Learning and Instruction 65 (2020) 101248

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Using a picture-embedded method to support acquisition of sight words

C. Benjamin Strauber*, Piya Sorcar, Claire Howlett, Shelley Goldman

Stanford University, USA

ARTICLE INFO

Keywords:
Literacy
Sight words
Early readers
Modified orthography
Multimodal

ABSTRACT

This study investigated whether an intervention using words embedded with pictures can be more effective in sight word instruction than one using words alone. Participants included sixty-nine children in junior kindergarten (ages 4–5) enrolled in school in Ontario, Canada. Children were split randomly into treatment and control groups; the treatment group was taught four words using picture-embedded words, and the control group was taught using text alone. Both groups also received phonics instruction to support sight word acquisition. Children in the picture-embedded word condition performed significantly higher than those in the word-alone condition on an immediate post-training test and later retention tests. This outcome, which contrasts with previous studies using picture-embedded words, may result from this method's use of a relevant linking phrase and action that help build an association between picture and word, as well as its incorporation of phonics instruction, with future work needed to test this hypothesis.

1. Introduction

All written material consists of a high proportion of common words referred to as high-frequency words. The 100 most common words make up approximately half of all material we read, with the 25 most common words comprising about a third (Fry & Kress, 2012). Given the ubiquity of high-frequency words, it is advantageous for beginning readers to learn to recognize them: a child who can instantly recognize the 100 most common words has the base needed to read the majority of a typical text. In many classrooms, children are taught to immediately recognize these high-frequency words by sight. Recognizing words by sight is how experienced readers process written words with which they are familiar (Ehri, 2005), and explicitly teaching high-frequency sight words to children can help jumpstart this strategy of reading. While phonics instruction is often used both to bolster acquisition of high-frequency sight words and to allow for decoding of other words, explicit memorization is especially advantageous for words that are not phonetically transparent, such as “one,” “are,” or “look.” With a core of basic sight words committed to memory, early readers can then use phonetic strategies to decode unknown words in a text and ultimately learn to recognize them by sight as well. With a set of sight words in memory, reading becomes more fluent because the learner does not have to decode every word (Ehri, 2014). Given the value of teaching high-frequency words to young children, teachers are encouraged to promote mastery of these words with great efficiency in the early years of schooling (Fry & Kress, 2012). Because of the value and wide use of sight word training in reading acquisition, a significant body of research has focused on teaching both typically developing children and children with learning disabilities to recognize sight words (Blackwell & Laman, 2013; Browder & Lalli, 1991; Ehri, 2005; Blischak & McDaniel, 1995; Van der Bijl, Alant, & Lloyd, 2006; and many others).

1.1. Using pictures to teach sight words

One strategy for teaching sight words that has been the focus of a number of studies is to pair vocabulary words with corresponding pictures to help children form an association between the written word and its meaning (Elliott & Zhang, 1998). Pictures can increase the relevance of a sight word and children's motivation to learn it, and they can also provide an extra cue that assists learners in determining the word's meaning. Additionally, consistent with dual coding theory (Paivio, 1991), having both visual and verbal input when learning a sight word can help learners form a stronger memory trace, deepening their association between the sight word and its meaning.

Despite the apparent advantages of using pictures in teaching sight words, most studies on this strategy's effectiveness have in fact found that placing pictures next to sight words provides no learning benefit, and can often be detrimental to learning, in comparison to teaching written words alone (Singer, Samuels, & Spiroff, 1973; Harzem, Lee, & Miles, 1976; Singer, 1980; Singh & Solman, 1990; Hill, 1995; Solman & Wu, 1995; Elliott & Zhang, 1998; Pufpaff, Blischak, & Lloyd, 2000;

* Corresponding author. 505 Lasuen Mall, Stanford, CA, 94305, USA.
E-mail address: strauber@stanford.edu (C.B. Strauber).

https://doi.org/10.1016/j.learninstruc.2019.101248
Received 16 November 2018; Received in revised form 21 August 2019; Accepted 22 August 2019
Available online 24 September 2019

Meadan, Stoner, & Parette, 2008; Dittlinger & Lerman, 2011).[1] One notable exception has been dismissed by other papers as having methodological flaws (Arlin, Scott, & Webster, 1978).[2] Several reasons have been proposed for the common observation that side-by-side pictures do not assist in sight word acquisition. First, in what has been referred to as the focal attention hypothesis, a number of researchers have suggested that pictures distract learners from the word to be learned (Didden, Prinsen, & Sigafoos, 2000; Samuels, 1967; Singer et al., 1973; Solman, Singh, & Kehoe, 1992). In this view, putting pictures next to sight words can result in an overshadowing effect, whereby the picture distracts the child's attention from the word, and/or a blocking effect, whereby the child relies on the picture to identify the spoken word it is associated with, ignoring the written word (Dittlinger & Lerman, 2011). Second, if pictures and words are placed next to each other or even integrated, there is nothing to guarantee that children will learn an association between the two. Researchers have suggested that the relationship between the picture and word should be emphasized for the child (Fossett & Mirenda, 2006).

1.2. Modified orthography

To avoid the aforementioned issue of pictures distracting from the words they represent, attempts have been made to integrate words and pictures in what has been called “modified orthography” or “embedded words” (Blischak & McDaniel, 1995; Didden, de Graaff, Nelemans, Vooren, & Lancioni, 2006; Pufpaff et al., 2000; Tabe & Jackson, 1989; Van der Bijl et al., 2006). In this method, a word is superimposed onto a picture that represents it, or a picture is drawn into one or more of the letters of the word. In theory, this causes the learner's attention to be focused both on the word and the picture together, allowing for the picture to assist in the learning of the word rather than distracting from it. While research has shown that integrating words with pictures results in greater learning than placing words next to pictures (Blischak & McDaniel, 1995), nearly all studies using embedded pictures have found they provide no learning advantage compared to using written words alone.

One reason for this may be that children must be taught to associate the word with the meaning represented by the picture it contains, a step not explicitly taken in many of the studies conducted on picture-embedded words. Past work in related areas has used a variety of methods to help learners form an association between the word and the concept represented by the picture. One approach has involved explicitly telling children that a picture presented next to a word is a representation of that word; Lang and Solman (1979) found that this approach led to increased learning relative to when the association was not made explicit. A second approach is picture-to-text matching, in which children are asked to actively match words with the pictures that represent them; Fossett and Mirenda (2006) found this method was more effective than merely teaching with pictures next to their corresponding words. A third approach is picture fading, in which a picture is presented in or around a word and faded out over a series of images. Studies by Smeets, Lancioni, and Hoogeveen (1984) and Tabe and Jackson (1989) used picture fading to teach sight words to children with some success, though the latter study found no difference in outcome between picture fading and pictures next to words. Using individualized sessions with seven participants, Richardson et al. (2017) evaluated three separate training methods that employed either picture fading, picture matching, or text alone with echoic prompts. They observed that some individual participants acquired words more quickly using the picture-matching method than the text-alone method, though neither picture matching nor picture fading resulted in greater overall mastery than text alone at final posttest. A fourth approach, word morphing, employs a series of images or animations in which a picture morphs into a word, providing a gradual transition intended to help the child form an association between the two. Miller and Miller (1971) found that animations morphing from a picture to a word resulted in greater learning than when the picture was next to the word. Sheehy (2005) later found a similar approach to result in greater learning than a word-alone teaching method.

Hoping to build this association using picture-embedded words, Van der Bijl et al. (2006) explored the impact of multiple presentation modes. They found that using cards with just picture-embedded words during the acquisition phase resulted ultimately in lower identification scores than using cards with the text-only word. By contrast, using cards with the picture-embedded word on one side and text-only word on the other side, and emphasizing to children that the two words were the same, resulted in higher identification scores than presenting the modified word alone with a reported effect size of 0.56 SD, though this difference was not statistically significant at $p < 0.05$. Beyond emphasizing the words were the same, this method did not seek to build an explicit semantic association between the picture and the word.

1.3. Orthographic mapping

Another association relevant to sight word acquisition is orthographic mapping, the ability of readers to associate sounds with symbols, but explicit focus on orthographic mapping has been conspicuously absent from past studies that have used pictures to teach sight words to children. Previous work suggests that early readers likely rely on such mapping in the process of learning sight words (Ehri, 2014). In Ehri's formulation of the phases of reading, children initially use orthographic mapping of individual letters and larger multi-letter units to form and retain associations between the appearance of a word and its sound. Some studies have successfully used pictures embedded in or integrated with individual letters to teach sound-symbol mappings for these individual letters (Ehri, Deffner, & Wilce, 1984; Shmidman & Ehri, 2010), but no study to our knowledge has used words with embedded pictures in conjunction with phonics instruction to teach whole sight words to early readers.

1.4. The present study

As discussed, missing from modified orthography methods reported until now has been an explicit association between the word and the picture that represents it, as well as integration with phonics instruction. To address this gap, we present here a first evaluation of a picture-embedded word teaching method that incorporates these two components. Given these supports connecting the word with the picture, we hypothesize that children in this study will learn sight words when using embedded pictures as well as or better than when using text alone. We also hypothesize that this method will be particularly effective for words whose initial sounds are more difficult for children as compared to those that are easier because, in the case of the former, reading words through phonics and storing them using orthographic mapping may be particularly challenging for beginning readers.

2. Method

2.1. Participants

Sixty-nine children in junior kindergarten (ages 4–5; average age 4.82 years; 41 females) were selected to participate in the acquisition study.[3] Informed consent was obtained based on the requirements of

[1] Among studies that reported data allowing for a calculation of effects sizes, effect sizes for methods using pictures next to words compared to methods using words alone ranged from −0.66 to 0 SD.

[2] This study reports a remarkable effect size of close to 4 standard deviations.

[3] With alpha = 0.05, beta = 0.8, and R-squared = 0.4, we estimated that we

Thames Valley School Board in a process approved by the Institutional Review Board of Stanford University. Participants complied with the following selection criteria: English as home language, normal visual and hearing ability, normal speech and language development, no obvious learning difficulties, and no obvious emotional/behavioral problems. Information was obtained through interviews with the relevant educator and school records. Students had been taught some sight words using word-alone methods but not any of the sight words used in this study. Furthermore, students had received instruction in letter recognition and letter-sound association, though teachers of these classes generally observed that students could not associate all letters with a sound. In the acquisition study, the 69 children were randomly placed into a treatment group of 35 students, who learned words using cards with picture-embedded words, and a control group of 34 students, who learned words using cards with text alone. In addition to the 69 children who participated in the acquisition study, a separate 40 students matching the selection criteria participated in the stimulus word selection procedure described below, and a further 60 students matching the selection criteria participated in a later assessment intended to confirm these words are generally not known to students in this population.

2.2. Materials and equipment

2.2.1. Stimulus word selection

A list of 24 high-frequency words were selected from Fry's Instant Word List (Fry & Kress, 2012). Forty children across 9 classrooms who complied with the selection criteria described above participated in the word selection procedure. In the word selection procedure, the text-only form of each of the 24 words was shown individually on flashcards to each of the 40 children. The children were asked to read each of them with the following instructions: “Read each word to me if you happen to know it. It is okay if you don't know the word; just say, ‘I don't know'." Each child's response for each word was recorded. From this list of 24 words, the four words that were most commonly missed by students were selected for inclusion in the full study. The selection of four words for training and testing is consistent with previous seminal work in this area (Singer et al., 1973). These four words (little, his, all, out) were unreadable to all but three of the 40 students participating in the word selection procedure, two of whom recognized one word and one of whom recognized two words. Because knowledge of these words was so low among students in our target demographic, and because we randomized the treatment and control groups, we chose not to use a pre-test in the main study.

Amid doubts about students' initial word knowledge, to confirm the representativeness of the knowledge levels identified in the selection procedure, after running the main study we tested a separate sample of 60 junior kindergarten students meeting the selection criteria on their knowledge of these four words. Consistent with our earlier finding of low knowledge levels for these four words among junior kindergarten students, we found that only two of these 60 children knew one of the words, and no child knew more than one of the words.

2.2.2. Stimulus material

The picture-embedded words used in the experiment were created by an organization called Eyewords (http://www.eyewords.com). The picture-embedded words were printed in lowercase letters using size 175 pt. Century Gothic font and modified by embedding colored graphics into the text. Pictures were chosen both to fit the shapes of the letters and to provide material for an association to be built between the written word and its meaning. The shapes of the letters were used as much as possible in creating the picture; for example, in the word his, the s was stylized into a snake, and the i was stylized into a boy holding the snake (Fig. 1). The image of the boy holding a snake indicated possession, allowing for the use of the associated phrase “his snake,” which helped build a connection between the word's appearance and meaning. The modified words were mounted on 18 cm × 11 cm flashcards. The text-only words were printed in lowercase letters using size 175 pt. Century Gothic font and mounted on 18 cm × 11 cm flashcards.

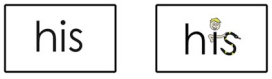

Fig. 1. Example of word-alone flashcard (left) and picture-embedded word flashcard (right) for the word “his.”

2.3. Experimental procedure

The experimental procedure consisted of randomization followed by three phases: acquisition period, immediate post-test, and retention tests. Participants were first randomly assigned to two groups, one receiving instruction with picture-embedded words and the other receiving instruction with text-only words. In the acquisition phase, each student attended two 15-min learning sessions individually, with one session each on two consecutive days. In each of the experimental conditions, the children were exposed to four high frequency words (little, his, all, out) in a random order for each child on each day. Following the learning session on the second day, an immediate post-training test was administered and the correct number of responses recorded. Retention tests were administered on days 4, 10, 15 and 18, and the correct number of responses was recorded to measure retention of the words taught. The participation schedule is represented in Table 1.

Table 1
Schedule for participation in study.

Day 1	Day 2	Day 4	Day 10	Day 15	Day 18
15-min acquisition session using Treatment Condition or Control Condition	Repeat Day 1 procedure followed by immediate post-test	Retention Test	Retention Test	Retention Test	Retention Test

2.3.1. Flashcard introduction

For both conditions, the child was presented with one word card at a time. In the control condition, upon presentation of the word card, the instructor said the word and asked the child to repeat the word (i.e., “This word says ‘his’. Can you say ‘his’?“). In the picture-embedded word condition, following word presentation, the instructor also introduced a phrase and action relevant to the embedded picture (i.e., “This word says ‘his’, his snake.” [instructor made a grabbing action at shoulder height] Can you say ‘his, his snake’? [time given for student to repeat the word and phrase.] Look at the picture in the word ‘his’. The letter ‘i’ looks like a little boy grabbing a snake. The letter ‘s’ is the

(*footnote continued*)
would need 62 participants (in an experiment with two treatment arms) for a minimum detectable effect size (MDES) of 0.50 SDs.

Table 2
Phrase, action, and description for each word in the embedded words condition.

Word	Phrase	Action	Description
all	all of the pizza	hands make a circle to represent an entire pizza	The letter 'a' looks like a pizza with **all** of its slices.
his	his snake	grabbing action at shoulder height	The letter 'i' looks like a little boy grabbing a snake. The letter 's' is the snake. It is **his** snake.
little	little boy	thumb and index finger placed an inch apart	The letter 'i' looks like a **little** boy.
out	out of the hole	thumb makes a motion indicating moving out	The letter 'u' looks like a hole with an arrow showing how to climb **out** of it.

Table 3
Phonics description for each word.

Word	Phonics description
all	The first letter in the word 'all' is 'a'. In this word it actually makes the sound makes the sound 'o'. The next two letters in 'all' are 'll'. They go together to make the sound 'l'. Let's stretch these sounds out like bubblegum 'o-l'. Stretch them out with me: o-l. And now we can snap them back together to say 'all'.
his	The first letter in 'his' is 'h'. It makes the sound 'h'. The next letter in 'his' is 'I'. In the word 'his' the 'I' makes the sound 'i'. The last letter in 'his' is 's'. In this word, the 's' sounds like 'z'. We can stretch these sounds out like bubblegum 'h-i-s'. Stretch them out with me: 'h-i-s'. And now we can snap them back together to say 'his'.
little	The first letter in 'little' is 'l'. It makes the sound 'l'. The next letter in 'little' is 'l'. In the word 'little' the 'I' makes the sound 'i'. The next two letters in 'little' are 'tt'. They go together and make the sound 't'. The next letter in 'little' is 'l'. It makes the sound 'l'. The last letter is an 'e' and is silent in this word. We can stretch these sounds out like bubblegum 'l-i-tt-l'. Stretch them out with me: 'l-i-tt-l'. And now we can snap them back together to say 'little'.
out	The first two letters in 'out' are 'ou'. They go together to make the sound 'ow'. The next letter in 'out' is 't'. It makes the sound 't'. We can stretch these sounds out like bubblegum 'ou-t'. Stretch them out with me: 'ou-t. And now we can snap them back together to say 'out'.

snake. It is **his** snake.") For each word, the phrase, action, and description used in the embedded words condition are provided in Table 2.

In both conditions, after initial introduction to the word, the instructor introduced the word using phonics, breaking down the word into sound segments and saying the sound each letter or letter combination makes. If letters didn't follow a rule, the student was told that the rule is broken in this case. For example, in the case of the word "his," phonics instruction was, "The first letter in 'his' is 'H'. It makes the sound 'h'. The next letter in 'his' is 'I'. In the word 'his' the 'I' makes the sound 'i'. The last letter in 'his' is 'S'. In this word, the 'S' sounds like 'z'. We can stretch these sounds out like bubblegum 'h-i-s'. Stretch them out with me: 'h-i-s'. And now, we can snap them back together to say 'his'. Can you look at this word and say it again with me? 'His'." Time was then given for student to repeat the word. The phonics description for each word is provided in Table 3.

2.3.2. Games

Students in both conditions played a series of three games involving the words: Bunny Hop, Wordmatch, and Word Detective.

2.3.2.1. Bunny Hop. The four word cards were spread out in a column on the floor. The instructor called out one of the four words and had the child hop to the card. If the child hopped to the incorrect word card, the instructor said, "Nice try, but that word says [chosen word]." Bunny Hop was played until all four words had been called out. The game was played through twice, randomizing the order of the words in the row and the order of the words called out. In the control condition, both rounds of the game were played using plain-text word cards. In the picture-embedded word condition, the first round of the game was played using picture-embedded word cards, and the instructor said the auditory phrase after each word being called out and demonstrated the action associated with the word. The second round was played using plain-text word cards, and the instructor did not provide the auditory phrase or action after calling each word. The layout for the Bunny Hop game is shown in Fig. 2.

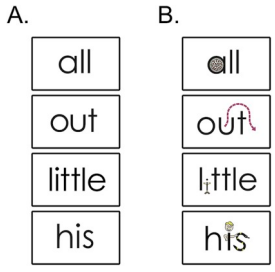

Fig. 2. Flashcard configuration for Bunny Hop game. Configuration A (left) was used in this game for rounds one and two of the control condition as well as round two of the picture-embedded word condition, and Configuration B (right) for round one of the picture-embedded word condition. Note: Though the same word order is shown in A and B to depict the contrast between the conditions, words were presented in a different randomized order in each round.

2.3.2.2. Word Match. A word card was displayed and beneath that word card was a 1 × 4 array of all four word cards. The top word card was a match to one of the four word cards below. The student was asked to place a star-shaped marker on the top of the matching word card in the 1 × 4 array. If the student marked the wrong word card, the instructor indicated the correct answer by saying "Nice try, but the matching word is [correct word]," while placing the star next to the correct word on its card. The instructor called out each of the four words per game. Wordmatch was played twice for a total of eight matching opportunities per student. With each round, the top word card was changed until all target words had been displayed at the top. The order of top word card and word cards below was randomized. In the control condition of this game, plain-text word cards were used above and below for all four rounds and in both games. In the picture-embedded word condition, in the first game, the top word card was presented in picture-embedded word form and the auditory phrase and action for the word was provided by the instructor when it was called. The below 1 × 4 array was presented in plain text so that the student had to match a plain-text word card to a picture-embedded word card.

The second time the game was played in the picture-embedded word condition, all word cards were presented in plain text and no auditory phrase or action was given after calling for a word. The layout for the Word Match game is shown in Fig. 3.

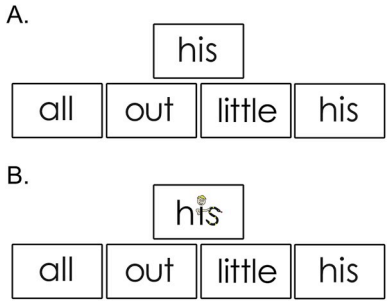

Fig. 3. Flashcard configuration for Word Match game. Configuration A (top) was used in this game for rounds one and two of the control condition as well as round two of the picture-embedded word condition, and Configuration B (bottom) was used for round one of the picture-embedded word condition. Note: Though the same bottom row word order and top word are shown in A and B to depict the contrast between conditions, the words in the bottom row were presented in a different randomized order in each round, and the top word was changed between individual trials.

2.3.2.3. Word Detective. The picture-embedded word cards were spread out in a 2 × 2 array on the floor. The instructor called out one of the four words and asked the learner to search for and identify the word being requested. Once identified, the student covered the word with a transparent purple card. In the case of a correct response, the child was provided positive feedback, the purple card was left on the word card, and the next word was called out. In the case of incorrect responses, the child was asked to try again until choosing the correct card. The instructor called out each of the four words in the game. The game was played twice so that each student made a total of eight selections. For the control condition, all words were displayed in plain-text word cards in both games. For the picture-embedded word condition, the first time the game was played, picture-embedded word cards were displayed, and the auditory phrase and action were provided. The second time the game was played, plain-text word cards were displayed, and neither an auditory phrase nor an action was provided. The layout for the Word Detective game is shown in Fig. 4.

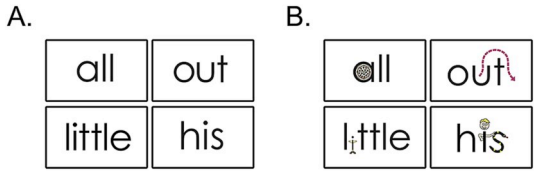

Fig. 4. Flashcard configuration for Word Detective game. Configuration A (left) was used in this game for rounds one and two of the control condition as well as round two of the picture-embedded word condition, and Configuration B (right) was used for round one of the picture-embedded word condition. Note: Though the same word order is shown in A and B to depict the contrast between the conditions, words were presented in a different randomized order in each round.

2.3.3. Immediate post-test

On day 2 of the study, immediately following the acquisition period, a post-test was conducted by the instructor in both experimental conditions. Participants in both conditions were shown the plain-text stimulus words on flashcards and asked to read them. The instructions were as follows: "Now, I am going to ask you to read the correct word after I show it to you. Look carefully at the word and try your best." Word presentation was randomized but the same for each student. Student responses for each word were recorded on a post-test recording sheet. Five students (four from the word-alone group, and one from the picture-embedded group) were absent on the day of the post-test, meaning they also missed the second half of the training; their data was thus excluded from all subsequent analysis, leaving 30 students in the word-alone group and 34 in the embedded pictures group for the remainder of the analysis.

2.3.4. Retention tests

Retention tests were carried out on days 4, 10, 15 and 18. Students from both conditions were withdrawn from their classrooms individually and asked to read each of the four plain text word cards as they were presented. The instructions were as follows: "A while ago, you played some word games and learned four words. Now, I am going to ask you to read the correct word after I show it to you. Look carefully at the word and try your best." The order of word presentation was randomized for each test day but was the same for each student on a given day. Student responses for each word were recorded on retention test recording sheets. There were a total of seven absences across all retention time points; as they had already undergone training, students were not excluded from the study if they missed a retention test day.

2.4. Baseline measures

To check that groups were matched on a variety of baseline measures, a survey asking about demographic features was distributed to parents of students participating in the study. The survey asked questions about number of books in the home, number of minutes the child is read to at home daily, the child's number of siblings, the nature of the child's commute to school, mother's education level, father's education level, and child race. Phonological awareness assessment scores from the semester prior to the study were also collected. The assessment included common measures of phonological awareness, including rhyming, initial and final sound identification, sound segmentation, sound deletion, sound substitution, and sound blending. Not all participants' parents agreed for information on baseline measures to be released; it was collected for roughly two-thirds of participants included in the analysis (19 out of 30 students in the word-alone condition, and 24 out of 34 students in the embedded words condition).[4]

3. Results

3.1. Baseline measures

Because groups were randomized, an ordinary least squares regression was conducted to check for differences in baseline measures between the treatment and control groups. While responses collected in the demographic survey were categorical, we converted them to binary variables for the sake of analysis. The binaries used were: whether the child was male, whether the child had more than 50 books in the home at the start of kindergarten, whether the child was read to for more than 30 min a day, whether the child had more than one sibling, whether the child was driven to school, whether the child's parents were married, whether the child's mother had at least a bachelor's degree, whether the child's father had at least a bachelor's degree, and whether the child's race was white. Percent correct on the phonological awareness test and the child's age were also included as baseline measures. Results of this regression, shown in Table 4, indicated that there were no significant

[4] Later analysis of immediate post-test scores revealed no significant difference on this outcome measure in either condition between children for whom baseline data was collected and for whom it was not collected.

Table 4
Descriptive statistics of the sample and balance test.

Variable	Mean in TreatmentGroup	Mean in Control Group	Difference (Treatment - Control)	Standard Error of Difference	t	p
Male (y/n)	0.38	0.43	−0.05	(0.09)	−0.41	0.68
More than 50 books in home (y/n)	0.84	0.95	−0.11	(0.09)	−1.15	0.26
More than 30 min reading day (y/n)	0.75	0.74	0.01	(0.14)	0.10	0.92
More than one sibling (y/n)	0.29	0.16	0.13	(0.13)	1.02	0.31
Driven to school (y/n)	0.38	0.33	0.05	(0.12)	0.40	0.69
Parents married (y/n)	0.79	0.94	−0.15	(0.08)	−1.40	0.17
Mother bachelor's degree (y/n)	0.34	0.32	0.02	(0.15)	0.12	0.91
Father bachelor's degree (y/n)	0.21	0.16	0.05	(0.12)	0.41	0.68
Race = white (y/n)	0.87	0.94	−0.07	(0.09)	−0.75	0.46
Phonological Awareness ratio correct	0.61	0.59	0.02	(0.10)	0.17	0.87
Age at onset of study (years)	4.85	4.81	0.04	(0.07)	0.59	0.56

Table 5
Number of participants, means, standard deviations, *t*-test value, degrees of freedom, significance of difference between groups, and Cohen's d at each time point.

	n	Mean	SD	t	df	p	d
Immediate post-test							
Embedded words	34	3.38	0.99	2.18	62	0.033	0.55
Text alone	30	2.80	1.16				
Retention test day 4							
Embedded words	34	3.29	1.19	3.12	62	0.003	0.79
Text alone	30	2.27	1.44				
Retention test day 10							
Embedded words	33	3.49	0.87	4.19	61	< 0.001	1.1
Text alone	30	2.30	1.34				
Retention test day 15							
Embedded words	33	3.48	0.90	2.89	61	0.005	0.74
Text alone	30	2.63	1.40				
Retention test day 18							
Embedded words	31	3.61	0.84	3.34	57	0.001	0.89
Text alone	28	2.68	1.28				

differences between groups for any of the baseline measures collected.

3.2. Immediate post-training test and retention tests

Two-tailed independent samples t-tests were conducted comparing treatment and control groups at each of the five time points. Means, standard deviations, and significance of difference between groups at each time point are shown in Table 5, and averaged raw scores from the immediate post-training and retention tests are shown in Fig. 5. Scores at the immediate post-training test were greater in the picture-embedded word condition than the word-alone condition by an average of 0.58 words, a difference significant at $p < 0.05$; scores on the retention tests at four, ten, fifteen, and eighteen days were respectively greater in the picture-embedded word condition than the word-alone condition by 1.02, 1.19, 0.85, and 0.93 words, with all differences significant at $p < 0.01$. Given that a total of four words were tested, this represents a difference in performance between conditions of 15 percentage points at the immediate post-training test and a difference of 26, 30, 21, and 23 percentage points for the retention tests at four, ten, fifteen, and eighteen days, respectively. The difference between conditions was larger in the retention tests than in the immediate post-training tests, largely because the scores of the control condition decreased while the scores of the treatment condition for the most part slightly increased.

The results of two ordinary least squares regressions that combine observations for each group across all five post-training tests, one with and one without controlling for baseline covariates, can be found in Table 6. Combining observations across all post-training tests, the embedded words intervention increased word acquisition by 0.91 words (out of 4) relative to the control without controlling for baseline covariates. The effect was statistically significant at $p < 0.01$. To account for any potential remaining imbalance in observable covariates and to increase power, we also performed a regression controlling for baseline covariates. The results, reported in Table 6, were substantively the same as those from the original regression, with the adjusted difference between conditions becoming 1.05 words and the effect also significant at $p < 0.01$.

3.3. Easy-initial and difficult-initial words

For the sake of further analysis, words were divided into two groups: words easier to sound out based on their initial letters (little, his), and words more difficult to sound out based on their initial letters (out, all). These words are referred to respectively here as easy-initial and difficult-initial words. The easy-initial words start with consonants that have one primary, common pronunciation in English (relatively consistent grapheme-phoneme mapping), while the difficult-initial words start with vowels that have several pronunciations in English dependent on context. While there was a significant effect of training condition on both word types, ordinary least squares regressions revealed a stronger effect of training condition on difficult-initial word acquisition (difference of 0.62 words) than on easy-initial word acquisition (difference of 0.29 words). Results of these regressions are presented in Table 6. A further regression confirmed the difference in outcomes between difficult-initial and easy-initial words was significantly larger for the picture-embedded words group than the word alone group ($p < 0.01$). This suggests that the picture-embedded words method was especially powerful for helping students learn difficult-initial words.

4. Discussion

We have reported here a method for using embedded pictures to teach sight words that, in contrast with methods used previously, resulted in this initial study in increased word learning compared to plain text alone. One feature most clearly distinguishing this method from those previous, and a potential reason for the increase observed in this study, is the explicit connection made between the picture and the word using a phrase and associated action. This connection could allow beginning readers to transfer meaning from the picture to the word, a step crucial for picture-assisted sight word learning (Van der Bijl et al., 2006) but one not apparent in many previous studies. Both placing the picture inside the word and using a connecting phrase and action may have helped avoid the blocking or overshadowing that has been hypothesized to hinder learning in earlier studies using picture cues (Dittlinger & Lerman, 2011; Singh & Solman, 1990). While this study's design does not allow us to pinpoint the reason for our observation that children learned better using this method as compared to a text-alone method, future work could include conditions intended to do so, some of which we highlight later in this discussion.

Another important feature distinguishing this method from previously used picture-embedded word methods is that the pictures presented here are often not naturally or obviously representative of the

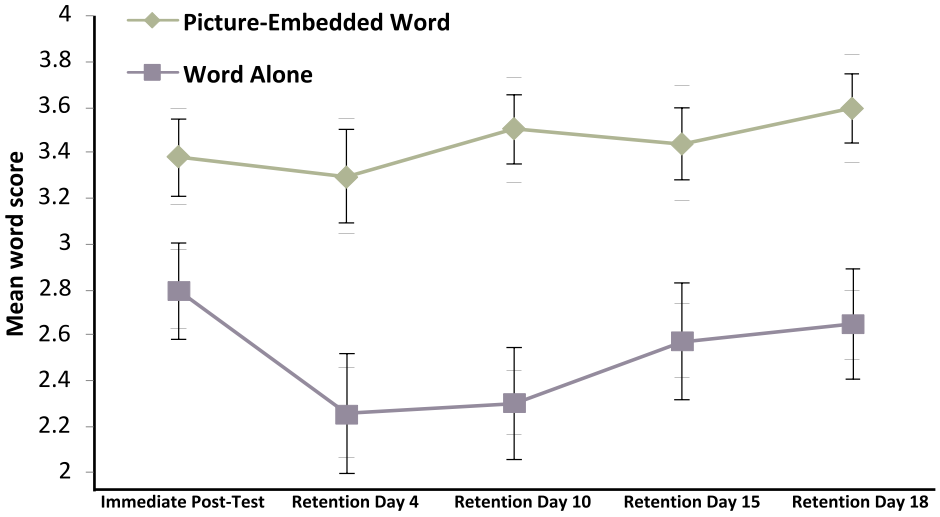

Fig. 5. Word score at immediate post-test and retention tests four, ten, fifteen, and eighteen days after picture-embedded word or word-alone training for all students. Comparisons between conditions at immediate post-training test and final retention test are significant at $p < 0.05$, and comparisons at all other time points are significant at $p < 0.01$.

Table 6
Average treatment effect of intervention on student word acquisition.

	(1)	(2)	(3)	(4)	(5)	(6)
	Total score	Total score	Easy-initial score	Easy-initial score	Difficult-initial score	Difficult-initial score
Treatment	0.91***	1.05***	0.29***	0.29***	0.62***	0.76***
	(0.13)	(0.13)	(0.08)	(0.09)	(0.07)	(0.08)
Controls	NO	YES	NO	YES	NO	YES
n	313	208	313	208	313	208
R-squared	0.149	0.444	0.058	0.348	0.198	0.406

Notes:
1) Heteroskedastic robust standard errors are in parentheses.
2) Treatment indicates how much higher (out of 4 points) average scores were in the embedded words group than in the text-alone group, combining all time points in the study.
3) Controls indicates whether covariates were included in the model. Covariates include all variables shown in Table 4: the child's phonological awareness score, the child's age, whether the child was male, whether the child had more than 50 books in the home at the start of kindergarten, whether the child was read to for more than 30 min a day, whether the child had more than one sibling, whether the child was driven to school, whether the child's parents were married, whether the child's mother was college-educated, whether the child's father was college-educated, and whether the child's race was white.
4) R-squared indicates the degree to which variation in the scores is explained by the independent variable (i.e. intervention condition) if the value in "Controls" is "NO" and by the independent variable plus the covariates if the value in "Controls" is "YES."
5) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

words they teach. For example, the picture for "all" is a whole pizza inside of the "a," and the phrase is "all the pizza." Relative to other methods, this lack of a natural connection increases the need for the bridge phrase. Importantly, it also increases the ease with which words without concrete representations – such as function words, which comprise well more than half of the 100 most frequent words (Fry & Kress, 2012) – can be depicted. This ability to move beyond concrete words overcomes a challenge present in previous formulations of picture-embedded word teaching methods, many of which have focused on concrete nouns with natural pictorial representations (Brown, 2014). As a result, the picture-embedded words method presented here could be effective for teaching all of the high-frequency words needed for success in reading, making it particularly useful in the context of the classroom.

Also in contrast with previous studies in the sight word literature, both conditions in this study drew children's attention to the phonetic pronunciation of each letter in each word during the initial exposure phase. This is consistent with the view that phonics and whole language approaches should be combined in instruction to maximize reading acquisition (Adams, 1994) and that orthographic mapping is an important component of sight word learning (Ehri, 2014), with children using grapheme-phoneme correspondences to form links between word spellings and pronunciations and ultimately retain them in memory. In this case, it is possible that the use of pictures and phonics feed into each other to enhance word learning beyond levels achieved with the use of word alone with phonics. For example, in the early stages of learning to read with phonics, it is often reported that readers will guess at the word after only seeing the first letter (Savage, Stuart, & Hill, 2001). Embedded pictures, which are spread throughout the word, serve to draw children's attention to parts of the word beyond the first letter. It is plausible, then, that embedded pictures enhance the practice of building grapheme-phoneme mappings across the word by drawing children's attention to non-initial parts of the word and emphasizing the importance of these components in word pronunciation. This may be especially relevant in children in the partial alphabetic phase, who have a limited knowledge of grapheme-phoneme connections (Ehri, 2014) and can use pictures as support for graphemes with whose pronunciations they are less familiar. Further study, including conditions with and without phonics instruction, would be required to confirm these hypotheses.

Relatedly, there was a difference in the degree to which the picture-embedded word method impacted acquisition of words that were easier to sound out based on initial letters (his, little) as compared to those that were harder to sound out based on initial letters (all, out). While the picture-embedded words method resulted in greater acquisition of both categories than the word-alone method, it resulted in greater relative acquisition of words that were harder to sound out based on initial letters. With easy-initial words, children were likely able to use phonics to sound out enough of the word to make decoding a helpful strategy, and to rely on these when building associations between the word's appearance and sound (Ehri, 2005). In cases where the word was harder to start decoding through phonics alone, the embedded picture and associated phrase may have served as a particularly useful cue to help children recognize the word. The picture and associated phrase may have helped support orthographic mapping and its use in learning to recognize sight words in cases where orthography was not particularly transparent. This indicates that the method presented here could be especially effective in helping children learn words that are difficult to read through phonic decoding alone. Such words comprise a large proportion of high-frequency words and more generally the English language, given its relative lack of phonetic transparency. However, it is important to note that, given the small set of words used in each category here, further study is needed to draw generalizations about the

effectiveness of this method for words of different degrees of phonetic transparency. It would be interesting and valuable for future work to examine this distinction with a larger set of words carefully divided into phonetically transparent and non-transparent categories, with attention to properties like consistency and regularity (Borleffs, Maassen, Lyytinen, & Zwarts, 2017).

Improvements conferred by the picture-embedded words training did not diminish during the eighteen-day retention period. Differences in test outcome between conditions were strongest at the final retention test, in large part because the treatment group's scores stayed relatively constant while the control group's scores decreased. Retention is rarely examined in this literature, and future work could examine in more detail the impact of embedded pictures as compared to text alone in sight word retention.

While the number of words tested and trained on in this study was consistent with seminal work in this area (Singer et al., 1973), it was nonetheless relatively small. Eyewords has used the same design principles to create picture-embedded words for 150 high-frequency words, and future studies using this method could train students on larger numbers of words to further strengthen the generalizability of the results. It would be interesting, for example, to see whether this method can be used effectively for words containing many of the same letters without children confusing the words.

While this study focused primarily on typically developing learners, many findings in this literature have focused on children with learning differences who face particular difficulty learning to decode and for whom early explicit sight word memorization is thus especially valuable (Blischak & McDaniel, 1995; Tabe & Jackson, 1989; Van der Bijl et al., 2006). A study using this method with children with learning differences would be useful both for determining its effectiveness for this population and for allowing more direct comparison with other studies in this literature.

Through reviewing a number of studies in this literature, we noticed that font of word presentation is highly variable across studies. While we recognize that regional differences exist in the fonts used to teach young children to read, it is possible that fonts used in some previous studies may not be familiar to young children. For example, many beginning readers in North America are used to reading the letter “a” as rendered in Century Gothic font but would not as easily recognize the letter “a” as rendered in Times New Roman font. This is especially important if phonics is used as part of the teaching method. We encourage researchers to consider the kinds of fonts that children are used to reading when designing word learning studies, as font choice is a small but important detail that can impact the ecological validity and generalizability of results.

A couple of limitations in this study prevent us from determining the specific reason for the outcome of this study, and we hope they can serve as inspiration for the design of future studies in this literature. First, it is possible that the semantic support phrase or action alone, and not the picture, led to higher scores in the treatment condition here. Future studies could include conditions with and without pictures and with and without semantic support such as a phrase and/or action. Second, and relatedly, children in the two conditions received different amounts and types of exposure to the words. Words were only said once in the initial introduction of the word alone condition but were repeated while explaining the associated phrase and action in the embedded words condition, with additional time also spent on instruction in the latter condition; while words were later repeated a number of times in the phonics description and throughout the games in both conditions, children in the embedded words condition nonetheless received more exposure to them as a result of these introductory differences. Future studies employing a phrase or action should ideally control carefully for exposure and instruction time to ensure they are not contributing to any observed effect. Acknowledging these limitations, we hope the primary contribution of this study will be to provide evidence that a picture-embedded method can be more effective than text alone while all other studies in this literature have found the opposite, allowing it to lay the groundwork for future studies that could test each of these individual variables to determine the reason for the improved outcomes observed here.

This study compared a condition that included a word, picture, phrase, and action to one that only included a word, with the former also involving a greater amount of exposure to the word. It is possible that not the picture but another feature of the picture-embedded word condition (for example, the phrase, or increased word exposure) was the primary factor underlying word learning. However, given that the picture is highly integrated into the word, and that pictures draw attention and serve as a source of motivation, it is plausible that children attend to and remember the picture in combination with the word. The picture may then provide a trigger for the phrase, which in turn serves as a bridge encouraging the association of the picture with the word. As described, the use of these cues in combination with phonics instruction may also have helped children more powerfully use grapheme-phoneme knowledge to form an association between the word and its sound. While further work is needed to shed greater light on the mechanism, we have observed in this initial study that the combination of these factors is effective in teaching sight words in comparison to text-alone methods, and we hope that this study can serve as a starting point for future exploration of the conditions in which picture-embedded words can be most effective for sight word acquisition.

Declarations of interest

None.

Funding

C. Benjamin Strauber was supported by a United States National Science Foundation Graduate Research Fellowship. This research did not otherwise receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgments

We would like to give special thanks to the students, teachers, and staff of Stoney Creek Public School and Woodland Heights Public School for participating in the study. We'd like to thank the Research and Assessment Department at Thames Valley District School Board, and in particular Steven Killip and Sarah Folino for their support with providing Phonological Assessment data on participants. We are grateful to our research assistants Ankita Patro, Emily Baxendale, Jo Ann Iantosca, Emma Brooks, Alyssa Nicholls, Larissa Pipe, June Doan, Lisa McKenzie, Katherine Reif, Crystal Gentle, Meagan Guy, Judy Walker and Mary Cahalan for their dedicated support. We are especially thankful to Professor Prashant Loyalka for his guidance in evaluating the study. This work would not have been possible without the late Professor Clifford Nass, a guiding light who helped design the study and serves as a continual source of inspiration to seek truth through evidence-based practices.

References

Adams, M. J. (1994). *Beginning to read: Thinking and learning about print.* Cambridge: MIT Press.

Arlin, M., Scott, M., & Webster, J. (1978). The effects of pictures on rate of learning sight words: A critique of the focal attention hypothesis. *Reading Research Quarterly, 14*(4), 645–660.

Blackwell, R., & Laman, S. (2013). Strategies to teach sight words in an elementary classroom. *International Journal of Education, 5*(4), 37–47.

Blischak, D. M., & McDaniel, M. A. (1995). Effects of picture size and placement on memory for written words. *Journal of Speech, Language, and Hearing Research, 38*(6), 1356–1362.

Borleffs, E., Maassen, B. A., Lyytinen, H., & Zwarts, F. (2017). Measuring orthographic transparency and morphological-syllabic complexity in alphabetic orthographies: A

narrative review. *Reading and Writing, 30*(8), 1617–1638.
Browder, D. M., & Lalli, J. S. (1991). Review of research on sight word instruction. *Research in Developmental Disabilities, 12*(3), 203–228.
Brown, A. A. (2014). *The effects of using morphophonic faces as a method for teaching sight words to low-performing kindergartners (Doctoral Dissertation)*LSU Digital Commons Retrieved from.
Didden, R., de Graaff, S., Nelemans, M., Vooren, M., & Lancioni, G. (2006). Teaching sight words to children with moderate to mild mental retardation: Comparison between instructional procedures. *American Journal on Mental Retardation, 111*(5), 357–365.
Didden, R., Prinsen, H., & Sigafoos, J. (2000). The blocking effect of pictorial prompts on sight- word reading. *Journal of Applied Behavior Analysis, 33*(3), 317–320.
Dittlinger, L. H., & Lerman, D. C. (2011). Further analysis of picture interference when teaching word recognition to children with autism. *Journal of Applied Behavior Analysis, 44*(2), 341–349.
Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167–188.
Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading, 18*(1), 5–21.
Ehri, L. C., Deffner, N. D., & Wilce, L. S. (1984). Pictorial mnemonics for phonics. *Journal of Educational Psychology, 76*(5), 880.
Elliott, R. T., & Zhang, Q. (1998). Interference in learning context-dependent words. *Educational Psychology, 18*(1), 5–25.
Fossett, B., & Mirenda, P. (2006). Sight word reading in children with developmental disabilities: A comparison of paired associate and picture-to-text matching instruction. *Research in Developmental Disabilities, 27*(4), 411–429.
Fry, E. B., & Kress, J. E. (2012). *The reading teacher's book of lists.* Hoboken, NJ: John Wiley & Sons.
Harzem, P., Lee, I., & Miles, T. R. (1976). The effects of pictures on learning to read. *British Journal of Educational Psychology, 46*(3), 318–322.
Hill, L. (1995). An exploratory study to investigate different methods for teaching sight vocabulary to people with learning disabilities of different aetiologies. *Down Syndrome: Research and Practice, 3*(1), 23–28.
Lang, R. J., & Solman, R. T. (1979). Effect of pictures on learning to read common nouns. *British Journal of Educational Psychology, 49*(2), 138–149.
Meadan, H., Stoner, J. B., & Parette, H. P. (2008). Sight word recognition among young children at-risk: Picture-supported vs. word-only. *Assistive Technology Outcomes and Benefits, 5*(1), 45–58.
Miller, A., & Miller, E. E. (1971). Symbol accentuation, single-track functioning and early reading. *American Journal of Mental Deficiency, 76*(1), 110–117.
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255.
Pufpaff, L. A., Blischak, D. M., & Lloyd, L. L. (2000). Effects of modified orthography on the identification of printed words. *American Journal on Mental Retardation, 105*(1), 14–24.
Richardson, A. R., Lerman, D. C., Nissen, M. A., Luck, K. M., Neal, A. E., Bao, S., et al. (2017). Can pictures promote the acquisition of sight-word reading? An evaluation of two potential instructional strategies. *Journal of Applied Behavior Analysis, 50*(1), 67–86.
Samuels, S. J. (1967). Attentional process in reading: The effect of pictures on the acquisition of reading responses. *Journal of Educational Psychology, 58*(6), 337.
Savage, R., Stuart, M., & Hill, V. (2001). The role of scaffolding errors in reading development: Evidence from a longitudinal and a correlational study. *British Journal of Educational Psychology, 71*(1), 1–13.
Sheehy, K. (2005). Morphing images: A potential tool for teaching word recognition to children with severe learning difficulties. *British Journal of Educational Technology, 36*(2), 293–301.
Shmidman, A., & Ehri, L. (2010). Embedded picture mnemonics to learn letters. *Scientific Studies of Reading, 14*(2), 159–182.
Singer, H. (1980). Commentary: Sight word learning with and without pictures: A critique of Arlin, Scott, and Webster's research. *Reading Research Quarterly, 15*(2), 290–298.
Singer, H., Samuels, S. J., & Spiroff, J. (1973). The effect of pictures and contextual conditions on learning responses to printed words. *Reading Research Quarterly, 9*(4), 555–567.
Singh, N. N., & Solman, R. T. (1990). A stimulus control analysis of the picture-word problem in children who are mentally retarded: The blocking effect. *Journal of Applied Behavior Analysis, 23*(4), 525–532.
Smeets, P. M., Lancioni, G. E., & Hoogeveen, F. R. (1984). Using stimulus shaping and fading to establish stimulus control in normal and retarded children. *Journal of Intellectual Disability Research, 28*(3), 207–218.
Solman, R. T., Singh, N. N., & Kehoe, E. J. (1992). Pictures block the learning of sight-words. *Educational Psychology, 12*(2), 143–153.
Solman, R. T., & Wu, H. M. (1995). Pictures as feedback in single word learning. *Educational Psychology, 15*(3), 227–244.
Tabe, N., & Jackson, M. (1989). Teaching sight vocabulary to children with developmental disabilities. *Australia & New Zealand Journal of Developmental Disabilities, 15*(1), 27–39.
Van der Bijl, C., Alant, E., & Lloyd, L. (2006). A comparison of two strategies of sight word instruction in children with mental disability. *Research in Developmental Disabilities, 27*(1), 43–55.