Holmes & Gathercole, 2013). Consequently, our investigation is designed to ask whether the effects of the intervention extended from targeted reading abilities, to non-targeted reading-related abilities (near effects), or even to more general cognitive abilities (far effects). This methodology provides useful information about the generality of any intervention effects.

Two abilities that are targeted by this computer intervention involved decoding and phonological awareness so that these are seen as key indicators of success. In addition, assessments are made of two near abilities not targeted by the intervention, but may have benefitted from it. These are spelling, which is important aspect of literacy, and naming speed, which is highly related to literacy abilities (Kirby, Georgiou, Martinussen, & Parrila, 2010), although there are still uncertainties about the reasons for these associations (Stainthorp, Stuart, Powell, Quinlan & Garwood, 2010). Lastly, the researchers decide to assess two aspects of the working memory system that are usually seen as less directly associated with decoding. The two are phonological short-term memory (PSTM) and executive loaded working memory (ELWM). These were chosen as it has been suggested that PSTM is a phonological ability associated with reading (Hulme & Snowling, 2009, p. 10) and EWLM is often impaired in children who have literacy difficulties (Booth, Boyle, & Kelly, 2010), with evidence that training in this ability helps children’s reading (Loosli, Buschkuehl, Perrig, & Jaegg, 2012).

METHOD

Design

The design of the current study is quasi-experimental by considering the control over groups and randomization. First, A RCT (randomized controlled trial) was conducted with a Pre-test and 13 months later Post-test1 during this period only the experimental group received the intervention, the waiting list comparison group received teaching and support as normal in the school. Thereafter both groups received the intervention with the Post-test2 assessments seven months later, allowing evaluation of the effects of shorter versus longer interventions in what has sometimes been termed a dose-response analysis (Morrison & Connor, 2002; the actual duration of school-based tutorials was slightly shorter because of school holidays).

Similar RCT designs to this with waiting list control groups have been used to in a number of important evaluations of reading interventions as it addresses questions about the added value of the intervention beyond any existing and usual support that is available. Moreover, comparing an intervention to ‘teaching as usual’ does not always result in large or lasting effects (Bachelor et al., 2006; Duff et al., 2014). As in previous evaluations, regression analyses were used to investigate group differences (e.g. Hatcher et al., 2006) and effect sizes were calculated so that our findings can be compared with the effectiveness of other interventions.

Participants

Six schools agreed to take part in the research and this was an opportunity sample. Within each school, the children were randomly assigned to the experimental and control conditions to minimise any effects of school on the findings. The schools were located within. There was no evidence of an over-representation of social deprivation in the sample.

All children, who were identified by these schools as needing support because of their poor progress in reading related activities, were invited to take part (parental permission was first obtained), as a result, the participants corresponded to the target population of concern to many schools. These children were most likely to have delays in decoding, but in some cases the children may have had other reading related delays such as poor comprehension, which could have made an intervention which targeted decoding less effective (see Potocki, Magnan & Ecalle, 2015). Details of the children’s reading related abilities are given in the pre-test scores in Table 1; the mean standardised scores were in the low typical range for decoding, phonological awareness, PSTM and ELWM. All the children had percentile scores above 30 on the BAS II Matrices Scale of non-verbal intelligence (Elliott, Smith, & McCulloch, 1996).
The identification of children with reading difficulties by the six schools followed Iran national guidelines involving three levels of need and support. At the levels of school action or school action plus teachers identified children who were making slower progress in reading than their peers. School action involved support within the school, and school action plus involved professionals from outside the school providing additional advice or assistance, no information was systematically collected about the role of these professionals although our impression was very few if any children received additional provision from outside the schools. For some children, clinical assessments had resulted in a formal statement of special educational need which constituted the third level. Two additional children were included in the sample because an independent assessment had indicated that they had poor decoding abilities (they had standardised scores on the Test of Word Reading Efficiency below 86 (TOWRE; Torgesen et al., 1999)).

Studies of reading show that the level of word knowledge is related to reading skills (Rohwer & Snow, 1989). Children with reading difficulties are more likely to have an impoverished vocabulary (McKenna & Wyness, 1992).

**Assessments**

The following assessments were administrated using the guidelines provided by the test manuals. The sessions were conducted by a very experienced assessor in a quiet room at the schools.

**Decoding.** Decoding was assessed using the TOWRE (Torgesen et al., 1999), a test of the ability to read and pronounce lists of words with accuracy and fluency. There are two subtests, each with two alternate forms (A and B). One subtest assesses the accurate decoding of real words, the second assesses the decoding of pronounceable non-words. After the practice items the children read the test items. Assessment starts when the child reads the first word and ends after 45 seconds. Any words/non-words skipped or read incorrectly are marked and only those read correctly are scored. The manual gives the average test-retest (time sampling) coefficient for the same form and for alternate forms as above.90.

**Phonological awareness.** Phonological awareness was assessed using three tests from the Phonological Assessment Battery (Frederickson, Frith and Reason, 1997). The tests of rhyme and alliteration awareness were used at Pre-test and Post-test1. One subtest assesses the accurate decoding of real words, the second assesses the decoding of pronounceable non-words. After the practice items the children read the test items. Assessment starts when the child reads the first word and ends after 45 seconds. Any words/non-words skipped or read incorrectly are marked and only those read correctly are scored. The manual gives the average test-retest (time sampling) coefficient for the same form and for alternate forms as above.90.

<table>
<thead>
<tr>
<th>Table 1. Means and standard deviations at Pre-test, Post-test1 and Post-test2 of reading-related abilities in the experimental and waiting list control groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental group</strong></td>
</tr>
<tr>
<td>Pre-test</td>
</tr>
<tr>
<td>Decoding Mean</td>
</tr>
<tr>
<td>S.D.</td>
</tr>
<tr>
<td>Phon. awareness Mean</td>
</tr>
<tr>
<td>S.D.</td>
</tr>
<tr>
<td>Naming speed Mean</td>
</tr>
<tr>
<td>S.D.</td>
</tr>
<tr>
<td>Spelling Mean</td>
</tr>
<tr>
<td>S.D.</td>
</tr>
<tr>
<td>PSTM Mean</td>
</tr>
<tr>
<td>S.D.</td>
</tr>
<tr>
<td>ELWM Mean</td>
</tr>
<tr>
<td>S.D.</td>
</tr>
</tbody>
</table>

RETRACTED
sound. The Spoonerism test was used at Post-test2 to avoid ceiling effects. There were two tasks. A word was spoken and the child was asked to replace the beginning sound with another one (e.g. cat with a/f/gives?); the other task involved transposing the beginning sounds of two words (e.g. King John gives Jing Kohn). The manual reports the internal consistency as being above 0.8.

**Naming Speed (Rapid Automatized Naming, RAN).** Naming speed was assessed using a subtest from the PhAB (Fredrickson et al., 1997). The children had to name as quickly as possible, 50 pictures of 5 objects (e.g., hat, ball, door) which were presented in 5 rows with there being no discernible sequence to their order. The time taken and any errors were recorded. The internal consistency is reported as being above 0.8.

**Spelling.** The BAS II Spelling Scale (Elliott et al., 1996) was used to assess the spelling of single words. The experimenter dictated the words one at a time and the children wrote these down. The words were grouped into age-related blocks of ten, and starting and stopping points are according to age and ability. The manual gives test-retest reliability as 0.93.

**Phonological Short Term Memory (PSTM).** The word recall test from the Working Memory Test Battery for Children (WMTB-C, Pickering, & Gathercole, 2001) was used to assess PSTM. Children had to repeat the single syllable words spoken by the experimenter immediately and in the correct order. Children started with a set of 6 single words, the number of words was increased by 1 until there were three errors within a set. The children’s word span was used in the analyses. Test-retest reliability is reported as 0.8.

**Executive Loaded Working Memory (ELWM).** ELWM was assessing using the listening recall test from the WMTB-C. The experimenter read a sequence of short sentences, which did, or did not make sense (e.g. flowers play football). The child said whether the sentence was true or false. Then the child recalled the final word of the sentence(s) in the correct order. As with PSTM, trials started with single item sequences and were repeated until there were three errors made within a test. Listening span was used in the analyses. Test-retest reliability 0.738-0.83 are reported for relevant ages.

**Non-verbal Intelligence.** The BAS II Matrices Scale (Elliott et al., 1996) was used to assess non-verbal intelligence. The child was shown an incomplete abstract matrix, they had to select from six choices, the figure that correctly completed the matrix.

The Intervention

The tutorials were usually 10-15 minutes long and usually took place 2-3 times a week during the school terms. An adult, usually a teaching assistant, was present during the tutorials which often involved a small group of children (2-3) working on their own individual computer. The adults had taken part in an online training package. A Pre-test occurred in October/November of the first year and a Post-test1 occurred 13 months later; during this period only the experimental group received the intervention. Additionally, both groups could receive the intervention for 7 months after Post-test1 and before Post-test2. During this second phase of 7 months, the tutorials could be terminated for children in the experimental group if the teachers considered they were no longer necessary, this usually happened if children had made good progress and there were other demands made on the time of teaching assistants. Because of holidays, the children had school based sessions for 10 months between pre-test and post-test1 and for 6 months between post-test1 and post-test2. In the experimental group, the average number of tutorials between Pre-test and Post-test1 was 110.2 (S.D. 24.9), the number between Post-test1 and Post-test2 was 50.2 (S.D. 28.0), in the waiting list control group the corresponding number was 55.0 (S.D. 15.8).

The intervention and waiting list control groups continued to receive the usual support that the schools deemed appropriate, most of the support involved additional reading related activities provided by teaching assistants. The intervention was designed to have high fidelity; adults were always present, they checked that the children were engaging in the tasks and that the tasks were completed. Regular visits confirmed that the intervention was administered appropriately.

The core feature of the computer tutorials was the use of Trainertext and interactive, multimedia materials. As already described, different, amusing, visual characters are linked to the same grapheme (see figure 1) it also should be noted that Trainertext involves an alphabetic approach, by giving the children a single character to represent phoneme pairs. An example is the Star in a car, to represent the phonemes /s/ and /t/. This type of presentation is different to conventional phonics instruction in the U.K. because it uses implicit instruction (Easyread, 2014) rather than explicit decoding rules.

The tutorial experience involved two key reading activities with Trainertext. This is the time when the facilitator was most likely to provide assistance by overseeing effective decoding of the Trainertext materials and inputting a child’s performance into the computer:

Activity 1. Four phrases of 5-6 words are presented as Trainertext images with no letters, thereby ensuring children decodes each word phonetically.

Activity 2. In the next session children read aloud the same phrases in plain text and can click on letters so as to view the Trainertext mnemonics when needed. An important part of the adult’s role is to ensure they are able to decode the word with fluent accuracy. When this is achieved the children return to Activity 1 with a different set of words and the sequence is repeated for 5-10 minutes. Once the learner has finished these activities there is a final game, which is randomly picked from 14 different games in the system. The games are fun and provide a reward at the end of the lesson, but also involve some decoding for successful performance. As the adults are working with small groups of children it is possible to provide individual attention during the important activities as the children were usually progressing at different rates and it was possible to and/or briefly delay the activities of one child while focussing on another.
Thus, children learn by decoding words for a few minutes each day to establish procedural knowledge (Messer, & Pine, 2000). After 95 tutorials, by which time this decoding approach has become very familiar, the amount of text is gradually increased, until the learner is reading around 150 words per day (recorded by the adult facilitator).

The intervention is structured to maximise initial success with, and mastery of, basic reading related tasks. The first 14 lessons give a foundation with gaming activities to build the child’s confidence, particularly in relation to decoding. Then there are reading activities with Trainertext, with a game before and after the activity to help maintain motivation. The early games are simple decoding related activities to help children become comfortable with these tasks. For instance, children changed a word into another word (e.g., lamb to limb) by choosing one of several graphemes that are positioned above the word and dragging the grapheme onto the letter that needs to be replaced. The learners are supported through these gamified exercises with audio instructions and affirmations. These are suitably randomised to give a natural feel to the experience. When a mistake is made in one of the gamified exercises, spoken feedback generated by the computer is given to the learner, along with some guidance on completing the task correctly. In most instances this involves a friendly spoken response (e.g. “Oops, it wasn’t that one”) and spoken encouragement from the computer to attend to relevant information.

The computer tutorials involve a multifaceted approach to reading difficulties and are designed to deal with a range of challenges that children can face, such as weaknesses with processing sounds, memory for sounds, fluency, ability to follow a moving target, and convergence. The system also has been designed to reduce stress and anxiety about reading. Usually the progress is monitored centrally so individual assistance can be provided if needed, but no such assistance was provided during the investigation.

Data Checking and Inspection

The children’s non-verbal ability in the two groups as assessed by the British Ability Scales, Matrices Scale (BAS II; Elliott et al., 1996) were similar at the beginning of the study (mean and SD: at Pre-test, experimental 52.2, 5.6; control 53.8, 4.6; at Post-test1 experimental 53.7, 5.2; control 53.7, 5.4). A 2 (groups) x 2 (times Pre-test, Post-test1) repeated ANOVA failed to identify any significant effects of group, time or interaction. This suggests that any improvements are not attributable to a change in general non-verbal ability.

For all variables, skewness and kurtosis were within acceptable limits (George & Mallory, 2010). Some outliers were identified, therefore, further statistical checks were carried out in relation to the regression analyses (Durbin-Watson, tolerance/VIF statistics, and Cook’s/Mahalanobis distances), all were within acceptable limits (Field, 2009).

RESULTS AND DISCUSSION

Standardised scores were used as these take account of the effects of age related improvements. Table 1 show the standardised scores for the two groups at the three time points. At Pre-test both groups had mean standardised scores on six reading related abilities that were below those expected for typically developing children. After the intervention, at Post-test1 the experimental group had mean scores on most of the assessments that were close to or above 100, and these were maintained over the next 7 months. In contrast, the waiting list control group made little if any progress between Pre-test and Post-test1. After the control group had experienced the intervention between Post-test1 to Post-test2, most of standardised scores increased, but not to the levels achieved by the experimental group at either post-test.

A separate multiple hierarchical multiple regression analysis was conducted to investigate whether group had a significant effect on each of the six dependent variables at Post-test1, while controlling for initial abilities at Pre-test1; this was similar to conducting ANCOVAs (see Hatcher et al., 2006; Duff et al., 2014). In each of the six regression analyses, BAS matrices T-score was entered at Step 1 to control for general cognitive ability. The standardised score at Pre-test of the relevant independent variable was entered at Step 2 to control for pre-intervention ability and autoregressive effects (e.g. when the Post-test1, TOWRE scores were the dependent variable, the Pre-test TOWRE scores were entered at Step 2 to control for initial abilities). Group was dummy coded and entered at Step 3 to investigate whether the intervention resulted in group differences in reading related abilities.

A summary of these regression analyses is provided in Table 2. The adjusted $R^2$ values indicate that the three independent variables entered into each of the six regression analyses together accounted for a reasonably high percentage of the variance of the dependent variable (range 0.4 to 0.7). At Step 1, the matrices T-score assessment, only was a significant predictor of phonological awareness and was not a significant predictor of any of the other dependent variables, this indicates non-verbal ability was not an important influence on most of the dependent variables at Post-test1. In contrast, all the Pre-test variables entered at Step 2 resulted in a significant $R^2$ change; this is to be expected, as initial abilities often are the best predictor of the same later abilities due to autoregression. At Step 3, group resulted in a significant reduction in $R^2$ in five analyses, the exception being spelling, this shows that after taking account of BAS matrices ability and initial ability, the experimental group still had a significant effect on five of the Post-test1 scores (decoding, phonological awareness, naming speed, PSTM and ELWM). The size and significance levels of the standardised Beta coefficients at Step 3 confirmed that our findings were not simply the result of the order of entry of the variables.

A second set of regression analyses was conducted for each of the six dependent variables to investigate whether the longer exposure to the intervention in the experimental group resulted in higher gains. In each regression, the dependent variable was either the gains of the experimental group between Pre-test and Post-test1, or the gains of the waiting list control group between Post-test1 and Post-test2. The time points for the experimental group were chosen because there...
were no appreciable increases in the standardised scores between Post-test1 and Post-test2. The entry of variables into each regression was the same as in the previous analysis.

The adjusted $R^2$ values were lower than in the previous analysis indicating that it was more difficult to predict gain scores from the three independent variables. The findings from adjusted $R^2$ change and Beta coefficients gave a consistent picture (see Table 3). BAS matrices was not a significant predictor of any of the gain scores, as in the previous analyses this suggests that non-verbal ability was not an important influence on the children’s gains. Nor were the Pre-test scores of the dependent variable particularly good predictors of the gain scores, only Pre-test rapid naming and Pre-test spelling were significant predictors of the relevant gains. This suggests that the sizes of the gains were not related to the initial reading-related abilities and therefore length of intervention (10 months versus 6 months), was a significant predictor of decoding, rapid naming, PSTM and ELWM. Consequently, for these variables the intervention with a longer duration and more sessions appears to have been significantly more effective than the shorter intervention.

To provide information about the effect sizes of the increases in reading related abilities Cohen’s $d$ was calculated for the gain scores relevant to the two regression analyses (see Table 4). Between Pre-test and Post-test1, apart from spelling, there were large effect sizes, all above 0.80, indicating that the intervention was highly effective in changing the children’s abilities; the highest effect sizes were for decoding and phonological awareness. In the comparison of the interventions which were of different lengths (i.e. experimental group between Pre-test and Post-test1 versus control group between Post-test1 and Post-test2), there were large effect sizes for the improvement of the TOWRE and RAN, with medium effect sizes for PSTM and ELWM, there were negligible or no effects for phonological awareness and spelling.

**CONCLUSION**

The investigation has provided evidence that English-speaking children considered by their schools as in need of support for their reading can benefit from computer-assisted learning. After the children in the experimental group received an intervention for 10 months they had mean standardised scores on a range of reading related abilities that were close to or above 100, and these scores were maintained over the next 7 months. Consequently, the intervention appeared to

<table>
<thead>
<tr>
<th>Variables</th>
<th>Adjusted $R^2$</th>
<th>$R^2$ change at steps 1-3</th>
<th>Standardised Betas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Step 1 (BAS matrices)</td>
<td>Step 2 (Relevant independent variable)</td>
</tr>
<tr>
<td>Decoding</td>
<td>0.67</td>
<td>0.01</td>
<td>0.49***</td>
</tr>
<tr>
<td>Phonological awareness</td>
<td>0.58</td>
<td>0.17***</td>
<td>0.28***</td>
</tr>
<tr>
<td>Rapid naming</td>
<td>0.54</td>
<td>0.00</td>
<td>0.48***</td>
</tr>
<tr>
<td>Spelling</td>
<td>0.73</td>
<td>0.05</td>
<td>0.69***</td>
</tr>
<tr>
<td>PSTM</td>
<td>0.40</td>
<td>0.01</td>
<td>0.20***</td>
</tr>
<tr>
<td>ELWM</td>
<td>0.47</td>
<td>0.05</td>
<td>0.35***</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.01; ***p<0.001

<table>
<thead>
<tr>
<th>Variables</th>
<th>Adjusted $R^2$</th>
<th>$R^2$ change at steps 1-3</th>
<th>Standardised Betas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Step 1 (BAS matrices)</td>
<td>Step 2 (Relevant independent variable)</td>
</tr>
<tr>
<td>Decoding</td>
<td>0.20</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>Phonological awareness</td>
<td>0.09</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>Rapid naming</td>
<td>0.24</td>
<td>0.02</td>
<td>0.12**</td>
</tr>
<tr>
<td>Spelling</td>
<td>0.07</td>
<td>0.02</td>
<td>0.05*</td>
</tr>
<tr>
<td>PSTM</td>
<td>0.08</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>ELWM</td>
<td>0.13</td>
<td>0.00</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.01; ***p<0.001
have overcome the Matthew effect where children fail to catch up with their typically developing peers.

The first regression analysis revealed that between Pre-test and Post-test1, the children in the intervention group had significantly higher scores than those in the control group in decoding, phonological awareness, naming speed, PSTM and ELWM. The differences were present despite the earlier entry of non-verbal ability and the relevant Pre-test ability into the regression to control for initial abilities and auto-regression effects. Furthermore, the effect sizes for the gains between Pre-test and Post-test1 were large (except for spelling). This indicates that the intervention had a significant impact on the reading of the children and these effect sizes were much higher than those generally reported in reviews (e.g., Slavin et al., 2011; Cheung & Slavin, 2013).

A comparison of the gains made by the two groups revealed that a longer 10-month intervention had more extensive and significantly higher gains in reading-related abilities than the 6-month intervention. This effect could have been due to a longer period for the consolidation of skills, or the higher number of sessions or the intensity of the sessions. This increases our confidence in the findings of the RCT analysis as there appears to have been a ‘dose-related’ increase. Previous informal observations suggest that there needs to be several sessions a week for the system to be effective and that noticeable gains occur between sessions 60 and 90. Somewhat surprisingly, phonological awareness failed to show an effect of duration, suggesting that the shorter intervention was equally effective as the longer one. As the standardised scores of phonological awareness at the end of both interventions remained in the mid-90s, this may reflect a difficulty in increasing children’s phonological abilities even with longer interventions (Hulme & Snowling, 2009).

Our investigation also showed that the intervention had effects on targeted, near and far reading-related abilities. To our knowledge, this has not been shown before with this range of variables (see Cheung & Slavin, 2013). Decoding and phonological awareness were specifically targeted by the intervention and as might be expected these showed the largest effect sizes in the comparison of groups between Pre-test and Post-test1. There were some mixed findings for the two near abilities that were not directly targeted by the intervention (naming speed and spelling). The intervention group had significantly higher naming speed than the control group at Post-test1. Because the intervention did not focus on speed related tasks, the improvements might have been because children became more efficient at accessing phonological verbal and/or translating visual stimuli into a phonological form. It is also worth noting that Torgesen et al., (2010, p, 75) have reported positive effects of a computer intervention on naming speed with children in the U.S.A., however, Heikkila et al., (2013) failed to produce a similar effect with Finnish children.

Spelling was the only variable not to show a significant group difference at Post-test1. This was unexpected as decoding and spelling are closely associated (e.g., Christopher et al., 2012) and the failure could be due to several possibilities including the relatively high standardised scores for spelling at Pre-test and the need for lengthy consolidation before decoding gains translate into spelling gains. In addition, improvements in decoding may not necessarily translate to spelling (Hughes et al., 2013) and even computer interventions which target spelling have not been successful (Bishop, Adams, Lehtonen & Rosen, 2005).

There were significant improvements in the two components of the working memory system that constituted far abilities. The increase in PSTM could have been because the learning of letter sound relationships helped the accurate storage of non-words. This in itself is notable as few previous interventions have been able to increase PSTM abilities (Hulme & Muir, 1985; Henry, Messer & Nash, 2014). There was also a significant effect of group on ELWM. Given the interest in finding ways to improve ELWM these findings deserve further attention (see Melby-Lervåg & Hulme, 2012). ELWM involves the simultaneous processing of current information and the storage/retrieval of other information; consequently, it seems likely that the intervention had an effect on ELWM because the computer tasks often involved having to simultaneously process two forms of visual information (graphemes and mnemonic devices) as well as the concurrent retrieval of phonological information from memory. To assess whether these effects extend beyond the verbal domain and involve a general effect on ELWM (Pimperton & Nation, 2014).

The overall success of the intervention contrasts with many previous investigations and the general conclusions from reviews about computer-assisted interventions (Cheung & Slavin, 2012). Consequently, our findings indicate that it is possible for computer-assisted interventions to have significant effects on English-speaking children’s reading related abilities. This adds to the existing evidence that some computer-based interventions can be effective with these children. It is useful to speculate about the reasons for the effectiveness of the intervention as this can provide a basis for further research and for software developments. The most obvious possibility for the success of the intervention

### Table 4. Effect sizes on group differences in reading related abilities using Cohen’s D on (i) gain scores for Pre-test to Post-test1 and (ii) gain scores for experimental group Pre-test to Post-test1 compared to control group Post-test1 to Post-test2

<table>
<thead>
<tr>
<th>Variables</th>
<th>Both groups Pre-test to Post-test1</th>
<th>Experimental Pre-test to Post-test1 cf Control Post-test1 to Post-test2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decoding</td>
<td>1.34</td>
<td>0.97</td>
</tr>
<tr>
<td>Phonological awareness</td>
<td>1.03</td>
<td>0.69</td>
</tr>
<tr>
<td>Rapid naming</td>
<td>0.84</td>
<td>0.80</td>
</tr>
<tr>
<td>Spelling</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>PSTM</td>
<td>0.90</td>
<td>0.69</td>
</tr>
<tr>
<td>ELWM</td>
<td>0.74</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Our investigation also showed that the intervention had effects on targeted, near and far reading-related abilities. To our knowledge, this has not been shown before with this range of variables (see Cheung & Slavin, 2013). Decoding and phonological awareness were specifically targeted by the intervention and as might be expected these showed the largest effect sizes in the comparison of groups between Pre-test and Post-test1. There were some mixed findings for the two near abilities that were not directly targeted by the intervention (naming speed and spelling). The intervention group had significantly higher naming speed than the control group at Post-test1. Because the intervention did not focus on speed related tasks, the improvements might have been because children became more efficient at accessing phonological verbal and/or translating visual stimuli into a phonological form. It is also worth noting that Torgesen et al., (2010, p, 75) have reported positive effects of a computer intervention on naming speed with children in the U.S.A., however, Heikkila et al., (2013) failed to produce a similar effect with Finnish children.

Spelling was the only variable not to show a significant group difference at Post-test1. This was unexpected as decoding and spelling are closely associated (e.g., Christopher et al., 2012) and the failure could be due to several possibilities including the relatively high standardised scores for spelling at Pre-test and the need for lengthy consolidation before decoding gains translate into spelling gains. In addition, improvements in decoding may not necessarily translate to spelling (Hughes et al., 2013) and even computer interventions which target spelling have not been successful (Bishop, Adams, Lehtonen & Rosen, 2005).

There were significant improvements in the two components of the working memory system that constituted far abilities. The increase in PSTM could have been because the learning of letter sound relationships helped the accurate storage of non-words. This in itself is notable as few previous interventions have been able to increase PSTM abilities (Hulme & Muir, 1985; Henry, Messer & Nash, 2014). There was also a significant effect of group on ELWM. Given the interest in finding ways to improve ELWM these findings deserve further attention (see Melby-Lervåg & Hulme, 2012). ELWM involves the simultaneous processing of current information and the storage/retrieval of other information; consequently, it seems likely that the intervention had an effect on ELWM because the computer tasks often involved having to simultaneously process two forms of visual information (graphemes and mnemonic devices) as well as the concurrent retrieval of phonological information from memory. To assess whether these effects extend beyond the verbal domain and involve a general effect on ELWM (Pimperton & Nation, 2014).

The overall success of the intervention contrasts with many previous investigations and the general conclusions from reviews about computer-assisted interventions (Cheung & Slavin, 2012). Consequently, our findings indicate that it is possible for computer-assisted interventions to have significant effects on English-speaking children’s reading related abilities. This adds to the existing evidence that some computer-based interventions can be effective with these children. It is useful to speculate about the reasons for the effectiveness of the intervention as this can provide a basis for further research and for software developments. The most obvious possibility for the success of the intervention

### Table 4. Effect sizes on group differences in reading related abilities using Cohen’s D on (i) gain scores for Pre-test to Post-test1 and (ii) gain scores for experimental group Pre-test to Post-test1 compared to control group Post-test1 to Post-test2

<table>
<thead>
<tr>
<th>Variables</th>
<th>Both groups Pre-test to Post-test1</th>
<th>Experimental Pre-test to Post-test1 cf Control Post-test1 to Post-test2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decoding</td>
<td>1.34</td>
<td>0.97</td>
</tr>
<tr>
<td>Phonological awareness</td>
<td>1.03</td>
<td>0.69</td>
</tr>
<tr>
<td>Rapid naming</td>
<td>0.84</td>
<td>0.80</td>
</tr>
<tr>
<td>Spelling</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>PSTM</td>
<td>0.90</td>
<td>0.69</td>
</tr>
<tr>
<td>ELWM</td>
<td>0.74</td>
<td>0.65</td>
</tr>
</tbody>
</table>
was the novel use of visual imagery/mnemonics involving Trainertext to help children remember the relationship between graphemes and phonemes and thereby assist decoding in the presence of a supportive adult. There is broad agreement that this mastery of the phonetic coding is a crucial aspect of the development of reading and that this mastery is particularly problematic in English (Slavin et al., 2011), and computer interventions can effectively target phonological abilities (Comaskey, Savage & Abrami, 2009; Wild, 2009). A similar technique to Trainertext has been used in books for beginning readers (e.g. the Letterland series), but this system goes beyond a single relationship between a letter and sound that occurs in many readers (e.g. the ‘s’ sound in snake), to provide for the same letter, different visual mnemonics when the letter can be pronounced in different ways in different words. In addition, previous findings suggest that interventions involving phonological awareness are more effective when these are associated with the use of letters than when confined to speech processes (Ehri, Nunes, Willows, Schuster, Yaghoub-Zarabieh, & Shenan, 2001), and other interventions which use different methods to address the diversity of grapheme-phoneme relations in English have been effective (Hughes et al., 2013; Dyyle et al., 2013). Therefore, as the Trainertext technique is different from conventional phonics instruction it deserves further evaluation.

There are also several general features of the intervention, in addition to Trainertext, that might have resulted in the increase in children’s reading related abilities. Because the computer tutorials have been developed over 12 years in the light of feedback and observations, they have been fine-tuned to make the learning experience enjoyable, motivating and effective. Related to this, several of the successful interventions for non-English speaking children have been the result of a sustained programme of research and interventions (e.g., Ecalle et al., 2013; Lyytinen, Ronimus, Alanko, Poiikkeus, & Taanila, 2007). It also is possible that because this computer system is designed to provide support for a range of literacy related impairments that this contributed to its effectiveness. Furthermore, he relatively long 10-months duration and the number of sessions of the experimental intervention and of the 6-month period between Post-test2 and Post-test3 is also likely to have enhanced their effectiveness, and it should be noted that these are higher than in many other interventions. However, some interventions over similar periods have failed to produce significant gains (Cheung & Slavin, 2013), so it is unlikely that simply the duration, number of session were the main reason for the improvements.

A limitation of the study was that there was no assessment of reading comprehension (see Lysenko & Abrami, 2014). However, related to this, it is worth noting that the intervention had significant effects on decoding and ELWM; these often are regarded as key to reading comprehension abilities (Cain, Oakhill, & Bryant, 2004; but see Pimperton & Nation, 2014). Consequently, it would be worth investigating reading comprehension in future research, although there are mixed findings about the effects of interventions designed to increase this ability (Savage et al., 2013; Paul & Clarke, 2016). Another limitation was that the evaluation of the dose-response effect did not involve the same time-period and so findings about this effect should be treated with caution, especially as the comparison did not involve a RCT.

To summarise, this evaluation has shown that computer-assisted interventions can be effective with children identified in need of literacy support. There are several features of the intervention that might have contributed to this success: the duration of the intervention; the fine-tuning of the tutorials over a long period and perhaps most importantly the use of visual mnemonics in the form of Trainertext. This investigation has also provided new information on the important of the duration of interventions and the way that they can influence a range of near and far reading related abilities.

REFERENCES


