Evaluation of an online reading programme to improve pupils' reading skills in primary schools

Watkins, R. C.; Hulson-Jones, Amy; Tyler, Emily; Hastings, R.; Beverley, Michael; Hughes, Carl

Wales Journal of Education

Evaluation of an online reading programme to improve pupils' reading skills in primary schools: Outcomes from two implementation studies

Richard Watkins[1,2], Amy Hulson-Jones[2], Emily Tyler[3,2], Richard P. Hastings[4], Michael Beverley[2] and J. Carl Hughes[2]

[1]GwE, Regional School Effectiveness and Improvement Service for North Wales

[2]School of Psychology, Bangor University

[3]School of Education, Bangor University

[4]Centre for Educational Development Appraisal and Research, University of Warwick

Corresponding author: Dr Richard Watkins. Email: richardwatkins@gwegogledd.cymru

ABSTRACT

Headsprout Early Reading (HER) is an online reading programme designed to teach pupils early reading skills. It uses adaptive technology to deliver systematic, synthetic phonics instruction, and is an effective supplementary teaching aid for beginning readers in mainstream and special schools (aged 4-7 years). In the current study, an evaluation of HER was conducted with two mainstream primary schools in North Wales. The primary aim was to investigate whether pupils receiving HER would improve early reading skills (i.e. catch up) in school A compared to a control group in the same school. The secondary aim was to describe the use of HER without implementation support (School A) and with support (School B) following initial training of teaching assistants, and to present evidence about the reading outcomes for the pupils in the two schools. We report findings from 28 Year 2 pupils (6-7 years) in school A and 14 Year 2-6 pupils (6-11 years) in School B. Analysis of pre- and post-test standardised reading scores indicate statistically significant improvements in reading comprehension scores for HER pupils within School A compared to pupils not receiving intervention. There are also statistically significant improvements in pupils' DIBELS (correct words) sub-test measures in favour of the school receiving the implementation support (School B) compared to the school not receiving implementation support (School A). These results support previously reported findings that HER is an effective supplementary reading programme for struggling readers. It also indicates that HER can be delivered effectively by schools with minimal support.

Key words: Online reading programme, basic reading skills, Headsprout, education.

Introduction

The acquisition of early reading skills is an essential element of a child's early primary education. While many children learn to read at the expected rate without the need for supplementary tuition, a significant number fail to acquire early reading skills and subsequently require additional intervention and support at school (Vaughn and Fuchs, 2003; Institute of Education Sciences, 2010).

Those pupils who fall below the typical reading achievement expected in mainstream education are unlikely to access the full range of curriculum experiences as they move through primary and into secondary education, and are subsequently at a greater risk of encountering difficulties throughout their academic experience (Kamil, 2003). Intervention to remediate reading deficiencies is most successfully achieved early (Cooke, Kretlow and Helf, 2010), and should enable a pupil to learn a greater amount in a shorter space of time (Johnson and Layng, 1994; Johnson and Street, 2004).

Following disappointing results in the internationally comparative PISA tests in 2009 (OECD, 2010), Welsh Government identified the need to raise standards in reading, mathematics, and science (Welsh Government, 2013). Furthermore, the need to improve standards of literacy and numeracy across the curriculum led to the creation of a new National Literacy and Numeracy Framework (LNF) in 2013 to provide schools with a progressive framework for the teaching and assessment of literacy and numeracy skills for learners aged 5 to 14 years (Welsh Government, 2013). Alongside the introduction of the LNF, Welsh Government also introduced annual standardised numeracy and reading tests in 2013, with expectation that schools would use these data for diagnostic and/or formative purposes.

Since the introduction of the LNF in 2013, improving the level of reading has remained a major priority for Welsh Government and the newly established regional consortia (Hill, 2013). Although the percentage of pupils reaching the expected level in language, literacy and communication skills at the end of Year 2 in Wales has increased steadily since 2012, this figure still remains below the corresponding indicators for personal and social development, well-being and cultural diversity, and mathematical development (Estyn, 2016). The Welsh Government's flagship vision for improving Welsh schools up to 2020 (Welsh Government, 2014) highlights the need to improve the quality of provision for learners of all abilities, including the use of digital technologies to develop approaches to more personalised learning.

Over the past 20 years, a growing body of evidence has indicated the effectiveness of systematic phonics as a method of teaching early reading skills (Education Endowment Foundation, 2016). In the USA, the National Reading Panel report (2000) highlighted the importance of teaching reading using a systematic phonics approach, and identified the five key skills that produce functional readers:

- phonemic awareness (the ability to recognise that sound are made up of separate sounds called phonemes);
- reading phonics (understanding that certain letters combinations are linked to particular sounds);
- reading vocabulary (understanding that words that are read have meaning);
- reading fluency (the ability to read quickly and accurately); and
- reading comprehension (the ability to understand what is read).

Further reports in the UK and the USA (Department for Education, 2015; Slavin et al., 2009; Estyn, 2007) have provided schools with clear guidance on the use of

systematic phonics as an effective, evidence-based method for teaching reading to children between the ages of 5 and 7 years. Additionally, there are many other sources of evidence for teachers and school leaders on the effectiveness not only of phonic teaching, but on specific phonic programmes that are currently available (Slavin et al., 2009; Institute of Education Sciences, 2009; Pang et al., 2003; Education Endowment Foundation, 2016).

In spite of this accumulated evidence-based knowledge on reading instruction and the use of phonic-based instruction in primary schools (Estyn, 2007), recent studies have shown a persistent gap in the performance of pupils from different socioeconomic backgrounds in Wales (Estyn, 2016; Save the Children, 2015). The review of standards of reading in Wales commissioned by Read On. Get On. (Save the Children, 2015) drew on longitudinal data from the Millennium Cohort Study and found that the early effects of struggling to read at the age of 5 are likely to impact negatively on pupils' attainment at both the end of primary school (11 years of age) and into adult life. More significantly, the report also indicated that pupils living in persistent poverty are twice as likely to score below average for language development at the age of five. These pupils also score 22% lower on reading comprehension tests at the age of 11 compared a child who has not experienced poverty.

The reasons why some pupils fail to acquire early reading skills are difficult to specify and are likely to represent a combination of complex and interrelating socioeconomic and experiential factors. However, those pupils deemed at risk, or demonstrably falling behind, with their reading skills are likely to require some form of supplementary tuition if they are to reach their expected reading levels, especially when differentiated classroom provision has not led to improvements.

The current evidence base for supplementary reading instruction indicates that explicit, systematic, and intensive instruction in the early years for children considered to be 'at-risk' of reading difficulties can have significant and sustained effects on reading skills (Coyne et al., 2004). A teaching assistant typically delivers such supplementary instruction in UK schools in designated catch-up sessions. However, the effectiveness of teaching assistants in improving pupil progress has been found to be variable (Webster and Blatchford, 2012; Education Endowment Foundation, 2016a). Where teaching assistants have been trained to deliver focused interventions in small groups or for individual tuition, they may have moderate positive benefits. Conversely, when the work of teaching assistants lacks focus, or they are not appropriately trained to deliver a specific intervention, then there is unlikely to be positive impact on learning. In some cases, lower ability pupils may even perform less well in classrooms where teaching assistants are not appropriately deployed (Education Endowment Foundation, 2016a; Webster and Blatchford, 2012). This suggests that the quality of training teaching assistants receive to deliver effective supplementary instruction (including reading instruction) is a powerful indicator of the impact of those interventions on pupil progress.

Estyn (2011) identified deficiencies in the provision of high quality, effective reading tuition in a minority of schools. The lack of a progressive systematic phonics scheme and infrequent opportunities for regular practice were cited as characteristics of poor provision. Central to the effectiveness of a school's provision for teaching reading is their methodology for teaching beginning reading skills and/or choice of phonics programme. With the launch of the Sutton Trust teaching and learning toolkit (Higgins et al., 2012) and its successor the Education Endowment Foundation Toolkit (Education Endowment Foundation, no date) school leaders and teachers now have

access to a wealth of information on the impact of school and teaching influences on attainment. Complementary websites such as the Best Evidence Encyclopaedia (University of York, no date; John Hopkins University, no date) and What Works Clearinghouse (Institute of Education Sciences, no date) now provide schools with independent, evaluative reviews of educational programmes based on research evidence, including information on catch-up and supplementary reading programmes for struggling readers. Once a school has selected an effective supplementary reading programme, it is a challenge to ensure that struggling readers receive consistently high quality tuition from teaching assistants with the appropriate skills and knowledge (including pedagogical content knowledge, PCK) to deliver the programme with fidelity.

Computer-assisted instruction (CAI)

Computer Assisted Instruction (CAI) has become a successful intervention to teach and remediate reading skills (Fletcher-Flinn and Gravatt, 1995). A more comprehensive review of the effects of CAI's as a tool to deliver reading instruction has been published by Blok et al. (2002) and Higgins et al. (2012). Although the evidence on the positive impact of using technological approaches to learning is equivocal, there is general agreement that CAI's can be beneficial when used to deliver short, focused interventions for lower attaining and/or 'at risk' pupils as a supplementary provision alongside normal classroom teaching (Hall, Hughes and Filbert, 2000; Higgins et al., 2012a).

A CAI known as *Headsprout* is an online instructional reading programme comprised of two sequential programmes; *Headsprout Early Reading* (HER) and *Headsprout Reading Comprehension* (HRC). *Headsprout* provides a responsive technology designed to teach early reading and comprehension skills. HER is an early

literacy curriculum consisting of 80 animated episodes, designed for typically developing children between the ages of 4 and 7. The episodes are designed to teach phonemic awareness, phonics, fluency, vocabulary, and comprehension, and pupils monitor their progress through the episodes through an interactive map. Each episode is designed to last between 15-20 minutes, and pupils are encouraged to complete at least 3 episodes per week during the course of the intervention. Although HER offers pupils opportunities for cumulative review and application, teachers are also encouraged to deliver supplementary fluency practice if required. HER also provides pupils with regular stories based on the episodes to reinforce instruction provided in the lessons (Layng, Twyman and Stikeleather, 2003). Headsprout Reading Comprehension (HRC) is designed to teach children strategies necessary for success in reading comprehension. Comprising 50 animated episodes, HRC is designed for typically developing children aged 8-9 years old.

The inbuilt algorithms in *Headsprout* programmes serve to adapt instruction according to how each individual is performing. Therefore, the programme's adaptive technology provides additional instruction and error correction, and aims to achieve mastery of the skills taught in each episode.

Huffstetter et al. (2010) identified the effectiveness of *HER* online reading programme in improving the reading skills of 'at risk' pre-school pupils in the United States. Further studies by Grindle et al. (2013) and Tyler et al. (2015) also indicate that HER is an effective programme for teaching early reading skills to primary ages pupils in the UK (including pupils with autism and mainstream pupils, aged 4-7 years). The Regional School Effectiveness and Improvement Service for North Wales (GwE) is supporting the implementation of the Welsh Government's flagship education improvement plan, *Qualified for Life* (Welsh Government, 2014). The first strategic

objective of this plan is to produce '*an excellent professional workforce with strong pedagogy based on an understanding of what works.*' The report indicated that Wales has yet to fully realise the potential that digital technologies offer learners, and recommended continued support for schools in the effective use of digital technologies to develop approaches to more personalised learning. It also encouraged schools and regional school improvement consortia to create environments where teachers and educators are supported to innovate and evaluate educational practice. Following the positive outcomes reported in two *Headsprout* studies in schools in North Wales with participants aged 4-7 years (Grindle et al., 2013; Tyler et al., 2015), GwE requested support from the School of Psychology, Bangor University, to evaluate whether HER would be an effective supplementary reading programme for older primary pupils (up to 11 years) in other mainstream primary schools.

In this paper, we describe the implementation of HER in two North Wales schools with implementation support or no such support following initial training of teaching assistants. Within one school, we were able to compare HER outcome data for the children with other children in the same school not identified as in need of additional reading intervention. Thus, we were able to explore whether HER can be used to help pupils catch up with reading. We also compared reading outcomes between the two schools. This enabled us to examine whether there was evidence that additional implementation support led to improved HER outcomes.

Method

Participants and settings

The sample of pupils used in this study was drawn from two mainstream primary schools in North Wales (School A and School B). The language of instruction in

school A is predominantly English with significant use of Welsh. In School B the language of instruction is predominantly English medium with Welsh taught as a second language. Both schools use well-established phonic programmes as their main method of teaching reading.

School A enrolled 29 Year 2 pupils (6-7 years of age) to receive *Headsprout* (male = 14, female = 15). Most were either on the school's 'at risk' register for struggling readers and/or scored well below the mean score of 100 in the 2014 national reading and numeracy tests. Following the *Headsprout* placement tests (see below), all of the pupils were placed within HER. School B enrolled 35 pupils from Years 2-6 (6-11 years of age) to receive *Headsprout*; 24 pupils were enrolled in HER (male = 12, female = 12), and the remaining 11 were enrolled on HRC (male = 6, female = 5).

Settings

Both Schools A and B allocated three 30 minute morning sessions into the school week to allow pupils to access *Headsprout*. Both schools chose to run the sessions outside of the classroom. School A utilised a small storage room that was used for other withdrawal groups and set up six laptop computers to run the programme. School B had access to an information and communications and technology (ICT) suite that housed 20 desktop PC's to run the *Headsprout* sessions.

Materials and Apparatus

Materials included the *Headsprout Early Reading* (HER) programme and the sequel *Headsprout Reading Comprehension* (HRC) programme. HER is comprised of 80

online episodes that, on average, took 15-20 minutes per episode to complete. HRC comprises 50 online episodes that take approximately 20 minutes to complete. Accompanying materials to the online programmes include progress maps, fluency building resources (flash cards and words and sound sheets), printable stories and completion and progress certificates. Students were awarded stickers to mark each completed episode on their progress map. Apparatus included headphones, a computer with Internet access, and a web browser with a Macromedia Flash plug. Neither school was encouraged to share the *Headsprout* programmes with parents during the course of this study.

Measures

Intervention pupils were screened using pre- and post-test assessments outlined below.

Dynamic Indicator of Basic Early Literacy Skills (DIBELS)

HER pupils in both schools were screened using this assessment. Sub-tests taken from the DIBELS (Good and Kaminski, 2007) were used as pre- and post-intervention reading performance measures. The sub-tests included were the *Nonsense Word Fluency* (NWF) test with a focus on the number of sounds read correctly, and the number of whole words read correctly. Performance on these subtests is measured by calculating the number of correct responses from pupils during a one-minute timing. The DIBELS assessment provides multiple parallel assessment forms, conducive for repeated usage over time, thus reducing practice effects. In both schools, the research officer administered the DIBELS pre-test assessments in December 2014, with post-tests carried out in July 2015 (also by the researcher).

New Group Reading Test (NGRT)

School A administered this standardised reading test as an additional pre- and post-intervention measure to all pupils in Year 2. The NGRT assesses phonic knowledge, decoding ability, sight word knowledge and comprehension (retrieval, simple inference and writer's use of language) through sentence retrieval through sentence completion questions and a passage comprehension task (GL Assessment, 2013). Results from NGRT have been converted into age standardised scores (population mean 100, standard deviation 15). The NGRT pre-test assessments were undertaken in October 2014, with post-tests carried out in June 2015. The class teacher rather than the researcher administered both tests. Not all pupils were available for pre- and post-testing, and pupils without two data points have been excluded from this study.

Intervention and procedure

Training and support

Both schools designated a member of staff to become trained to deliver *Headsprout* and oversee the project within the respective school. These staff were released from school to attend a 1.5 hour group training session on *Headsprout* and how to implement the intervention, delivered by a research officer experienced in the use of *Headsprout*. Supplementary training was also made available for both schools individually, providing an opportunity for the research officer to see where *Headsprout* would take place, clarify any queries the member of staff had, and help to set up and schedule *Headsprout* into the school timetable. Trained members of staff were also made aware of the resources that are available to consult or download from the *Headsprout* website. A 'Helpful Tools and Tips' and a 'Getting Pupils Started'

webpage provides information on how to navigate *Headsprout* and how to get the most out of the programmes.

In addition to the initial training, schools had the choice to opt into a package of 'on-going implementation support'. This support took the form of school visits to offer advice and assistance with quality implementation, together with email and phone assistance for technical issues. School A chose to deliver *Headsprout* without implementation support from the research officer. School B chose to deliver *Headsprout* with implementation support.

Headsprout Placement-Test

The *Headsprout* placement test is a brief reading assessment that is taken prior to starting the pupils on the programme. The test, which is downloadable from the *Headsprout* website, ensures that pupils begin the programme at their developmentally appropriate level. All pupils across both schools took the test during the time allocated to take their pre-test reading assessments. The guidelines provided with the *Headsprout Placement-Test* were used to determine which starting episode the pupil should begin at.

Headsprout sessions

During a typical *Headsprout* session, pupils were collected from their classrooms and taken to the designated *Headsprout* area. Pupils sat at a laptop or desktop PC equipped with a set of headphones, and logged onto their personal accounts. The majority of the older pupils were able to log into the system and continue without the prompt or aid of the teacher/teaching assistant. Younger students

required more support from the member of staff, who would facilitate the child logging in and ensuring that the headphones were connected and positioned appropriately. Pupils were reminded not to talk to one another while they completed the episodes, but to put their hands up if they could no longer hear the instructions through the headphones or if they had a question.

Once the pupils were logged in, they followed the instructions provided by the programme. Teachers/teaching assistants awarded stickers ('stars') to those who had completed an episode during the session, and if there was enough time pupils were able to spend up to five minutes spending the stars that they had earned during the episode on their personal robot avatar.

During the training session the teachers/teaching assistants were informed that *Headsprout* continuously collects data on students' performance on each episode. For both programmes, *Headsprout* provides an overall percentage score at the end of the episode. The teachers/teaching assistants were encouraged to monitor the Performance Reports for each pupil. For the HER programme, they were advised to repeat episodes in which children scored below 90%. For the HRC programme, they were advised to monitor students more closely if they scored below 80% on three consecutive episodes. As part of the ongoing implementation support for School B, the research officer monitored the *Headsprout* data remotely. However, the teaching assistant at School B also regularly monitored the pupil's performance data. Upon identification that a pupil was struggling (i.e. scoring below 80% on three consecutive episodes) the research officer arranged to visit the school to decide whether the student would need to complete some of the targeted practice fluency building exercises (for HER) or strategy review exercises (for HRC) before continuing with the programme.

To adhere to the programme guidelines, both schools planned to deliver three *Headsprout* sessions per week, in addition to usual classroom provision. Pupils were enrolled from January 2015 to July 2015, giving an intervention period of 19 school weeks.

Approach to data analysis

This study aimed to evaluate whether HER can help children with poorer reading skills to catch up (within School A), and then whether the changes in pupils' DIBELS scores between School A and School B (using the two different implementation/support models) differed. The latter analysis addresses the question of the best way to implement HER as a catch up intervention.

In school A, a control group was established to allow performance to be compared to the HER intervention pupils. The control group consisted of all the non-HER pupils in the 2014-2015 Year 2 cohort in the school (12 pupils; male = 7, female = 5). The control group pupils followed the same taught curriculum delivered by the same class teachers, but did not receive HER intervention because they were not identified as in need of catch-up reading support. It was not possible to create a similar control group in School B due to the varying ages of the pupils receiving HER and the relatively small school cohort sizes.

Data analysis focused on mixed Analysis of Variance models (ANOVA) in which time (pre to post-intervention) was the repeated measures factor. For the first research question, the group of children (HER, non-HER) was the between subjects factor. For the second research question, school (A or B) was the between subjects

factor. In both analyses, the interaction effect with time was the focus. Partial eta squared is reported here as a measure of effect size (Pallant, 2010).

In School B, the research officer gathered qualitative field notes and observations, and both schools completed a post-study questionnaire focused on the quality of implementation fidelity and the practical arrangements for delivering HER.

Results

Headsprout delivery and episode progress

School A

All of the 29 pupils were placed within in HER; 26 beginning at episode 1, two at episode 19, and one pupil was placed at episode 41. The final evaluation NGRT data are based on the 26 pupils who commenced HER at episode 1, and who possess both pre- and post-test scores (male = 12, female = 14). The evaluation of the control group NGRT data is based on 12 pupils (male = 7, female = 5). The evaluation of DIBELS scores is based on 28 pupils with pre- and post-test scores. The 29 *Headsprout* pupils completed an average of 48 episodes (range = 39-56 episodes) during the intervention period. On average pupils from this cohort completed an average of 2.4 episodes each week over the duration of the study. The average episode accuracy across all pupils during the intervention period was 95%. (range = 88-98%; median value = 96%).

School B

The DIBELS evaluation is based on 14 HER pupils with pre- and post-test scores (male = 5, female = 9). The 24 pupils placed within HER completed an average of 26 episodes (range = 5-40 episodes). The 11 pupils placed on HRC completed an average of 22 episodes, however only six completed more than 10 episodes. Pupils in School B completed an average of 1.4 HER episodes per week. The average HER episode accuracy was 96% (range = 88-99%; median value = 97%).

Did HER help children in School A to catch up with their peers?

The mean scores for the NGRT at pre- and post-intervention for the HER catch up intervention group and the control group in School A are shown in Table 1 and in Figure 1. These data show that the group of children, identified for catch up reading intervention, were indeed performing more poorly on the reading test before intervention compared to other pupils in the same school and school year. Both groups of children improve over time, and there appears to be a narrowing of the reading performance gap post-HER intervention.

[Insert table 1 here]

These data were subjected to statistical analysis using mixed ANOVA. There was an overall significant main effect of time ($F (1, 36) = 51.1$, $p < .001$, partial $\eta^2 = 0.59$, associated with a moderate effect size – overall the children in the school improved on the reading test over time as would be expected. There was also a significant main effect of intervention group ($F (1, 36) = 24.3$, $p < .001$, partial $\eta^2 = 0.40$), also a moderate effect size – overall, the children in the catch up group had lower reading scores than the comparison group. However, there was also a significant

interaction effect (F (1, 36) = 5.2, p= .028, partial η^2 = 0.13) representing a small

effect size. This interaction effect showed that increases in reading scores over time

differed between the two groups. From Table 1 and Figure 1, it is clear that children in

the catch-up group improved more in their reading scores than their peers in the

comparison group. There was thus evidence of a catch-up effect for the poorer readers.

[Insert figure1 here]

Were better HER outcomes apparent between Schools A and B?

Mean scores for the children who received HER in both School A and School B on the

two DIBELS reading measures for pre and post-intervention are summarised in Table

2 and in Figures 2 and 3. These data suggest a steeper increase in reading scores for

children in School B – where additional implementation support was received as

described earlier.

[Insert figures 2 & 3 here]

These data were also subjected to statistical analysis using mixed ANOVA. For

the DIBELS measure of nonsense sounds read correctly in one minute, there was an

overall significant main effect of time (F (1, 40) = 55.0, p< .001, partial η^2 = 0.58),

associated with a moderate effect size – overall the children in the study improved on

this reading test over time as would be expected. There was no main effect of school

(F (1, 40) = 2.5, p= .121, partial η^2 = 0.06) – overall, the children in the two schools

did not differ significantly on this reading measure. There was also no interaction

effect (F (1, 40) = 2.8, p= .104, partial η^2 = 0.07).

For the DIBELS measure of nonsense words read correctly in one minute, there was also an overall significant main effect of time ($F_{(1, 40)} = 45.1$, $p < .001$, partial $\eta^2 = 0.53$) representing a moderate effect – overall the children in the study improved on this reading test over time as would be expected. There was also a main effect of school ($F_{(1, 40)} = 4.3$, $p = .044$, partial $\eta^2 = 0.10$), a small effect size – overall, the children in the two schools differed significantly on this reading measure. This effect has to be interpreted in the context also of a significant interaction effect ($F_{(1, 40)} = 5.3$, $p = .026$, partial $\eta^2 = 0.12$). This interaction effect was associated with a small effect size, and shows that increases in reading scores over time differed between the two schools. From Table 2 and Figure 3, there is a steeper increase in reading scores on this measure in School B where additional implementation support was provided.

[Insert Table 2 here]

Discussion

The present study provides an insight into the feasibility and practicalities of delivering the *Headsprout* online reading package as a supplementary reading instruction for struggling readers in a mainstream school setting. The outcomes of this study add to and build upon the research conducted in North Wales by Tyler et al. (2015).

This study has provided evidence that supplementary HER tuition can help children with poorer reading skills catch up with pupils not deemed to require additional support. The pupils detected to receive HER in School A were all identified as 'at risk' readers, and achieved a statistically significant improvement in their average reading comprehension (NGRT) score over time (+14.73 standard score). In comparison, the pupils in this cohort acting as a control group (readers who were not

struggling and so did not receive catch-up intervention) achieved a significant, but smaller, average improvement in their reading comprehension (NGRT) scores (+7.58 standard score). Care needs to be exercised in the overall interpretation of these NGRT data as the control pupils were selected from the residual pool of Year 2 pupils and, therefore, were a higher performing cohort of pupils at pre-testing. Therefore, their potential to demonstrate increases in reading scores similar to those achieved by the HER pupils may have been limited by a ceiling effect. In School A, the HER pupils received an average of 2.4 HER episodes each week, and this equates to an additional 50 minutes of instructional time each week. As the control pupils in School A did not receive any additional instructional time on a reading CAI (or another educationally-based CAI) it is not possible to be certain that the provision of supplementary HER or the provision of additional CAI time is the determining feature for success. However, it is promising to note that the HER pupils in School A were successful in improving their reading skills relative to the pupils not deemed to require supplementary HER provision. In this case, there is some evidence that the school's use of HER as a catch-up intervention was successful.

We also found some evidence that ongoing implementation support to assist School B produced statistically additional improvements in pupils' reading skills over and above the improvements seen in School A (with no implementation support). Although no statistically significant interaction (time x school A/B) was found for the DIBELS (correct sounds) sub-measure, the data in Figure 2 do support the pattern of findings that improvement was greater (steeper graph line) in School B. We did find a significant interaction effect for the DIBELS (correct words per minute) score,

suggesting that the improvement in School B was larger than that for School A (see Figure 3).

The descriptive data from both schools indicate it is possible to integrate *Headsprout* sessions effectively into the school week. Although children in both schools exceeded the average episode accuracy target of 90%, neither school was able to achieve the minimum advised number of three weekly episodes for each child. School A chose not to receive on-going supplementary support following the initial training; whereas School B chose to receive supplementary support in addition to the initial training and received on-going email, telephone and face-to-face contact when required. School A demonstrated that it is possible to train an elected member of staff to deliver the *Headsprout* online reading programme with minimal implementation support and achieve positive impact. In addition to the information of episodes per week and average accuracy discussed previously, a review of the field notes and school questionnaires indicates differences in the quality and consistency of HER implementation. These features may, in part, explain the difference between the reading scores. These key aspects have been summarised as follows:

- **Undertake regular benchmarked assessments.** School A did not always complete regular benchmarked assessments. School B completed these at the defined intervals and entered pupil data online.
- **Ensuring pupils read the *Headsprout* paper-based stories alongside episodes (in school and at home).** Neither school completed this task.
- **Ensuring pupils receive fluency building practice and/or repeating episodes if they are rated 'needs practice'.** School A did not always repeat episodes if 'needs practice' was indicated, nor did they offer systematic fluency practice. School B encouraged pupils to repeat episodes and offered

fluency building practice when required (although a minority of pupils were not keen to repeat episodes).

- **Ensuring the Headsprout checklist is completed after every episode, and note pupils' scores.** School A adopted an informal checklist procedure in the form of notes made by the teaching assistant. School B completed a regular and systematic checklist of pupils' scores and episode information.

- **Provide a progress wall map so help pupils visualise their progress through the episodes.** School A knew where the progress map is on the *Headsprout* website, but did not utilise this resource. School B provided each pupil with a large *Headsprout* progress map, and ensured pupils were awarded stickers as a motivational reward.

- **Awarding *Headsprout* certificates to acknowledge good progress**. School A did not award certificates to pupils. School B provided certificates to pupils during weekly assemblies. The pupils in School B responded very positively to this reward.

- **Ensure the progress wall map is populated with reward stickers after each episode, and ensure pupils access the 'star zone'**. School A allowed pupils to access the 'star zone' at the end of the project only. School B allowed pupils to access the 'start zone' as a reward for effort and achievement between episodes.

It is also worth noting that neither school has reported any significant problems with either the small number of American English spellings in HER or the computerised American accent that the programme employs. Pupils encounter a

significant amount of American culture and/or accents on television, films and on the internet, and their use on *Headsprout* programmes is not an unusual feature.

Previous studies on the impact of reading interventions have shown that explicit provision of basic reading instruction can have a positive impact on the standards achieved by 'at risk' pupils, and can accelerate reading development (Scamacca et al., 2007). More recent findings have indicated that HER has the potential to be a very effective supplementary reading programme for struggling readers (Tyler et al., 2015). The findings presented in this study support the conclusion that HER may be an effective 'catch-up' intervention to help struggling readers.

In addressing the questions in this paper, we have identified that some key differences in HER implementation may account for differences in impact between the two schools. Further research is now required to more fully assess the impact of training and fidelity of implementation on pupil outcomes. With this information on how and why implementation models matter, HER offers the potential to be a cost effective method of delivering high quality reading instruction across many schools.

References

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted
    instruction in support of beginning reading instruction: a review. *Review of
    Educational Research*, 72, 101–130.

Brown, G. T., and Hattie, J. A. C. (2012). The benefits of regular standardized
    assessment in childhood education: Guiding improved instruction and learning.
    In: Suggate, S. and Reese, E. ed. *Contemporary Debates in Childhood
    Education and Development*. London: Routledge.

Coe, R., Kime, S., Nevill, C., and Coleman, R. (2013). *The DIY Evaluation Guide.*
    London: Education Endowment Foundation.

Cook, N. L., Kretlow, A, G., and Helf, S. (2010). Supplemental reading help for
    kindergarten students: How early should you start? *Preventing School Failure,
    54,* 137-144.

Coyne, M. D., Kame'enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning
    reading intervention as inoculation or insulin: first-grade reading performance
    of strong responders to kindergarten intervention. *Journal of Learning
    Disabilities*, 37(2), 90–104.

Department for Education (2015). *Reading: the next steps. Supporting higher
    standards in schools*. London: Department for Education.

Education Endowment Foundation (no date). *Teaching and Learning Toolkit*.
    Available at: https://educationendowmentfoundation.org.uk/evidence/teaching-
    learning-toolkit/ [accessed on 17 February 2016].

Education Endowment Foundation (2016). *Phonics*. Available at:

> https://educationendowmentfoundation.org.uk/evidence/teaching-learning-
>
> toolkit/phonics/ [accessed on 17 February 2016].

Education Endowment Foundation (2016a). *Teaching Assistants*. Available at:

> https://educationendowmentfoundation.org.uk/evidence/teaching-learning-
>
> toolkit/teaching-assistants [accessed on 17 February 2016].

Estyn (2007). *Improving the learning and teaching of early reading skills*. Cardiff:

> Estyn.

Estyn (2009). *Best practice in the reading and writing of pupils aged five to seven*

> *years*. Cardiff: Estyn.

Estyn (2011). *Literacy and the Foundation Phase. An evaluation of the*

> *implementation of the Foundation Phase for five to six-year-olds in primary*
>
> *schools, with special reference to literacy*. Cardiff: Estyn.

Estyn (2016). *The Annual Report of Her Majesty's Chief Inspector of Education and*

> *Training in Wales 2014-2015*. Cardiff: Estyn.

Fletcher-Flinn, C. M., & Gravatt, B. (1995). The efficacy of computer assisted

> instruction (CAI): A meta-analysis. *Journal of Educational Computing*
>
> *Research, 12,* 219-233.

GL Assessment (2013). *York Assessment of Reading for Comprehension*. Available at:

> http://www.yarcsupport.co.uk/index.html [accessed on February 10 2016].

GL Assessment (2013a). *A Short Guide to Standardised Tests*. London: GL

> Assessment Limited.

Good, R. H., and Kaminski, R. A. (2007). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugine, OR: Institute for the Development of Educational Achievement.

Grindle, C. F., Hughes, J. C., Saville, M., Huxley, K. and Hastings, R. P. (2013). Teaching Early Reading Skills to Children with Autism using Mimiosprout. *Behavioral Interventions*, 28, 203–224.

Hall, T.E., Hughes, C. A. and Filbert, M. (2000) Computer assisted instruction in reading for students with learning disabilities: A research synthesis. *Education and Treatment of Children*, 23, 173-193.

Higgins, S., Kokotsaki, D. and Coe, R. (2012) *The Sutton Trust/EEF Teaching and Learning Toolkit*. London: Education Endowment Foundation.

Higgins, S., Xiao, Z., and Katsipataki, M. (2012a). *The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation*. London: Education Endowment Foundation.

Hill, R. (2013) *The future delivery of education services in Wales*. Cardiff: Welsh Government Consultation.

Huffstetter, M., King, J. R., Onwuegbuzie, A. J., Schneider, J. J., and Powell-Smith, K. A. (2010). Effects of a computer-based early reading program on the early reading and oral language skills of at risk preschool children. *Journal of Education for Students Placed at Risk*, 15, 279–298.

Institute of Education Sciences (no date). *What Works Clearinghouse*. Available at: http://ies.ed.gov/ncee/wwc/default.aspx [accessed on 13 February 2016].

Institute of Education Sciences (2009). *What Works Clearinghouse: Phonics*.

>   Available at: http://ies.ed.gov/ncee/wwc/default.aspx [accessed on 13 February

>   2016].

Institute of Education Sciences (2010). *Effectiveness of Selected Supplemental*

>   *Reading Comprehension Interventions: Findings From Two Student Cohorts.*

>   Alexandria, VA. Available at:

>   http://ies.ed.gov/ncee/pubs/20104015/pdf/20104015.pdf [accessed March 10

>   2016].

John Hopkins University (no date) *Best Evidence Encyclopedia*. School of Education

>   Center for Data-Driven Reform in Education (CDDRE), John Hopkins

>   University. Available at: http://www.bestevidence.org/index.cfm [accessed on

>   March 10 2016].

Johnson, K. R., & Layng, T. V. (1994). *The Morningside Model of Generative*

>   *Instruction*. In: Gardner, R., Sainato, D.M., Cooper, J.O., Heron, T.E., Heward,

>   W.L., Eshleman, J.W., and Grossi, T. ed. *Behavior analysis in education:*

>   *Focus on measurably superior instruction.* Belmont, CA: Brooks/Cole.

Johnson, K. R., and Street, E. M., (2004). *The Morningside Model of Generative*

>   *Instruction: What It Means to Leave No Child Behind.* Concord, MA:

>   Cambridge Center for Behavioral Studies.

Kamil, M. L. (2003). *Adolescents and Literacy: Reading for the 21st Century. Alliance*

>   *for Excellent Education.* Available at:

>   http://www.all4ed.org/files/archive/publications/AdolescentsAndLiteracy.pdf

>   [accessed on 15 February 2016].

Kucirkova, N. (2014). iPads in early education: separating assumptions and evidence. *Frontiers in Psychology,* 5, 715. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4085648/ [accessed on 17 February 2016].

Layng, T. V. J., Twyman, J. S., and Stikeleather, G. (2003). Headsprout early reading: reliably teaching children to read. *Behavioral Technology Today*, 3, 7–20.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*, Washington, DC: National Institute for child health and development.

OECD (2010) *PISA 2009 Results: Executive Summary*. Paris: Organisation for Economic co-operation and Development.

Pallant, J. (2010) *SPSS Survival Manual (4th edition): A Step by Step Guide to Data Analysis using SPSS (Version 15).* Maidenhead: Open University Press.

Pang, E. S., Muaka, A., Bernhardt, E. B., and Kamil, M. L. (2003). *Teaching Reading. International Academy of Education.* Geneva: International Bureau of Education. Available at: http://unesdoc.unesco.org/images/0013/001313/131370e.pdf [accessed on February 16 2016]

Pitchford, N. (2014). *Unlocking Talent: Evaluation of a tablet-based Masamu intervention in a Malawian Primary School*. Available at:

https://onebillion.org/downloads/unlocking-talent-final-report.pdf [accessed

on 13 February 2016].

Save the Children (2015). *Ready to read. Closing the gap in early language skills so

that every child in Wales can read well*. Available at:

http://www.savethechildren.org.uk/resources/online-library/ready-read-wales

[accessed on 15 February 2016].

Scammaca, N., Vaughn, S., Roberts, G., Wanzek, J. and Torgesen, J. K. (2007).

Extensive reading interventions in grades k -3: From research to practice.

Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Slavin, R. E., Lake, C., Chambers, B., Cheung, A. and Davies, S. (2009). *Effective

Beginning Reading Programs: A Best-Evidence Synthesis*. School of Education

Center for Data-Driven Reform in Education (CDDRE), John Hopkins

University. Available at:

http://www.bestevidence.org.uk/reviews/begin_reading/index.html [accessed

on 17 February 2016].

Tyler E. J., Hughes, J. C., Beverley, M. and Hastings, R. P. (2014). Improving early

reading skills for beginning readers using an online programme as

supplementary instruction. *European Journal of Psychology of Education*, 30

(3), 281-294.

University of York (no date). *Best Evidence Encyclopaedia*. The Institute for Effective

Education, The University of York. Available at:

http://www.bestevidence.org.uk/index.html [accessed on 17 February 2016].

Vaughn S. and Fuchs L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice*, 18, 137–146.

Webster, R. and Blatchford, P. (2012). Supportive Learning?: How effective are teaching assistants? In: Adey, P. and Dillion, J. ed. *Debunking Myths in Education*, Maidenhead: Open University Press.

Welsh Government (2013). Achievement of 15 year olds: Program for International Student Assessment (PISA) National Report . Available at: http://gov.wales/statistics-and-research/achievement-15-year-olds/?lang=en [accessed on 17 February 2016].

Welsh Government (2013a). *National Literacy and Numeracy Framework*, Cardiff: Welsh Government.

Welsh Government (2014). *Qualified for Life*, Cardiff: Welsh Government.

Welsh Government (2015). Effective use of data and research evidence, *Cardiff: Welsh Government*.

Table 1. Changes in NGRT standard reading scores in school A (see text for explanation).

Intervention type	Number of pupils (n)	Reading Standard Score			
		Pre-test		Post-test	
		Mean (M)	Group standard deviation (SD)	Mean (M)	Group standard deviation (SD)
2014-15 HER pupils	26	86.27	15.71	101.00	11.82
2014-15 non-HER pupils (control group)	12	108.92	5.99	116.50	4.50

Table 2. Changes in DIBELS scores in school A and school B (see text for explanation).

| DIBELS sub-tests	School A n = 28				School B n = 14			
	Pre-test		Post-test		Pre-test		Post-test	
	Mean (M)	Group standard deviation (SD)	Mean (M)	Group standard deviation (SD)	Mean (M)	Group standard deviation (SD)	Mean (M)	Group standard deviation (SD)
Nonsense words (correct sounds) per minute	44.14	17.69	60.57	11.61	31.29	17.72	57.21	27.06
Nonsense words (correct words) per minute	16.46	8.30	21.18	4.95	9.57	5.97	19.21	9.85
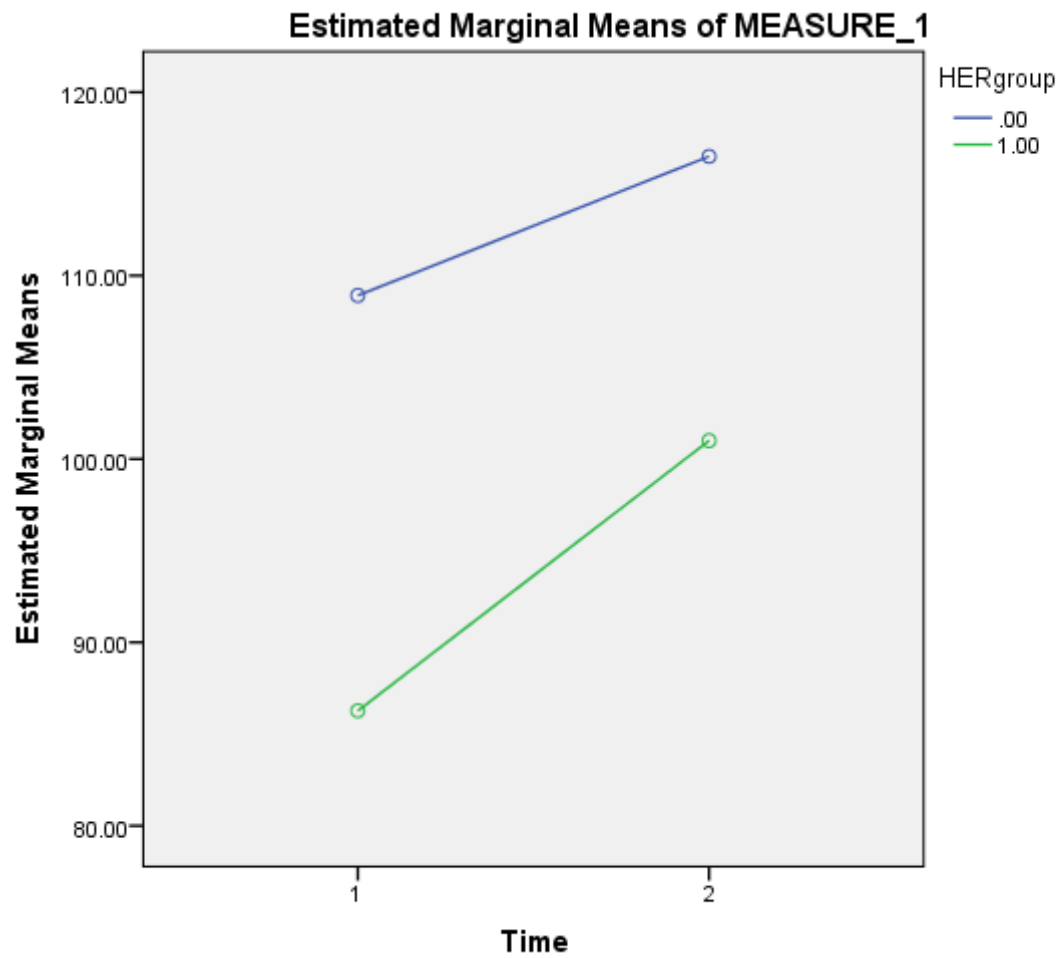
Figure 1. Mean NGRT scores between HER pupils (lower line) and control group pupils (upper line) in school A.
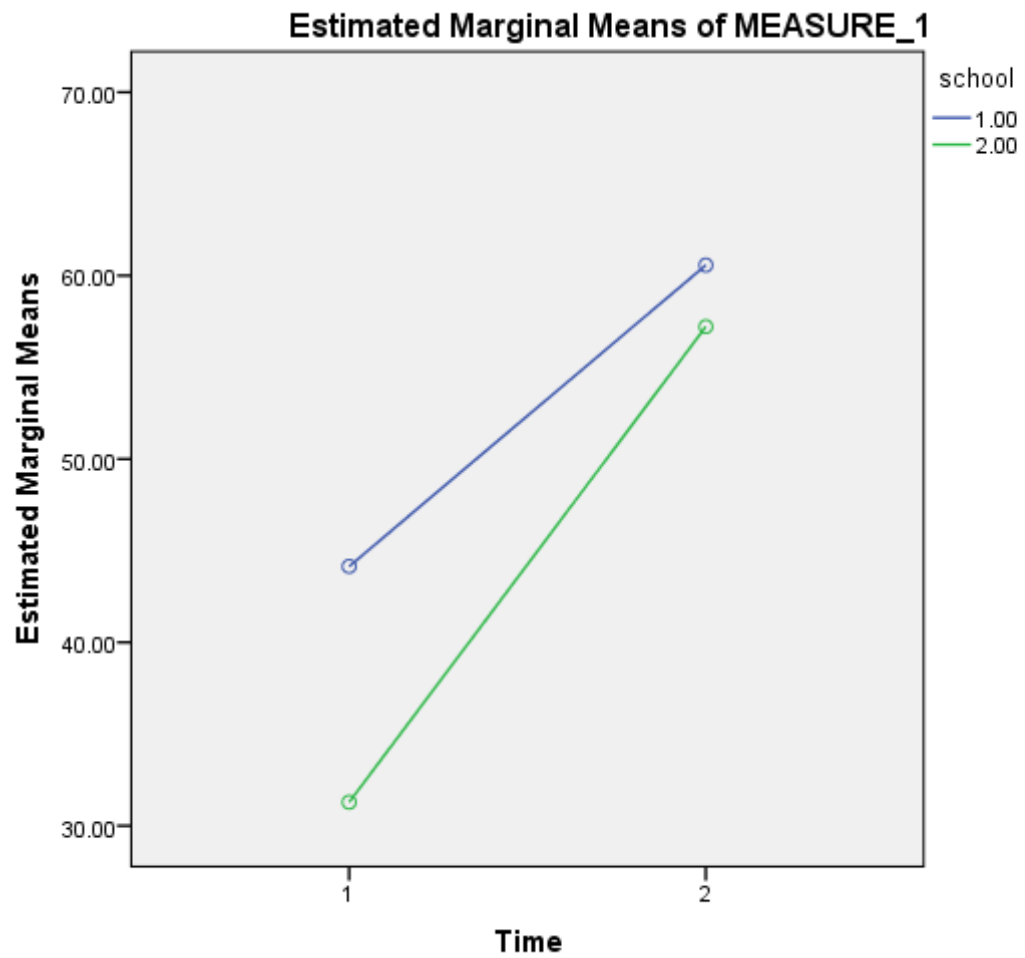
Figure 2. Mean scores for DIBELS (correct sounds per minute) between school A (upper line) and school B (lower line).
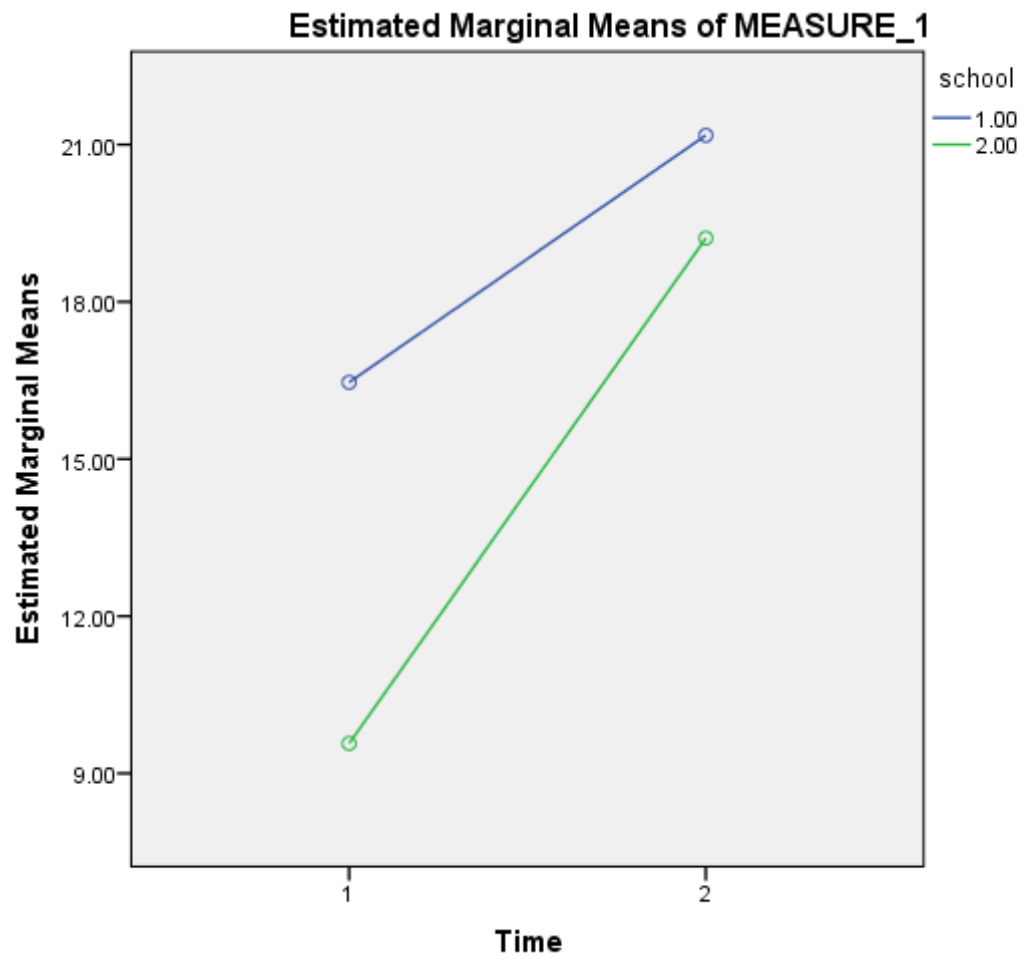
Figure 3. Mean scores for DIBELS (correct words per minute) between school A (upper line) and school B (lower line)