




**Original citation:** Price-Mohr, R. and Price, Colin  (2019) *A Comparison of Children Aged 4–5 Years Learning to Read Through Instructional Texts Containing Either a High or a Low Proportion of Phonically-Decodable Words*. *Early Childhood Education Journal*. ISSN Print: 1082-3301 Online: 1573-1707 (In Press)

**Permanent WRaP URL:** <https://eprints.worc.ac.uk/id/eprint/8868>

### Copyright and reuse:

The Worcester Research and Publications (WRaP) makes this work available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRaP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

### Publisher's statement:

This is a post-peer-review, pre-copyedit version of an article published in *Early Childhood Education Journal*. The final authenticated version is available online at:

<https://doi.org/10.1007/s10643-019-00970-4>

### A note on versions:

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRaP URL' above for details on accessing the published version and note that access may require a subscription.

**For more information, please contact [wrapteam@worc.ac.uk](mailto:wrapteam@worc.ac.uk)**

High-percentage decodable text versus low-percentage decodable text

## **A Comparison of children aged 4-5 years learning to read through instructional texts containing either a high or a low proportion of phonically-decodable words**

### **Introduction**

In this paper, we present the results from a highly novel approach to comparing instructional reading texts. Two sets of twelve books were purposely written and illustrated for this experiment and additional novel teaching activities developed including original board games. To our knowledge this has not previously been attempted.

Historically, and for more than a century, the writing of instructional reading books for beginners has been based on the assumption that vocabulary should be controlled. Reading books with texts that offered repeated practise with a small set of decodable words has long been thought to assist word reading abilities (Hoffman, Sailors & Patterson, 2002). The arguments given for using decodable readers in the early stages of reading instruction include: the opportunity for children to practise what they have been explicitly taught; early independence; and early success (Castles, Rastle & Nation, 2018). These decodable readers are designed to provide reading material in which specified phonemes can be practised (Wilson & Colmar, 2008). The text is therefore carefully controlled, only to contain phonemes that children should already be familiar with. However, others have argued against the development of reading schemes that use a contrived text. As far back as 1908, Edmund Huey expressed his concerns about the kind of language found in contrived text: “The language used often shows a patronizing attempt to ‘get down to the child’s level’, and results in a mongrel combination of points of view and of expression that is natural neither to an adult nor to a child” (Huey, 1908, p. 279). By contrast, the work of Kenneth Goodman in the 1960s promoted the view that reading should be more about comprehension of an author’s meaning, than a process of recognising letter strings in words (Pearson, 2004).

In 2005 in the UK, the House of Commons Education and Skills committee (House of Commons, 2005) commissioned a review of the methods used to teach reading, and at the same time recommended that

## High-percentage decodable text versus low-percentage decodable text

research be commissioned to compare the use of decodable instructional text with text that goes beyond a child's current decoding ability. The review of teaching methods was duly carried out (Torgerson, Brookes & Hall, 2006), but to our knowledge, the research to compare instructional texts was never commissioned. Subsequent policy decisions were based on existing reviews and consultation. New reading schemes were written to comply with the UK government's insistence that only words that can be sounded out should be introduced in the beginning stages of learning to read (Marshall, 2011). This was based on the recommendations of the House of Commons Education Select Committee (2005) which states that:

“Children are only taught to read through texts fully within their current phonological ability. So, although children might encounter words they do not understand, they are not given texts they cannot decode and are therefore not expected to infer words from context or syntax.”  
(HMSO, 2005, p. 14)

In New Zealand, the method of choice for many years from the 1980s has been the whole-word, real-books approach, endorsed by consecutive governments, as a result of the perceived success of the Reading Recovery programme (Solar & Openshaw, 2007). The whole language/real-books remedy for reading failure remains in place, while phonics has been side-lined. By contrast, in 1998, California prohibited the use of books where children could use contextual cues, although, in the US National Reading Panel report (NICHD, 2000) there was no explicit support for decodable text (Pearson, 2004). In 2000, the state of Texas introduced new rules for choosing texts for beginning readers. These were required to be decodable according to specific rules: each word had to be decodable (a word was considered to be decodable if all the letter-sound associations in that word had been previously explicitly taught); words had to follow a specified sequence of learning. In this instance, texts were to be analysed according to the number of rimes, repetition of high-frequency words, and the density of the text (Hoffman et al, 2002). More recently in the US, the Common Core State Standards (NGACBP, 2010) have raised the importance of learning the skills of inference generation and reading for comprehension; there is accumulating evidence of the benefit of building up topic-related vocabulary knowledge that is likely to be encountered in reading texts (Hall, 2016).

## High-percentage decodable text versus low-percentage decodable text

The Common Core State Standards challenge the US tradition of matching texts to children's instructional reading levels and demand an increase in text complexity although there appears to be no clear consensus regarding the definition of text complexity (Amendum, Conradi & Hiebert, 2018).

The core criteria that the Department for Education in the UK (DfE, 2010) issued for publishers of early reading books include that: "children (...) practise reading using texts which are entirely decodable" (DfE, 2012, p. 1). The National Curriculum in England Framework document (DfE, 2013), statutory since September 2014, describes in detail the requirements for the reading books that are expected to be used, including that children are: "[...] supported by practising their reading with books consistent with their developing phonic knowledge" (p. 18); and "[...] that do not require them to use other strategies to work out words" (p. 20). The draft inspection framework due to be rolled out in England in September 2019 includes the statement that inspectors will "consider how well staff teach children to read systematically by using synthetic phonics and books that match the children's phonics knowledge" (Ofsted, 2019, p. 74). In Australia there is currently a call, amongst some, to promote the use of decodable books on the grounds that 'Decodable books allow children to practise their decoding skills in a methodical way' (Victoria Liberal, 2018).

Decodable texts that are typically produced emphasise common letter-sound-correspondences and high-frequency words within simple sentences and basic story lines, as well as showing some lesson-to-text match (Brown, 2000; Mesmer, 2009). The rationale is that this will reinforce students' current alphabetic knowledge and increase word identification (Beverly, Giles and Buck, 2009; Mesmer, 2009). The assumption is that texts that follow regular letter-sound-correspondences will be read more easily. It has been suggested in the past that reading material should be readable with at least 98% accuracy (Betts, 1946 cited in Allington, 2013).

During the 1970s, there was a movement to promote the use of 'authentic' texts, written by well-established children's authors, in which the content is neither constrained nor restricted in its use of phonics, or vocabulary. For those who promoted the use of 'authentic' texts as instructional readers, the expectation

## High-percentage decodable text versus low-percentage decodable text

was that children would develop skills through repeated exposure to words, in a motivating context (Coles, 2004; Goodman, 1976; Smith, 1973). However, organising these texts into a learning sequence was both difficult and time consuming, and there was a risk that young readers could find themselves faced with text that comprised inaccessible words and over-complex sentence structures.

A number of studies and reviews have found no significant evidence of positive effects from using decodable reading texts (Slavin, Lake, Chambers, Cheung & Davis, 2009; Jenkins, Peyton, Sanders & Vadasy, 2004). There is also some evidence that decodable texts can be less motivating to read (Coles, 2004; Mosely, 2004; Solity & Vousden, 2009). In a small study of 32 young readers in the 1960s, children who were already reading fluently before starting school were questioned about their strategies when encountering unknown words (Clark, 1976). Clark suggests that the evidence from these children indicates that the complexity of a text may, in fact, aid understanding. She goes on to suggest that an apparently simplified text may make the reading material not only less stimulating, but also more difficult to comprehend. Reflecting back on her work, Clark concludes that complexity of text does not necessarily make it more difficult to read, and that simplicity does not necessarily make it easier (Clark, 2014). More recent research indicates that decoding instruction that focuses on the ability to sound out a word, can lead to successful word reading without comprehension (Michaud et al, 2017). Below are two examples of text from decodable readers that demonstrate how this may be challenging for young readers to comprehend:

“Sid and Nick plot to stop back at the long ship at the dock.”

(Greene & Wood, 2000)

“Tab the cat has had a kitten. “Tab! Tab! I’ve got fish and milk,” calls Meg.”

(Munton, 2006)

The authors of a two-year study of fluency-oriented reading instruction (Stahl & Heubach, 2005) found that when children, including struggling readers, were given greater support and exposure to vocabulary prior to reading they were able to read texts of much greater difficulty than expected. A more

## High-percentage decodable text versus low-percentage decodable text

recent review has suggested that it may be more effective to scaffold more complex texts rather than to use easily decodable texts (Fisher & Frey, 2014). Moses and Kelly (2018) draw attention to the potential negative impact of a skills-based approach to reading associated with decodable books; that it can lead to children perceiving reading as a set of skills to perform rather than an enjoyable activity. They found that first-grade children chose books to read that were above their expected grade level when they were interested in the content.

A review of the research conducted in 2012 (Cheatham & Allor, 2012), found only seven studies where text type was the sole independent variable, of which only three were intervention studies. Their conclusions were that results: were inconsistent and inconclusive across the studies; demonstrated that levelled and predictable texts were read more fluently than decodable text; and that the use of decodable text aids readers in grasping the alphabetic principles. There was, however, no evidence that decodable text supported comprehension; comprehension was measured in only one of the studies. In a subsequent study, Allor, Giffod, Al Otaiba, Miller and Cheatham (2013) re-wrote texts that included familiar vocabulary, high frequency words and natural syntax. They found that the children progressed more quickly with these texts than the decodable texts. A follow up study of students with special needs evidenced the benefit of using texts with natural sentence structure and prior learning of sight vocabulary (Allor et al, 2018).

Following a more recent review of the evidence, Castles et al (2018), whilst accepting that decodable readers may be appropriate for the earliest stages of reading, conclude that decodable books have little benefit beyond the initial stage of reading and that any benefits may be outweighed by their limitations, stressing the importance of exposure to complex words and sentence structure. Additionally, Goswami et al (2016) report findings that show children with dyslexia having greater difficulty with words of similar prosody compared to age-matched controls, due to impaired short-term/working memory. It is possible that controlling the text, with the intention of simplifying the vocabulary, may result in many of the words having a similar prosody or stress pattern, thereby making it more difficult to comprehend for some.

## High-percentage decodable text versus low-percentage decodable text

A recent study of beginner readers learning irregular words as sight words found evidence suggesting that children using a ‘hierarchical’ (recognising and using letter groups) rather than a ‘sequential’ approach to decoding were likely to be more successful. They go on to suggest that it would nevertheless be useful to continue to use decodable texts until children have learned spelling patterns and rules (Murray , McIlwain, Wang, Murray & Finley, 2019).

### **Theoretical Framework**

The literature remains unclear as to the justification for the continued emphasis on decodable texts for beginner readers; there is a paucity of empirical evidence making direct comparisons between more and less decodable texts. Effectiveness may depend on the level of integration with the method of instruction with whichever text type is used. There is clearly a need to consider potential confounding variables in any comparison of instructional texts, such as instructional method. Therefore, one of the aims of this study was to create ideal conditions for a direct comparison, whereby teaching methods, materials, and the teacher, were all sufficiently controlled that it would be possible to explore the central question of vocabulary.

The research question asked how Intervention B (high-phonically decodable vocabulary) compared with Intervention A (low-phonically decodable vocabulary) on specific measures of word reading and comprehension. Primary outcome measures, chosen for this trial, were word recognition, phoneme awareness and passage-reading comprehension. Secondary outcome measures were constructed for the process evaluation.

### **The study**

The study involved using books-A and books-B in three primary school Reception classes (age 4-5 years) to teach children to read; there were 12 books of each type. Importantly, a corresponding set of board games and other activities was developed to support the use of both types of books. Children played the games and completed the activities before they read the books. Teaching assistants delivered the lessons in two schools; the class teacher opted to deliver the lessons in one school. In this section we consider the design of the

## High-percentage decodable text versus low-percentage decodable text

books, the design of the games, the design of the study, training provided to the teacher and Teaching Assistants, and finally the participants involved, and the research instrument used.

### **Materials: Books**

Predictable text was used in both books-A and B, enabling the use of vocabulary in context. Selection of the vocabulary for books-B was made first as this had to conform to a pre-determined sequence of grapheme-phoneme correspondences (GPCs) outlined in the English National Curriculum (DfE, 2014). More complex GPCs were also introduced earlier than recommended, to reflect the number of new words being introduced for each book. Irregular and high frequency words were prioritised according to the ‘key words to literacy’ list (McNally & Murray, 1962).

One set of twelve books was provided for each intervention group for each of the participating classes (twenty-four in total). Books-A and B contained the same illustrations and storyline and introduced the same number of new words in each book, the same length of sentences and number of pages. The early books, (in the sequence one to twelve) had the same total number of words, but this was relaxed for books later in the sequence to avoid compromising the authenticity of the storyline. However, there were no more than ten words difference between books-A and B later in the sequence. Full-colour, full-page illustrations were used in the books to aid inference generation (Arizpe & Styles, 2003; Kelly & Moses, 2018; Kachorsky, Moses, Serafini & Hoelting, 2017; Lacey et al, 2007; Walsh, 2003). All the storybook characters were original designs, having similarities with cartoon characters found in popular culture. The characters had no obvious gender and by setting the stories in a woodland habitat, cultural issues were minimised. The advantage of writing original books was that we could ensure the percentages of decodable and non-decodable words were the same in both book types and for each book in sequence.

Each book and its associated activities introduced twenty new words. We calculated that the percentage of words that were common to books A and B was 38.7%. Of the remaining words, for books B, there were 12.2% non-decodable words that were included as high frequency words children were expected



## High-percentage decodable text versus low-percentage decodable text

to learn at this stage and for books A there were 64.2% non-decodable words. A sample from each text is shown below. This sample from text A has fewer words than text B; however the number of syllables differs by just one. The books overall had the same number of words ( $\pm 10$ ).

Sample of Text B (high-percentage decodable):

“Zon can see a man in a hat. The man looks very big.”

Corresponding sample from Text A (low-percentage decodable):

“Zon thinks the scarecrow is a monster. He is frightened.”

An analysis of the complete text of books 1-12 of each series (A and B) resulted in a clear difference of Flesch-Kincaid Grade Level: For the B series, a grade level of 0.308; for the A series, a grade level of 1.266.

### **Materials: Games**

As mentioned above, two parallel sets of games were also developed for interventions A and B. The words used in the games corresponded to the books. In England, the Reception class is still within the Early Years sector and many of the children were still only four years of age in the September in which they started school. Therefore, learning materials were designed to be informal play-based group activities; for children, play is an appropriate learning medium to engage in (Stephen, 2010). The games were a mix of traditional (e.g. snakes and ladders and snap) and purposely developed board and card games (e.g. ‘Pento’ and ‘Word Families’) that supported the development of manipulative skills and fine motor coordination. Vygotsky highlighted that children can learn not only from being taught, but from discovering skills in play situations (Vygotsky, 1978). This kind of learning has been described as ‘tangential learning’, whereby children learn through being covertly exposed to knowledge, rather than being actively taught in a structured way; a mechanism by which a child can learn, by being exposed to things in a context in which they are already highly engaged (Nahachewsky, 2013). The children were exposed to the new vocabulary for each book simply through playing games and not being directly taught.

## High-percentage decodable text versus low-percentage decodable text

The games were designed to be played in pairs or groups, and there were clear instructions to teachers to allow the children to assist each other, and discuss strategies for success. All the games relied on visual recognition, and were intended to encourage word recognition at the orthographic level, e.g. by using words with high neighbourhood density - such as ‘hat’ and ‘hot’ – that have been found to have a far greater relationship to sight vocabulary development than high frequency words (Stokes, Kern & DosSantos, 2012). A number of the activities were designed to direct the attention to onset-rime or initial letter positions. During the first term in which the programme began, activities focused on word recognition. The second and third term resources continued to build word recognition skills, but also had an additional emphasis on comprehension, (for a detailed description see Price-Mohr, 2016, chapter 3).

### Design

A split-cluster design was chosen (Hutchison & Styles, 2010). By having two arms (books A and books B) of the Intervention in the same classroom, it was possible to control for confounding factors such as the teacher and teaching style. There was a small risk of contamination across conditions and possible teacher bias; this was addressed in the training and monitoring. Each of three classes was randomly divided into two groups of equivalent numbers (see Table 1 below).

**TABLE 1** *Showing the two arms of the trial*

<b>Arm 1: Intervention A</b>	<b>Arm 2: Intervention B</b>
18 children	18 children
(randomly selected from each of three classes)	(randomly selected from each of three classes)
School 1: 5 children	School 1: 4 children
School 2: 6 children	School 2: 6 children
School 3: 7 children	School 3: 8 children

## High-percentage decodable text versus low-percentage decodable text

The Reading Programme was designed to run for three terms, following the first round of assessments and the teacher training, finishing before the final round of assessments. It was anticipated that each child would spend at least ninety minutes per week on Programme-related activities. This included a mix of whole-class and group work designed for groups of up to four children. Teachers were given the freedom to choose how they timetabled the sessions.

Random allocation of children to each arm of the Intervention was conducted by the Data Manager at the University of York; the researcher was kept blind to the allocation. Meetings with teachers were arranged at two time points: mid-way through the second term, and mid-way through the third term. At both of these time points there was a discussion regarding progress, and observations of a teaching session to monitor compliance to programme protocol. Monitoring of implementation fidelity was achieved through regular visits to the schools. Further information on implementation fidelity was collected at post-intervention. This included individual child reading logs and session-completion records.

### **Training**

With the intention of controlling for teaching method, all programme deliverers were given training with the learning resources and the desired approach. This was delivered in group seminars, with workshop activities, to permit feedback and questions (Wenz-Gross & Upshur, 2012). Participants in the training were class teachers for Reception and Year 1, plus Teaching Assistants and SENCOs (special needs coordinators). The training was designed to be accessible to both less experienced Teaching Assistants and experienced teachers. It was emphasised that the order of activities were first play, then reading each associated book. The actual reading of books was on a one-to-one basis with the teacher or TA.

### **Participants**

Three rural primary schools volunteered to participate in the trial. They were all single-form entry. All participating classes had Teaching Assistants available. All schools were mainstream state schools that would normally follow the National Curriculum for England. In one school the class teacher elected to

## High-percentage decodable text versus low-percentage decodable text

deliver the programme herself. A diagram of the flow of children through the trial is shown below in Figure 1.

- Fig 1 here -

The whole cohort of each Reception class from each school was recruited. The sample size (36 at pre-intervention) was calculated to have 80% power to detect a minimum effect size of 0.96. There were three Teaching Assistants participating. The number of additional helpers (part-time Teaching Assistants) varied from one to two, but they all received the same training. There were three participating teachers in the Intervention, who all participated in the initial training, in order to understand the programme and be able to take over the delivery if necessary. Class teachers were not expected to participate directly. Ethical approval was granted by the University of York Education Ethics Committee.

### **Measures and Data Collection**

Measures used included the British Picture Vocabulary Scale (Dunn et al, 2009), and Early Word Reading, Sound Deletion, Sound Isolation and Passage Reading Comprehension from the York Assessment of Reading for Comprehension (YARC) (Snowling et al, 2009). Measures are reported as standardised scores to control for age. Measures were chosen to assess phonic knowledge, word recognition and comprehension. The Passage Reading Comprehension measure was only administered at post-intervention as very few children had sufficient word recognition at pre-intervention. The YARC has a beginner passage that is a shared reading task that made it possible to measure comprehension at what is a very early stage of learning to read. The Passage Reading measure is designed to make a distinction between literal understanding, and comprehension by inference. The YARC was standardised in 2008, in the UK using a sample of 1,376 children from ten different regions.

Oral vocabulary has been found to be a strong predictor of semantic learning, and increases in oral vocabulary have been found to have a beneficial effect on reading comprehension (Clarke, Snowling, Truelove & Hulme, 2010; Protopapas, Mouzaki, Sideridis, Kotsolakou & Simos, 2013; Ricketts, Bishop,

## High-percentage decodable text versus low-percentage decodable text

Pimperton & Nation, 2011). Therefore, an assessment tool was included to measure receptive vocabulary. The British Picture Vocabulary Scale (BPVS-III) (Dunn et al, 2009) is norm-referenced and designed to measure the receptive (understood) vocabulary of children aged between 3 and 15 years using pictures. By using pictures, children do not require any pre-existing knowledge or reading skills. The BPVS III has been standardised in England using a sample of 3,278 children. Assessments were administered pre-intervention to establish baseline performance, and at post-intervention.

## Results

### Pupil Characteristics (at post-intervention)

There were no children with English as an additional language. There was no significant association between the balance of gender in each arm of the trial as calculated using Pearson's chi-square:  $\chi^2 (2) = 0.42$  (less than the critical value of 5.991 and therefore not significant at .05

Table 2 below shows the number of boys and girls for each arm of the trial.

**TABLE 2** *Distribution of gender*

Gender	Intervention B		Intervention A	
	Number	Percentage	Number	Percentage
Male	10	56	11	65
Female	8	44	6	35
Total	18	100	17	100

### Comparison of Texts

For each outcome we used a between-groups ANCOVA (analysis of covariance), in order to assess differences in post-intervention means after accounting for pre-intervention values, thereby controlling for pre-intervention differences. We report effect sizes, as well as statistical significance ( $p$  values), to provide the relative importance of an effect that  $p$  values cannot do (Nuzzo, 2014; Baker, 2016).

## High-percentage decodable text versus low-percentage decodable text

Because of the small sample size, and the exploratory nature of the study, according to Maxwell & Delaney (2008), the level of statistical significance need only be  $p < .10$  to indicate trends when sample size and power are limited and the purpose is to establish evidence upon which further larger-scale and more rigorous studies can be based. Where there are no significant differences, trends are reported which is considered to be a legitimate approach (Torgerson & Torgerson, 2008). Pre-intervention and post-intervention scores for all measures used are detailed in Table 3 below. Group means, standard deviations, and ANCOVA for each arm of the trial, for each test are shown.

**TABLE 3** *Standardised scores for group means (standard deviations) for conditions A and B*

Condition	Pre-intervention (standard scores)					Post-intervention (standard scores)					ANCOVA
	A (less)		B (more)		Effect size	A (less)		B (more)		Effect size	
Test	M	SD	M	SD	<i>d</i>	M	SD	M	SD	<i>d</i>	<i>p</i>
BPVS	106.94	10.10	103.05	9.39	0.39	105.29	8.49	99.94	10.17	0.57	.171
EWR	114.88	13.93	106.94	13.36	0.58	112.23	10.57	104.22	10.96	0.74	.051
SI	121.47	10.53	113.94	14.85	0.58	120.29	10.18	111.33	9.53	0.90	.069
SD	110.17	16.89	99.50	19.66	0.58	113.41	8.88	107.88	13.57	0.48	.477
PRC	-	-	-	-	-	112.52	10.33	103.05	9.31	0.96	.029

(EWR as covariate)

*Note.* Intervention A:  $n = 18$ ; intervention B:  $n = 18$ . BPVS = British Picture Vocabulary Scale; EWR = Early Word Reading; SI = Sound Isolation; SD = Sound Deletion; PRC = Passage Reading Comprehension

A preliminary analysis evaluating the homogeneity of regression assumption indicated that ANCOVA could be run for all reported measures. For the Passage Reading Comprehension measure (using Early Word Reading pre-intervention scores as covariate) there was a significant effect of the condition after

## High-percentage decodable text versus low-percentage decodable text

controlling for the covariate:  $F(1,32) = 5.16, p = .029, \eta_p^2 = .138$ . Coupled with a large effect size of 0.96 (Cohen's  $d$ ), valid for this sample size, this result suggests that a higher percentage of non-decodable words in predictable text can lead to significantly better outcomes in terms of reading comprehension.

For the Early Word Reading measure, the results were close to significance:  $F(1,32) = 4.11, p = .051, \eta_p^2 = .114$  with a medium effect size ( $d = 0.74$ ). This indicates a trend suggesting that a higher percentage of non-decodable words in instructional texts can also lead to better outcomes in terms of word recognition. Interestingly, the Sound Isolation measure, although not statistically significant:  $F(1,32) = 3.51, p = .069, \eta_p^2 = .099$ , showed a large effect size ( $d = 0.90$ ), indicating a trend that suggests that this element of phoneme awareness can be supported by a higher percentage of non-decodable words in texts. There were no significant observable effects or trends for the other phoneme awareness measure (Sound Deletion) or for the receptive vocabulary (British Picture Vocabulary Scale).

Fidelity of implementation and process was assessed at post-intervention; due to the split-cluster design, variation in fidelity was equal across the two intervention conditions. The quantity of resources made available and the regular support and monitoring led to high levels of implementation. Data collected from schools did evidence that higher levels of fidelity led to higher word-reading outcomes for the programme books.

## Discussion and Conclusions

Our results demonstrate a statistically significant difference for comprehension between the two text types and a large effect size in favour of books A that have been written to contain a low percentage of decodable text. In addition, there were observable trends indicating differences for Early Word Reading and Sound Isolation after controlling for pre-intervention scores. Books B, that contained a large proportion of decodable text were inferior in teaching children to read with understanding.

These results challenge long-held assumptions about using decodable books to support early independence and success. They challenge the idea that books should be used to practise specified letters or

## High-percentage decodable text versus low-percentage decodable text

letter groups and the view that decodable books lead to increased word identification and are simply easier to read. Rather, we feel that our results offer support for the Common Core State Standards (NGACBP, 2010) that raise the importance of reading for comprehension and developing inference generation. Additionally, we feel that these results support the view that children can be motivated to read above their expected (mandated) grade level when the content is enjoyable and gain more understanding from books that are not restricted to a vocabulary that is consistent with their developing phonic knowledge, thus challenging current policy in England, as well as other national policies that advocate using only decodable books.

These results highlight the potential mismatch between research and policy that can impact on practitioners. Concern has been expressed by some that often those researching in the area of education have rarely and sometimes never, been in a classroom or talked with teachers, and that global conclusions are based on localised evidence (Tesar, 2019). The drive towards a more standardised-based approach to teacher education, in which ‘what teaching is’ and ‘what teachers should know’ is made explicit (Menter, Valeeva & Kalimullin, 2017), alongside the increase in a culture of assessment and accountability, may be at the root of this mismatch. Concern has also been expressed about the increasing politicisation of the education agenda whereby research conclusions appear to be misunderstood as beliefs rather than understood as well-established knowledge (Pring, 2015). Estelle Morris (2012) has expressed similar concerns, however she also reminds researchers of the importance of reflecting educational changes and the needs of schools and policy makers; that they should contribute to the standards agenda.

In summary, these results add to the literature suggesting that decodable texts are of less value for children as they become more skilful at reading (Castles et al, 2018), but go further by demonstrating that decodable texts are of less benefit, in terms of comprehension, even for beginner readers. This clearly has implications for policy makers, practitioners and publishers of early reading texts.

There are some potential limitations to the study. The first was the risk of cross contamination, since by having both sets of books and resources in the same classrooms, they could become muddled. This was



## High-percentage decodable text versus low-percentage decodable text

mitigated by training, and by indicators on the rear cover of books and different coloured packaging for the associated activities. The small sample size reduced the power of the design, however, for a sample of our size, a reliable effect size needed to be at least 0.96, which we did obtain for Passage Reading Comprehension. In addition, in a small-*N* design, that treats individual participants as the unit of replication, the results can be robust and readily replicated; it is easier to make a large number of observations on a smaller population (Smith & Little, 2018). Nevertheless, we feel that the significant results from this small exploratory study warrant further investigation.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## REFERENCES

- Allington, R.L. (2013). What really matters when working with struggling readers. *The Reading Teacher*, 66(7), 520-530.
- Allor, J.H., Gifford, D.B., Al Otaiba, S., Miller, S.J. & Cheatham, J.P. (2013). Teaching students with intellectual disability to integrate reading skills: Effects of text and text based lessons. *Remedial and Special Education*, 34, 346-356.
- Allor, J.H., Gifford, D.B., Jones, F.G., Al Otaiba, S., Yovanoff, P., Ortiz, M.B. & Cheatham, J.P. (2018). The Effects of a Text-Centred Literacy Curriculum for Students With Intellectual Disability. *American Journal on Intellectual and Developmental Disabilities*, 123(5), 474-494.
- Amendum, S.J., Conradi, K. & Hiebert, E. (2018). Does Text Complexity Matter in the Elementary Grades? A Research Synthesis of Text Difficulty and Elementary Student's Reading Fluency and Comprehension. *Educational Psychological Review*, 30(1), 121-151.
- Arizpe, E., & Styles, M. (2003). *Children reading pictures: Interpreting visual texts*. London: RoutledgeFalmer.
- Baker, M. (2016). Statisticians issue warning on p values. *Nature*, 531, 151.

## High-percentage decodable text versus low-percentage decodable text

Beverly, B.L., Giles, M. & Buck, K.L. (2009). First-grade reading gains following enrichment: phonics plus decodable texts compared to authentic literature read aloud. *The Reading Teacher*, 53(4), 292-307.

Brown, K.J. (2000). What kind of text – For whom and when? Textual scaffolding for beginning readers. *The Reading Teacher*, 53(4), 292-307.

Castles, A., Rastle, K. & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, 19, 5-51.

Cheatham, P. & Allor, H. (2012). The influence of decodability in early reading text on reading achievement: a review of the evidence. *Reading and Writing*, 25, 2223-2246.

Clark, M.M. (1976). *Young Fluent Readers: What can they teach us?* London: Heineman Educational.

Clark, M.M. (2014). *Synthetic Phonics and Literacy Learning*, Birmingham: Glendale Education.

Coles, G. (2004). Real Books in the Caboose, *Knowledge Quest*, 33(2), 22-25.

Department for Education [DfE] (2010). *Phonics teaching materials: core criteria and the self-assessment process*. London: HMSO.

Department for Education [DfE] (2012). *Criteria for assuring high-quality phonic work*. London: HMSO.

Department for Education [DfE] (2014). *National Curriculum for English Key Stages 1 and 2*. London: HMSO.

Department for Education [DfE] (2015). *Reading: the next steps*. London: HMSO.

Dunn, L.M., Dunn, D.M., Sewell, J., Styles, B., Brzyska, B., Shamsan, Y., & Burge, B. (2009). *The British Picture Vocabulary Scale* (3<sup>rd</sup> ed.), London: GL Assessment.

Fisher, D. & Frey, N. (2014). Scaffolded reading instruction of content-area texts. *The Reading Teacher*, 67(5), 347-351.

Frith, U., (1985). Beneath the surface of developmental dyslexia. *Surface Dyslexia*, 32, 301-330.

Goodman, K.S. (1976). Reading: A Psycholinguistic Guessing Game. *Journal of the Reading Specialist*, 6, 126-135.

Goswami, U., Barnes, L., Mead, N., Power, A.J., & Leong, V. (2016). Prosodic Similarity Effects in Short-Term Memory in Developmental Dyslexia. *Dyslexia*. DOI: 10.1002/dys.1535.

High-percentage decodable text versus low-percentage decodable text

Greene, J.F. and Wood, J.F. (2000). *Language Readers: Language Readers Level 1, Book A, Units 1-6*. Longmont, CO: Sopris West. pp 4-7.

Hall, C.S. (2016). Inference Instruction for Struggling Readers: a Synthesis of Intervention Research, *Educational Psychological Review*, 28(1), 1-22.

Hoffman, J.V., Sailors, M., & Patterson, E.U. (2002). Decodable Texts for Beginning Reading Instruction: The Year 2000 Basals, *Journal of Literacy Research*. 34(3), 269-298

House of Commons Education and Skills Committee. (2005). *Teaching Children to Read, Eighth Report of Session 2004-5*. House of Commons, London: The Stationery Office.

Huey, E.D. (1908). *The Psychology and Pedagogy of Reading*, New York: The Macmillan Company.

Hutchison, D., & Styles, B. (2010). *A guide to running randomised controlled trials for educational researchers*. Slough: NFER.

Jenkins, J.R., Peyton, J.A., Sanders, E.A., & Vadasy, P.F. (2004). Effects of Reading Decodable Texts in Supplemental First-Grade Tutoring. *Scientific Studies of Reading*, 8(1), 53-85.

Johnston, R.S., & Morrison, M. (2007). Towards a Resolution of Inconsistencies in the Phonological Deficit Theory of Reading Disorders: Phonological Reading Difficulties Are More Severe in High-IQ Poor Readers. *Journal of Learning Disabilities*, 40(1), 66-79.

Kelly, L.B., & Moses, L. (2018). Children's Literature That Sparks Inferential Discussions. *The Reading Teacher*, 72(1), 21-29.

Kachorsky, D., Moses, L., Serafini, F., & Hoelting, M. (2017). Meaning Making With Picturebooks: Young Children's Use of Semiotic Resources. *Literacy Research and Instruction* 56(3), 231-249.

Lacey, P., Layton, L., Miller, C., Goldbart, J., & Lawson, H. (2007). What is literacy for students with severe learning difficulties? Exploring conventional and inclusive literacy. *Journal of Research in Special Educational Needs*, 7(3), 149-160.

Leong, V., Hamalainen, J., Soltesz, F., Goswami, U. (2011). Rise time perception and detection of syllable stress in adults with developmental dyslexia. *Journal of Memory and Language*, 64, 59-73.

## High-percentage decodable text versus low-percentage decodable text

- Marshall, B. (2011). English in the National Curriculum: a simple redraft or a major rewrite? *The Curriculum Journal*, 22(2), 187-199.
- Maxwell, S.E., & Delaney, H.D. (2008). *Designing experiments and analysing data* (2<sup>nd</sup> ed.). New York, NY: Routledge Academic.
- McNally, J., & Murray, W. (1962). *Key Words to Literacy*. London: The Schoolmaster Publishing Co.
- Menter, I., Valeeva, R. & Kalimullin, A. (2017). A tale of two countries – forty years on: politics and teacher education in Russia and England, *European Journal of Teacher Education*, 40(5), 616-629.
- Mesmer, H.A.E. (2009). Decodable text: A review of what we know. *Reading Research and Instruction*, 40(2), 121-141.
- Michaud, M., Dion, E., Barrette, A., Dupéré, V. & Toste, J. (2017). Does Knowing What a Word Means Influence How Easily its Decoding is Learned? *Reading & Writing Quarterly*, 33(1), 82-96.
- Morris, D. (2015). Preventing Early Reading Failure. *The Reading Teacher*, 68(7), 502-509.
- Morris, E. (2012). Managing change- The relationship between education and politics. *Better Evidence-based Education*, 4(2), 4-5.
- Mosely, D. (2004). The diagnostic assessment of word recognition and phonic skills in five-year-olds. *Journal of Research in Reading*, 27(2), 132-140.
- Moses, L. & Kelly, L.B. (2018). ‘We’re a little loud. That’s because we like to read!’: Developing positive views of reading in a diverse, urban first grade. *Journal of Early Childhood Literacy*, 18(3), 307-337.
- Munton, G. (2006). *Tab the cat*. Oxford University Press: Oxford
- Murray, B.A., McIlwain, M.J., Wang, C., Murray, G. & Finley, S. (2019). How do beginners learn to read irregular words as sight words? *Journal of Research in Reading*, 41(1), 123-136.
- National Institute of Child Health and Human Development [NICHD]. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

## High-percentage decodable text versus low-percentage decodable text

- National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). *Common Core State Standards for English language arts & literacy*. Washington D.C: National Governors Association Center for Best Practices, Council of Chief State School Officers.
- Nahachewsky, J. (2013). Understanding the importance of ethos in composing the “everyday” new literacies classroom. *Language and Literacy*, 15(1), 74-92.
- Nuzzo, R. (2014). Scientific method: Statistical errors. P values, the ‘gold standard’ of statistical validity, are not as reliable as many scientists assume. *Nature*, 506, 150-152.
- Office for Standards in Education, Children’s Services and Skills (Ofsted). (2019). <https://www.gov.uk/government/publications/ofsted-inspection-handbooks-drafts-for-consultation> Accessed 9 February 2019.
- Pearson, P.D. (2004). The Reading Wars. *Educational Policy*, 18(1), 216-252.
- Pring, R. (2015). *Philosophy of Educational Research* (Third Edition). London: Bloomsbury.
- Protopapas, A., Mouzaki, A., Sideridis, G.D., Kotsolakou, A., & Simos, P.G. (2013). The Role of Vocabulary in the Context of the Simple View of Reading. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 29(2), 168-202.
- Price-Mohr, R. M (2016). *Comparing a controlled levelled text with a language rich text in a beginner reading scheme*. University of York. <http://etheses.whiterose.ac.uk/12904/>
- Ricketts, J., Bishop, D.V.M., Pimperton, H., & Nation, K. (2011). The Role of Self-Teaching in Learning Orthographic and Semantic Aspects of New Words. *Scientific Studies of Reading*, 15(1), 47-70.
- Slavin, R.E, Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective Reading Programs for the Elementary Grades: A Best-Evidence Synthesis. *Review of Educational Research*, 79(4), 1391-1466.
- Smith, F. (1973). *Understanding reading: A psycholinguistic analysis of reading and learning to read* (2<sup>nd</sup> ed.). New York: Holt, Rinehart & Winston.
- Smith, P.L., & Little, D.R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic Bulletin and Review*, 25(6), 2083-2101.

## High-percentage decodable text versus low-percentage decodable text

- Snowling, M., Stothard, S.E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., Nation, K. & Hulme, C. (2009). *York Assessment of Reading for Comprehension*. London: GL Assessment.
- Solar, J., & Openshaw, R. (2007). 'To be or not to be?': The politics of teaching phonics in England and New Zealand. *Journal of Early Childhood Literacy*, 7(3), 333-352.
- Torgerson, C.J., Brookes, G., & Hall, J. (2006). *A Systematic Review of the Research Literature in the use of phonics in the teaching of reading and spelling*. London: Department for Education and Skills (DfES).
- Torgerson, D.J., & Torgerson, C.J. (2008). *Designing Randomised Trials*. New York: Palgrave Macmillan.
- Nation, K., & Hulme, C. (2009). *York Assessment of Reading for Comprehension*. London: GL Assessment.
- Solity, J., & Vousden, J. (2009). Real books vs reading schemes: a new perspective from instructional psychology. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 29(4), 469-511.
- Stahl, S.A., & Heubach, K.M. (2005). Fluency-oriented reading instruction. *Journal of Literacy Research*, 37(1), 25-60.
- Stephen, C. (2010). Pedagogy: the silent partner in early years learning. *Early Years*, 30(1), 1-14.
- Stokes, S.F., Kern, S., & DosSantos, C. (2012). Extended Statistical Learning as an account for slow vocabulary growth. *Journal of Child Language*, 39(1), 105-129.
- Tesar, M. (2019). Global politics and local impacts on educational policy. *Policy Futures in Education*, 17(3), 302-305.
- Victoria Liberal (2018). <https://vic.liberal.org.au/News/2018-10-10/liberal-nationals-to-fund-decodable-readers-for-victorian-schools> Accessed 9 February 2019.
- Vygotsky, L. (1978). *The Role of Play in Development, Mind in Society*. Cambridge, MA: Harvard University Press.
- Walsh, M. (2003). 'Reading' pictures: what do they reveal? Young children's reading of visual texts. *Literacy*, 37(3), 123-130.

## High-percentage decodable text versus low-percentage decodable text

Wenz-Gross, M., & Upshur, C. (2012). Implementing a Primary Prevention Social Skills Intervention in Urban Preschools: Factors Associated with Quality and Fidelity. *Early Education and Development, 23(4)*, 427-450.

Wilson, J. & Comar, S. (2008). Re-evaluating the Significance of Phonemic Awareness and Phonics in Literacy Teaching: The Shared Role of School Counsellors and Teachers. *Australian Journal of Guidance and Counselling, 18(2)*, 89-105.

Ziegler, J.C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin, 131(1)*, 3-29.