Article:

Evaluation of an explicit vocabulary teaching intervention for children learning English as an additional language in primary school

Chris Dixon[iD], Jenny Thomson and Silke Fricke
University of Sheffield, UK

Abstract
Many children learning English as an additional language (EAL) possess lower levels of English vocabulary knowledge relative to their non-EAL English-speaking peers. Longitudinal work suggests that this group discrepancy does not decrease markedly over time as a result of regular classroom teaching. Twelve EAL learners with English vocabulary weaknesses took part in a low-intensity, 10-week vocabulary intervention. Working one-to-one with speech and language therapy students, children took part in weekly activities designed to promote receptive and productive knowledge of 20 target words within the Tier-2 vocabulary category. When assessed on a bespoke word knowledge assessment, children made statistically significant gains between pretest and posttest in both receptive and productive knowledge of taught words, and maintained this knowledge six months later. Multi-component explicit vocabulary instruction offering opportunities for active engagement and discussion is thus presented as one potentially effective means of promoting the Tier-2 vocabulary knowledge of EAL learners in primary school in England.

Keywords
English as an additional language, explicit, intervention, Tier-2, vocabulary

I Introduction

In recent years a number of countries have reported increases in the proportions of school pupils possessing linguistic proficiency in a minority societal language (OECD, 2010). Bilingualism is a dimensional rather than categorical phenomenon, with bilingual experience differing considerably according to children's age of acquisition, patterns of language use, and societal and educational demands (Romaine, 1995). In England, bilingual pupils are referred to as learning 'English as an additional language' (EAL). EAL learners in England are a growing and diverse population, representing just over 1

Corresponding author:
Chris Dixon, Department of Human Communication Sciences, University of Sheffield, 362 Mushroom Lane, Western Bank, Sheffield, S10 2TS, UK; School of Psychology, University of Leeds, LS2 9JZ, UK.
Email: cddixon90@gmail.com

in 5 (21.2%) of primary school learners (Department for Education, 2019). Despite the binary nature of the EAL label, these children possess varying levels of English language proficiency at school entry, ranging from fully fluent to new to English. As EAL learners are educated alongside their non-EAL English-speaking peers, there is an expectation that English will be acquired solely through engagement with the curriculum and interaction with peers (Costley, 2014). However, EAL learners consistently underperform relative to non-EAL children on national assessments of reading and writing (Strand et al., 2015), suggesting that more could be done to cater towards the educational needs of these pupils.

International literature indicates that bilingual learners tend to exhibit certain profiles of strengths and weaknesses or 'profile effects' (Oller et al., 2007) across language and literacy skills. These learners often underperform relative to their monolingual peers on measures of receptive and expressive vocabulary, grammar, and listening and reading comprehension, but not measures of phonological awareness or single-word reading (Bialystok et al., 2010; Droop and Verhoeven, 2003; Geva and Farnia, 2012). Furthermore, longitudinal studies also tend to report that such group discrepancies remain in place or close only to a small degree over time (Droop and Verhoeven, 2003; Lervåg and Aukrust, 2010). This is particularly evident for progress in vocabulary knowledge, which is highly dependent upon patterns of linguistic input.

A similar picture of profile effects emerges among samples of EAL learners in England (Babayiğit, 2014; Burgoyne et al., 2011; Hutchinson et al., 2003). Lower levels of English vocabulary knowledge have also been found in EAL learners in secondary school (age 14–15), hinting at the persistent nature of vocabulary learning needs in this population (Cameron, 2002). Lower levels of vocabulary knowledge may have implications for other aspects of children's learning and attainment, as vocabulary is a well-established predictor of listening and reading comprehension (Verhoeven and van Leeuwe, 2008; Lervåg and Aukrust, 2010). In a study of 56 monolingual and 69 EAL learners aged 9 to 10, Babayiğit (2014) found a significant interaction between language group and vocabulary in a regression model predicting reading comprehension performance. This finding suggests that although vocabulary knowledge is important for all children, it may play a stronger role in the reading comprehension performance of EAL learners.

Although bilingual learners are characterized by a high degree of heterogeneity in terms of language and literacy skills, their lower levels of vocabulary knowledge and comprehension may be said to overlap to some extent with the profiles of other groups of children with speech, language and communication needs (SLCN), for instance those with specific reading comprehension difficulties (Nation et al., 2004). While children with EAL do not necessarily fall under the category of SLCN, the language and reading skills profile discussed above may certainly constitute communication needs for a subset of these children, particularly those who are not exposed to English in the home or are new to English at the point of school entry.

In summary, there is mounting evidence from the international literature as well as studies emerging from the UK that bilingual learners may be at risk for lower levels of target language vocabulary knowledge, which may impact upon other skills and ultimately educational attainment. Although children do acquire vocabulary incidentally through exposure to oral and written language, there is evidence to suggest that this is conditioned by pre-existing word knowledge and verbal skills (Swanborn and de Glopper, 1999); that is, children with lower levels of vocabulary knowledge may be less likely to acquire vocabulary successfully in an incidental fashion, justifying emphasis on explicit instruction.

1 Effective vocabulary instruction

Vocabulary intervention studies employ a wide range of word teaching strategies, including the provision of explicit definitions (McKeown et al., 1983; Nash and Snowling, 2006), semantic maps

and illustrations (Clarke et al., 2010; Nash and Snowling, 2006; Rupley and Nichols, 2005), connecting new words to previous knowledge and personal experiences (Goerrs et al., 1999), and affording opportunities to use new vocabulary (Gillanders et al., 2014).

Balance between definitional and contextual information is shown to impact the efficacy of vocabulary instruction. In definitional approaches, word meanings are presented in the form of concrete definitions, while in contextual approaches word meanings are provided through cues typically in story contexts (Stahl and Fairbanks, 1986). Although both approaches are shown to be effective, meta-analyses indicate significantly larger vocabulary gains from studies utilizing combined or definition-only approaches (Stahl and Fairbanks, 1986; Marulis and Neuman, 2010). Placement of target vocabulary in context more closely resembles free reading and may be more enjoyable for children (Nash and Snowling, 2006). There is evidence for the facilitative role of active or deep processing in word learning, broadly defined as the amount and quality of attention paid towards target words (Hulstijn and Laufer, 2001).

Deep processing is typically evidenced through conditions in which learners make use of sentential context, provide novel examples, form associations, and engage in discussion around word meanings (Bowyer-Crane et al., 2008; Dockrell et al., 2010; McKeown, et al., 1983; Nation, 2001). Again, meta-analyses provide evidence for a positive association between depth of processing and gains in vocabulary knowledge (Elleman et al., 2009; Stahl and Fairbanks, 1986), which may be achieved through reflection on personal experiences and recording encounters with target vocabulary (Clarke et al., 2010; Dockrell et al., 2010; Nash and Snowling, 2006).

2 Vocabulary intervention for bilingual learners

A number of successful oral language interventions for bilingual learners are reported in the international literature. Silverman (2007) compared the word-learning gains of 72 mono- and bilingual typically developing kindergartners in the US who engaged with novel vocabulary through children's literature, child-friendly definitions, visual aids, and opportunities to use words in context. The study showed significant gains for both groups of children, although bilingual learners gained relatively more word knowledge by the end of the 14-week programme. Other work has found similarly promising results among older children. For example, Carlo et al. (2004) carried out a 15-week vocabulary intervention among 169 mono- and bilingual students in the US aged 10–11 years. In each of a series of weekly cycles, students were exposed to 10–12 target words around topics encouraging debate (e.g. immigration) and took part in activities designed to promote depth of knowledge such as cloze tasks, associations and synonyms, and morphological analysis. Taking into account prior English receptive vocabulary knowledge, mono- and bilingual students in the intervention made significantly larger gains in knowledge of target words than control group peers.

Recent reviews indicate a lack of oral language interventions for EAL learners in the UK (Murphy and Unthiah, 2015; Oxley and de Cat, 2019). Despite this, a small number of studies do offer promise for oral language and vocabulary intervention with very young EAL learners in England. Dockrell et al. (2010) report an evaluation of the Talking Time intervention among a diverse sample of 96 nursery school-aged EAL learners (age 3–4) from relatively socially deprived communities. After 15 weeks of 15-minute small group oral language teaching twice a week, children showed significant improvements relative to a control group in vocabulary knowledge of taught words, as well as other aspects of verbal comprehension and production. Crucially, teachers were trained to employ strategies such as modelling, re-casting, and relating teaching to personal experiences. Similarly promising results are reported by St. John and Vance (2014). In this study, a sample of 18 Year 1 children (aged 5–6) with EAL and/or SLCN were selected by special educational needs co-ordinators in 22 schools to take part. Participants received vocabulary training for

| Baseline | 6 weeks | Pretest | 10-week intervention | Posttest | 6 months | Maintenance |
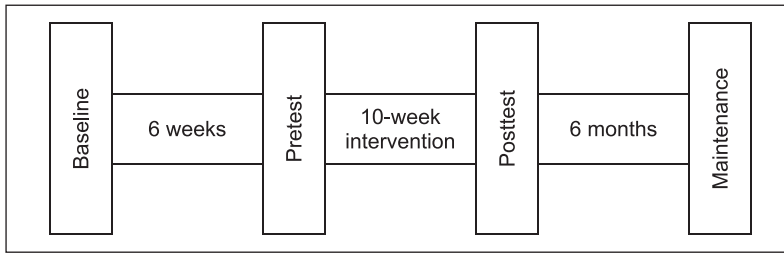
Figure 1. Intervention timeline indicating intervals between time points.

10 to 15 minutes per day, and vocabulary sourced from curriculum content was taught using definitions, mind maps, word association, true/false judgement, and use in context. Although target vocabulary was not held constant across participants, significant gains in word knowledge were observed for taught words after a period of 3–4 weeks of teaching relative to control words which were also assessed before teaching began.

Few studies have evaluated oral language intervention in older EAL learners in England. Kotler et al. (2001) assessed the efficacy of Talking Partners among a sample of 64 5- to 8-year-old EAL learners. In small, 20-minute, weekly group sessions delivered by parent volunteers, children took part in problem-solving activities affording opportunities for extended talk, but no specific vocabulary was targeted. After 10 weeks, children showed significant increases relative to a control group in the Information subscore of the Renfrew Action Picture Test (Renfrew, 1988), but not on measures of vocabulary or grammar.

In summary, the combination of multiple vocabulary teaching methods, as well as opportunities for extended talk and verbal engagement, appear to offer promise for supporting the lexical knowledge of bilingual learners. However, what little work has been conducted in England has typically been carried out with very young EAL learners, and therefore less is known concerning the efficacy of explicit vocabulary instruction for children beyond the initial stages of English language acquisition.

II The present study

The present study aimed to design, deploy, and evaluate a bespoke short-term vocabulary intervention for EAL learners in Year 4 (age 8–9) of primary school. The research questions were as follows:

1. To what extent does a short, low-intensity, one-to-one explicit vocabulary training programme improve target vocabulary knowledge in EAL learners who have vocabulary weaknesses?
2. What effect does such teaching have upon children's ability to give a definition and their ability to use words appropriately within a sentence?

III Method

Participants were assessed on their knowledge of preselected taught and untaught vocabulary at four time points (see Figure 1). Due to the lack of a control group, a 'double pre-test' design was employed in order to monitor maturational trends prior to teaching (Shadish et al., 2002), between

baseline and pretest. Additionally, children were assessed at each time point on a set of untaught control words in order to examine the specific effect of the intervention teaching.

1 Participants

Participants were recruited from a larger pool of 48 EAL learners taking part in a longitudinal study in South Yorkshire, England (Dixon, Thomson and Fricke (2020)). The purpose of this study was to follow the language and literacy development of these children over a period of 18 months. All EAL learners had been educated in mainstream primary schools since at least Year 1 (age 5–6), spoke a range of 15 different languages in the home, and had no history or statement[1] of special educational needs. Parents of EAL learners in the longitudinal study were asked to complete language background questionnaires: these were returned for 10 of the 12 intervention participants. Of these, only one child had been born outside of the UK, and English was not spoken at all at home for only two children. All but one child had begun their English-medium education in nursery (age 3–4) or Reception (age 4–5). To be eligible for the intervention, participants had to obtain a standard score of −1 SD on two out of three of the following measures at Time 1 of the longitudinal study: The British Picture Vocabulary Scale III (BPVS; Dunn et al., 2009); The Clinical Evaluation of Language Fundamentals Expressive Vocabulary subtest (CELF EV; Semel et al., 2006); and the Wechsler Intelligence Scale for Children IV Vocabulary definitions subtest (WISC; Wechsler, 2003). The reason for this criterion was to target intervention teaching at EAL learners with lower levels of receptive and expressive word knowledge. Informed parental consent was received for 12 of the 23 children identified as eligible to take part. Participants (3 boys; 9 girls) were aged between 8;9 and 9;7 years (mean = 9;1).

2 Selection of target words

Twenty Tier-2 target words were selected for teaching. Tier-2 describes words that are of 'high frequency for mature language users and are found across a variety of domains' (Beck et al., 2002: 8), with examples including more nuanced synonyms for concepts children are already likely to possess such as *fortunate* (lucky) or *benevolent* (kind). Tier-2 vocabulary was chosen due to its appropriateness for children past the initial stages of learning English (i.e. due to the likelihood of 8- to 9-year-old EAL learners having acquired a good deal of Tier-1 vocabulary). As the boundary between vocabulary tiers is not clear-cut, the present study also incorporated quantitative psycholinguistic information in order to select target words. A pool of potential words was selected from the database of Kuperman et al. (2012) – particularly, words with an age of acquisition (AoA) of between 6 and 10 years, and a frequency of occurrence of between 5 and 10 per one million words. A range of AoA and frequency values was chosen in order to vary the difficulty of target vocabulary. Due to the lack of a control group, a parallel list of 10 untaught words was selected in order to measure the specific effect of the intervention teaching. Untaught words were selected by matching additional word candidates from the Kuperman et al. (2012) database with target words as closely as possible in AoA and frequency. Target words were grouped into five themes in order to encourage discussion during teaching, but untaught words were not themed as these were matched purely on psycholinguistic variables (see Appendix 1).

3 Session content

The aim of the intervention was to impart and deepen knowledge of Tier-2 vocabulary using both definitional and contextual information through active engagement. In each session children

Table 1. Intervention structure activities and timing.

Activity	Details	Duration (minutes)
Introduction	State aims for session; play word game for warm-up; consolidation from last session	3
Vocabulary teaching (first word)	Passage, sentence judgement/completion, semantic map, sentence writing	10
Vocabulary teaching (second word)	Passage, sentence judgement/completion, semantic map, sentence writing	10
Plenary	Recap using flashcards and child-friendly definitions	2

worked one-to-one with an intervention co-ordinator and were exposed to two target words using the session structure outlined in Table 1.

Activities are described here in chronological order:

- Word game: Learners' motivation as well as their attention to and awareness of words are important factors in vocabulary teaching (Beck et al., 2002). Therefore, five different word games (e.g. hangman, word association) were employed throughout the ten weeks in order to promote children's engagement and awareness of words. Sessions 2 to 10 also began with a brief re-cap of words covered the previous week.
- Passage reading: In line with contextual approaches to vocabulary instruction (Stahl and Fairbanks, 1986), target words were presented within short written passages depicting stories of characters and events to illustrate use of the target word within context, for example a narrative of two friends who make a trip to a museum but are required to take public transport because it is located in a *distant* town. Passages ranged from 83 to 193 words and were written explicitly to contain cues to word meanings, e.g. for target word *distant*: 'Thorpetown is quite distant, so it's a long way to walk.'
- Sentence judgement / completion: In sentence judgement tasks, children were required to decide whether each of two sentences containing the target word made correct or incorrect use of that word. For example: 'London and Liverpool are *distant* from each other, so it's quick to travel between them' (Incorrect; it takes a long time because they are distant from one another). Sentence completion tasks were cloze-like activities in which children were asked to fill in missing content in sentence frames. For the target word *capital*, for example: 'I'm going to the capital to . . .' (e.g. see a landmark typical of a capital such as Big Ben).
- Semantic map: After the target word was written in the centre of a blank page, children were prompted to consider word associations, idioms, synonyms, antonyms, related concepts, and personal experiences relating to the word (e.g. *can you think of a time when you went somewhere distant?*) This allowed for a high degree of creativity, engagement, and interaction between the child and intervention co-ordinator.
- Sentence writing: Children were prompted to write their own sentence using the target word. Intervention co-ordinators prompted children to extend or enhance their sentences where possible.
- Flashcards and child-friendly definitions: At the end of each session, coordinators provided short 'child-friendly' definitions of the target words in an attempt to consolidate children's learning. Such definitions have been shown to result in greater improvements in children's word learning (McKeown, 1993). Additionally, visual aids are used in some vocabulary intervention studies (Clarke et al., 2010; Hairrell et al., 2011) and are recommended for

active processing of word meaning (Rupley and Nichols, 2005). Thus, at the end of each session children were shown flashcards containing the target word next to a colour photograph or illustration conveying the appropriate concept.

4 Intervention delivery and implementation

The intervention consisted of 10 weekly one-to-one sessions of approximately 25 minutes duration. Ten weeks was chosen in order to strike a balance between amount of instructional time and the logistic burden on participating schools. Additionally, vocabulary interventions of this duration or even shorter are found to result in significant word learning (Marulis and Neuman, 2010), as well as interventions delivered by trained university students (Elbaum et al., 2000) and in one-to-one or small groups of bilingual learners (Ross and Begeny, 2011). Teaching was carried out by the researcher and nine intervention co-ordinators enrolled in speech and language sciences/therapy degree courses. All co-ordinators possessed experience of working with young children in educational or clinical settings, and received two hours of training on intervention delivery and administration of the word learning measure (see below). Fidelity of implementation was measured through feedback forms completed by coordinators after each session. Specifically, this included dates, start and end times of sessions, the extent to which each activity was completed (zero; partial; full), and children's levels of attention and engagement (0–5), where attention was defined as making eye contact with the coordinator and the learning materials and not being distracted, and engagement was defined as willingness to provide answers to questions and engage in discussion around the topic. Co-ordinators could also provide free-text comments.

5 Measure of word learning

The decision was made to create a bespoke measure of word knowledge in the present study, as such measures are shown to be more sensitive than standardized measures to growth in word knowledge (Elleman et al., 2009). This measure assessed word knowledge both in terms of definitions and use within sentences, as detailed below.

a Word score (definitions). Performance on vocabulary definition tasks is said to require certain metalinguistic skills and is shown to be challenging for young children (Benelli et al., 2006). This has led other studies to adopt a different approach by acknowledging and awarding points for various aspects of word knowledge such as background knowledge, context, and gesture (e.g. Hadley et al., 2016), and this approach was adopted in the present study. Children were shown each word (verbally and in writing) and asked 'What does [this word] mean'? Points were awarded across four categories up to a maximum of 8.5 points for each word:

- Definitional information: 1 point awarded for partial or underspecified definition (e.g. *to rescue someone is to help them*), and 2 points awarded for a full or more highly abstract definition (e.g. *to rescue someone is to save them from a dangerous situation*);
- Background knowledge: up to 4 points per word, split into three categories:
  - Situational information: up to 1 point for a hypothetical or real situation demonstrating understanding of the target word (e.g. 'if your friend said bad things about you, you would feel *miserable*');
  - Contextually-related concepts or referents: up to 1 point (e.g. 'sky' or 'cockpit' for *pilot*);

- ○   Attributes or functions: 1 point for one attribute or function and 2 points for two or more attributes or functions (e.g. 'you have a red face when you're *furious*');
- Lexical knowledge: up to 1 point in each of two subcategories, including:
  - ○   Relevant synonyms such as 'excited' for *thrilled*, or antonyms such as 'agree is the opposite of *disagree*';
  - ○   Related words and phrases, for example morphologically related words such as 'distance' for *distant*, or collocational knowledge such as '*rescue* attempt' or 'the *coast* is clear';
- Non-verbal responses: up to 0.5 points (e.g. pointing to a far object to illustrate *distant*, or showing an angry face for *furious*).

b Sentence score.   Word knowledge may also be evidenced through understanding of grammatical function, collocation, and other constraints of use (Nation, 2001). For this reason, a sentence production task was incorporated in order to allow examination of children's productive knowledge of target words. Rather than provide cues to syntactic function (e.g. 'What would you be doing if you were *halting*? as reported in Coyne et al., 2010), children were simply asked 'Can you put [this word] in a sentence?' Points were awarded across three categories up to a maximum of 5 for each word:

- Syntax: 1 point for using target word as correct part of speech, for example correctly utilizing *distant* as an adjective.
- Morphology: 1 point for lack of any morphological error on target word (e.g. 'he *donate* money'). Any sentence scoring 0 for syntax automatically received a score of 0 for morphology.
- Semantics: One limitation of sentence production tasks is that a child may produce a sentence which, albeit correct, does not reveal the characteristic features of a word. The approach taken was to award more points to sentences that did reveal such characteristic features, while not entirely penalizing more 'generic' sentences (e.g. McKeown, 1993). 1 point was awarded for a very simple sentence such as 'I was *miserable*'; 2 points for a more explicit use of the target word with additional information, such as 'I was *miserable* because I was cold'; 3 points for a well-specified sentence giving reason, context, an example, or synonyms, e.g. 'I was *miserable* and felt like crying because my friend moved away'.

The word knowledge measure included two practice items, *library* and *remember*, on which children received feedback to illustrate examples of permissible responses. Assessments were carried out by the first author at baseline, posttest, and maintenance. Due to logistical restrictions, intervention coordinators carried out assessments at pretest. All baseline responses on the bespoke word knowledge assessment were independently scored by a doctoral student in the same university department. Cohen's kappa was calculated separately for all categories of word and sentence scores at baseline. All kappa values were statistically significant at the 0.01 level and fell within the 'substantial' to 'excellent' range of agreement (total word score: .79; total sentence score: .71), with the exception of lexical scores which represented 'moderate' or 'fair' agreement (.45; Landis and Koch, 1977). All disagreements were discussed and resolved.

IV Results

Out of the 12 children who received intervention teaching, one child was not available for testing at immediate posttest, and two children were absent at 6-month maintenance test. This resulted in

Table 2. Implementation fidelity.

Dosage	Mean (SD)	Range
Total sessions completed (maximum = 10)	8.67 (1.56)	6–10
Average session duration (minutes)	26.32 (6.41)	10–40
Mean activity completion rate (maximum = 2):		
Passage reading	1.98 (0.20)	0–2
Comprehension questions	1.98 (0.19)	0–2
Mind map	1.88 (0.41)	0–2
Sentence writing	1.83 (0.49)	0–2
Sentence judgement/completion	1.81 (0.59)	0–2
Attention and engagement:		
Mean level of engagement (maximum = 5)	3.96 (1.01)	1–5
Mean level of attention (maximum = 5)	3.79 (1.11)	1–5

sample sizes of $n = 12$ at baseline and pretest, $n = 11$ at posttest, and $n = 10$ at maintenance test. Therefore, the present study employs an intention-to-treat analysis (Gupta, 2011) by retaining all available data.

1 Implementation

As indicated in Table 2, the average number of completed sessions was high, and average session duration was around 26 minutes. Mean levels of engagement and attention were fairly high, and did not change appreciably throughout the intervention. Major themes in coordinators' text comments included noisy and disruptive working conditions in schools, the substantial amount of input and prompting required by some children, and some children's difficulty explaining or justifying their answers. However, evidence from coordinators' feedback forms suggested that children generally enjoyed discussing their personal experiences and that flashcards and child-friendly definitions helped to disambiguate word meanings.

Given the small sample size, the decision was made to use non-parametric statistics, including the use of medians and interquartile ranges as measures of central tendency and dispersion. Repeated measures (paired samples) Wilcoxon signed-rank tests were conducted to analyse the statistical significance and magnitude of changes in children's scores between each pair of time points. Interpretation of effect sizes (r) between each time point follows that of Cohen (1988), whereby 0.2 is considered small, 0.5 is medium, and 0.8 is large. Given a high number of comparisons, a Bonferroni adjusted alpha level was applied to analysis of taught and untaught words, respectively (.05 / 9 = .006). Results for taught and untaught words are presented in Table 3 and Table 4, respectively. Group and individual trajectories for taught and untaught words are also displayed in Figure 2 (also, for a table of participants' individual scores, see Appendix 2).

2 Progress in total score for taught and untaught control words

a   Baseline to pretest. Change in Total score (Word plus Sentence score) for the 20 taught words between baseline and pretest was not statistically significant ($Z = -0.49$, $p = .622$, $r = .14$). Similarly, no significant progress was observed for the 10 untaught words ($Z = -1.73$, $p = .083$, $r = .35$), suggesting that children were not making an appreciable rate of progress on the bespoke word scoring assessment prior to the start of intervention teaching.

Table 3. Descriptive and inferential statistics for taught words: median (interquartile range in parentheses).

	Baseline	Pretest	Posttest	Maintenance	Baseline to pretest			Prestest to Posttest			Posttest to Maintenance		
					Z	p	r	Z	p	r	Z	p	r
Total score	54.00 (20.25)	54.25 (5.25)	90.00 (36.00)	91.00 (39.50)	−0.49	.622	.14	−2.89	.004	.62	−0.95	.343	.21
Word score (maximum 170)	14.50 (9.25)	14.50 (5.62)	29.50 (16.25)	30.50 (19.25)	−0.85	.396	.17	−2.89	.004	.62	−1.12	.261	.24
Sentence score (maximum 100)	38.50 (11.25)	41.50 (8.25)	58.00 (24.50)	59.00 (19.50)	−0.83	.409	.17	−2.90	.004	.62	−0.83	.407	.18
n	12	12	11	10									

Note. r = effect size for Wilcoxon signed-rank test.

Table 4. Descriptive and inferential statistics for untaught words: median (interquartile range in parentheses).

	Baseline	Pretest	Posttest	Maintenance	Baseline to pretest			Prestest to Posttest			Posttest to Maintenance		
					Z	p	r	Z	p	r	Z	p	r
Total score	19.00 (10.25)	20.50 (12.75)	26.00 (8.50)	31.00 (13.25)	−1.73	.083	.35	−2.67	.008	.57	−1.25	.213	.27
Word score (maximum 85)	6.50 (4.25)	6.00 (6.50)	9.00 (4.50)	9.00 (7.75)	−0.67	.504	.14	−2.12	.034	.45	−1.49	.137	.33
Sentence score (maximum 50)	12.00 (3.50)	13.50 (7.50)	17.00 (8.50)	23.00 (6.00)	−1.66	.097	.34	−2.48	.013	.53	−1.06	.287	.23
n	12	12	11	10									

Note. r = effect size for Wilcoxon signed-rank test.
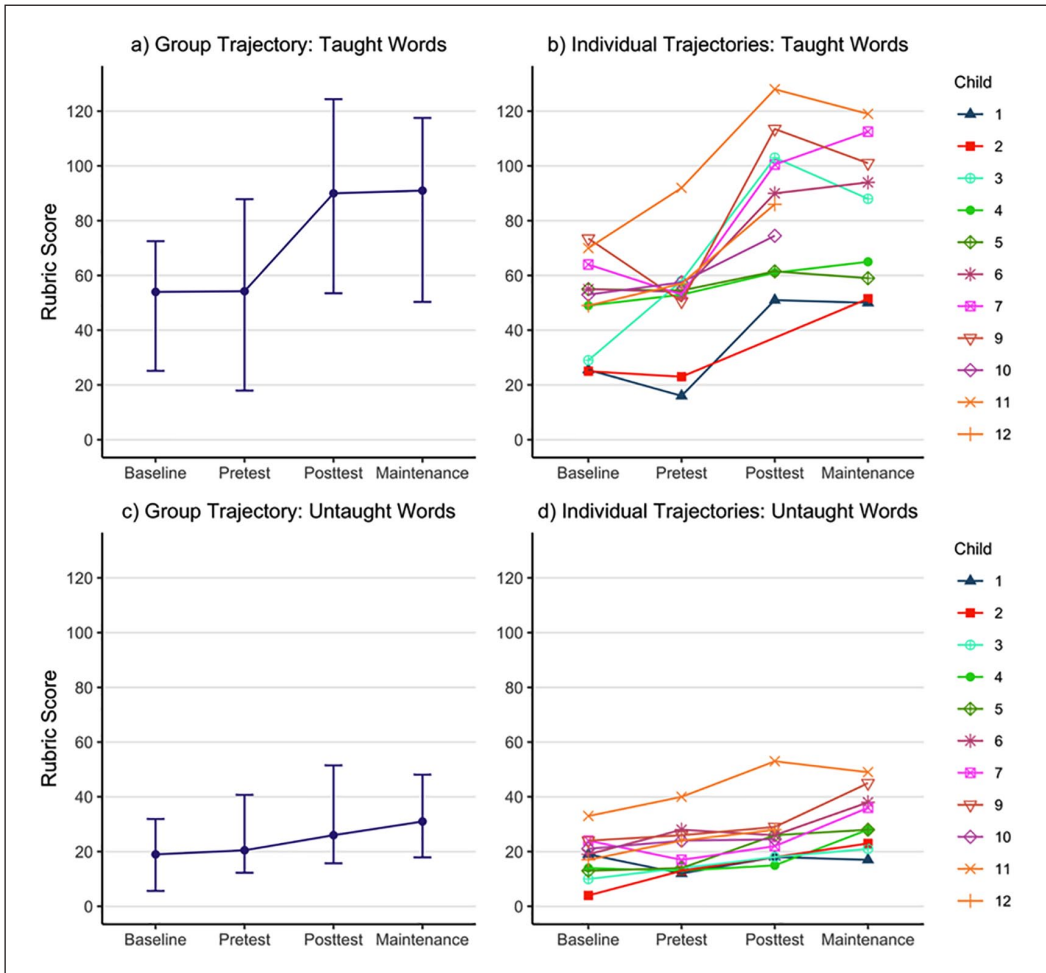
Figure 2. Group and individual trajectories for taught words and untaught words.
Notes. Scores represent total score (word score plus sentence score) for taught words (upper panels) and untaught words (lower panels).

b   Pretest to posttest. Total scores increased to a moderate and statistically significant degree between pretest and posttest for taught words ($Z = -2.89$, $p = .004$, $r = .62$). Although children made significant progress by posttest, this knowledge was by no means at ceiling level, with children scoring a median of 90 points out of a possible of 270 (range: 51–128). Progress in untaught words was observed, but did not reach the Bonferroni criterion ($Z = -2.67$, $p = .008$, $r = .57$).

c   Posttest to maintenance. Total scores did not change significantly between posttest and maintenance test six months later for either taught words ($Z = -0.95$, $p = .343$, $r = .21$) or untaught words ($Z = -1.25$, $p = .213$, $r = .27$). Therefore it appeared that for the most part children had retained knowledge of target words as a result of explicit instruction.

3 Progress in word score and sentence score

a   Baseline to pretest. As for Total score, children did not make significant progress in either Word score ($Z = -0.85$, $p = .396$, $r = .17$) or Sentence score ($Z = -0.83$, $p = .409$, $r = .17$). This pattern also applied to untaught words.

b   Pretest to posttest. By the end of intervention teaching, children had made a significant rate of progress in both Word score ($Z = -2.89$, $p = .004$, $r = .62$) and Sentence score ($Z = -2.90$, $p = .004$, $r = .62$). Again, although progress was observed in scores for untaught words, this did not reach the Bonferroni criterion.

c   Posttest to maintenance. As for the pattern observed in Total score, no significant change was observed between posttest and maintenance for either Word score ($Z = -1.12$, $p = .261$, $r = .24$) or Sentence score ($Z = -0.83$, $p = .407$, $r = .18$). This pattern applied equally to untaught words.

V Discussion

The present study assessed the effects of a short-term, low-intensity Tier-2 vocabulary intervention for EAL learners with English vocabulary weaknesses. Taking part in activities providing contextual and definitional information and focusing on active engagement and personal experiences, children showed a moderate and significant gain in knowledge of taught words immediately after teaching, and this knowledge was largely retained six months later. Furthermore, this gain was observed both in the ability to give definitions and use target words within sentences. In general, results accord with similarly-focused studies in the literature which show that vocabulary knowledge is responsive to explicit instruction (Elleman et al., 2009; Marulis and Neuman, 2010). However, it contrasts with some previous EAL-focused intervention studies in the UK by exclusively targeting vocabulary knowledge (e.g. Dockrell et al., 2010; Kotler et al., 2001).

Results of the intervention bear close comparison to those of St. John and Vance (2014), and build on some of the limitations of this study: first, by holding vocabulary constant across all intervention participants and therefore ensuring fair comparison across children in terms of words and, second, by selecting vocabulary independently of curriculum content and therefore reducing chances of reinforcement through classroom instruction. The specificity of the intervention teaching is demonstrated through lack of statistically significant improvements in knowledge of untaught words; however, as children continue to acquire word knowledge over time, improvements in untaught vocabulary have been reported elsewhere in the literature (St. John and Vance, 2014; Wilkinson and Houston-Price, 2013).

The intervention purposefully drew upon strategies that have been shown to result in effective word learning. Particularly, the presentation of target words within story contexts appeared to aid children's comprehension and retention of vocabulary with some children drawing on examples from passages when providing definitions or giving example sentences. It is likely that the wide range of activities helped to tease out and consolidate children's understanding of target words: analysis of completed mind-maps revealed evidence of discussions on, for instance, the difference between target word *coast* and other water-related vocabulary such as 'river' and 'lake', and the difference between target word *disagree* and 'dislike'. Such discussion may have encouraged depth of processing. Finally, certain aspects of intervention delivery may have accounted for gains in vocabulary knowledge. First, studies indicate that one-to-one working arrangements may be more effective than group- or class-level teaching (Chi et al., 2001), and that bilingual learners may benefit especially from this (Ross and Begeny, 2011). Second, studies show that trained student

volunteers can successfully deliver intervention teaching (Elbaum et al., 2000), and this is likely to have applied to student co-ordinators in the present study, the majority of whom were in training to become speech and language therapists.

The study employed a bespoke measure of word knowledge, allowing a more sensitive assessment of growth over time (Elleman et al., 2009). In particular, the inclusion of a productive element of vocabulary was advantageous in flagging children who, for instance, did not use target words as the correct part of speech (e.g. my friend *thrilled* at his brother; yesterday I *fraud*). Indeed, errors in productive knowledge appeared to be a sensitive measure of which children were likely to benefit from the intervention teaching, as children who made these types of errors made very little progress as a result of the explicit teaching. In a similar vein, allowing background knowledge such as situational information and contextually-related words or phrases as permissible responses in the definition task allowed children to demonstrate their word knowledge in other ways (Hadley et al., 2016), for example *the boats and the beach* (for 'coast') or *like if a cat's stuck in a tree, you ring the firefighter* (for 'rescue'). Adult-style definitions place high demand on metalinguistic knowledge and are likely to be even more demanding for children with low levels of vocabulary knowledge. This, along with story passages, appeared to provide children with background knowledge and context to support their understanding.

The lack of a control group is a limitation of the study. This was mitigated to some extent by the inclusion of a baseline period to measure growth prior to teaching, as well as a parallel list of untaught words to measure the specific effect of the intervention. Future work may seek to evaluate such an intervention with a larger sample of children and to incorporate a control group of EAL learners, or alternatively to compare efficacy of teaching among EAL and non-EAL pupils with low levels of English vocabulary knowledge. A bespoke word knowledge scoring rubric was created for the purposes of this study and as such, transfer beyond the specifically targeted words was not possible to measure. Although an attempt was made to assess the interrater reliability of the rubric, it was only possible to carry this out at baseline and not at additional time points, potentially affecting the reliability of children's word and sentence scores over time. The robustness of scores would also have been improved by blinding during the assessment process, which was not possible given time and resource limitations. As a result, caution must be used in the interpretation of the results presented here.

Although this study yielded tentatively positive results for EAL learners' word learning, it is unknown to what extent such gains would generalize to other settings or working patterns. Further evaluation would be required to establish whether such a programme could be provided to children in small groups, taught by teaching assistants or other paraprofessionals, and using different materials or target words. Such a vocabulary intervention programme may also be of utility to other children with SLCN as part of a package of ongoing support. For instance, such multifaceted instruction may promote the word knowledge of children lacking exposure to 'book language' in the home, or children who experience difficulties with word learning.

In conclusion, the results of this study provide some evidence that multi-component vocabulary instruction may be an effective means of improving the English vocabulary knowledge of 8- to 9-year-old EAL learners with English vocabulary weaknesses. Longitudinal work has shown that vocabulary knowledge discrepancies between EAL learners and their non-EAL peers are unlikely to narrow during primary school as a result of regular classroom instruction as EAL learners aim at a 'moving target'. Although not all vocabulary knowledge can be explicitly taught, there is scope for teaching selectively targeted Tier-2 vocabulary which may facilitate EAL learners' access to school curricula, and ultimately promote their educational attainment.

Declaration of conflicting interests

Funding

ORCID iD

Chris Dixon ⬤ https://orcid.org/0000-0002-5125-2107

Supplemental material

Supplemental material for this article is available online.

Note

1. Special educational needs (SEN) statements have now been replaced with Education, Health and Care plans. Previously, statements may have been issued to children with SEN whose needs fall outside the remit of the school-level support.

References

Babayiğit S (2014) The role of oral language skills in reading and listening comprehension of text: A comparison of monolingual (L1) and bilingual (L2) speakers of English language. *Journal of Research in Reading* 37: 22–47.

Beck IL, McKeown MG, and Kucan L (2002) *Bringing words to life: Robust vocabulary instruction*. New York: Guilford Press.

Benelli B, Belacchi C, Gini G, and Lucangeli D (2006) To define means to say what you know about things: The development of definitional skills as metalinguistic acquisition. *Journal of Child Language* 33: 71–97.

Bialystok E, Luk G, Peets KF, and Yang S (2010) Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition* 13: 525–31.

Bowyer-Crane C, Snowling MJ, Duff FJ, et al. (2008) Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry* 49: 422–32.

Burgoyne K, Whiteley HE, and Hutchinson JM (2011) The development of comprehension and reading-related skills in children learning English as an additional language and their monolingual, English-speaking peers. *British Journal of Educational Psychology* 81: 344–54.

Cameron L (2002) Measuring vocabulary size in English as an additional language. *Language Teaching Research* 6: 145–73.

Carlo MS, August D, McLaughlin B, et al. (2004) Closing the gap: Addressing the vocabulary needs of English-Language Learners in bilingual and mainstream classrooms. *Reading Research Quarterly* 39: 188–215.

Chi M, Siler S, Jeong H, Yamauchi T, and Hausmann R (2001) Learning from human tutoring. *Cognitive Science* 25: 471–533.

Clarke PJ, Snowling MJ, Truelove E, and Hulme C (2010) Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science* 21: 1106–16.

Cohen J (1988) *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Costley T (2014) English as an additional language, policy and the teaching and learning of English in England. *Language and Education* 28: 276–92.

Coyne M, McCoach B, Loftus S, et al. (2010) Direct and extended vocabulary instruction in kindergarten: Investigating transfer effects. *Journal of Research on Educational Effectiveness* 3: 93–120.

Department for Education (2019) *Schools, pupils and their characteristics: January 2019 (National tables)*. London: Department for Education.

Dixon C, Thomson J, and Fricke S (2020) Language and reading development in children learning English as an additional in primary school in England. *Journal of Research in Reading*. Epub ahead of print 30 April 2020. DOI: 10.1111/1467-9817.12305.

Dockrell JE, Stuart M, and King D (2010) Supporting early oral language skills for English language learners in inner city preschool provision. *British Journal of Educational Psychology* 80: 497–515.

Droop M and Verhoeven L (2003) Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly* 38: 78–103.

Dunn L, Dunn D, and NFER (2009) *British Picture Vocabulary Scale*. 3rd edition. London: GL Assessment.

Elbaum B, Vaughn S, Hughes MT, and Watson Moody S (2000) How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology* 92: 605–19.

Elleman AM, Lindo EJ, Morphy P, and Compton DL (2009) The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness* 2: 1–44.

Geva E and Farnia F (2012) Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Reading & Writing* 25: 1819–45.

Gillanders C, Castro DC, and Franco X (2014) Learning words for life: Promoting vocabulary in dual language learners. *Reading Teacher* 68: 213–21.

Goerrs B, Beck I, and McKeown M (1999) Increasing remedial students' ability to derive word meaning from context. *Reading Psychology* 20: 151–75.

Gupta S (2011) Intention-to-treat concept: A review. *Perspectives in Clinical Research* 2: 109–12.

Hadley EB, Dickinson DK, Hirsh-Pasek K, Golinkoff RM, and Nesbitt KT (2016) Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly* 51: 181–98.

Hairrell A, Rupley W, and Simmons D (2011) The state of vocabulary research. *Literacy Research and Instruction* 50: 253–71.

Hulstijn J and Laufer B (2001) Some empirical evidence for the Involvement Load Hypothesis in vocabulary acquisition. *Language Learning* 51: 539–58.

Hutchinson JM, Whiteley HE, Smith CD, and Connors L (2003) The developmental progression of comprehension-related skills in children learning EAL. *Journal of Research in Reading* 26: 19–32.

Kotler A, Wegerif R, and LeVoi M (2001) Oracy and the educational achievement of pupils with English as an additional language: The impact of bringing 'talking partners' into Bradford schools. *International Journal of Bilingual Education and Bilingualism* 4: 403–19.

Kuperman V, Stadthagen-Gonzalez H, and Brysbaert M (2012) Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods* 44: 978–90.

Landis J and Koch G (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159–74.

Lervåg A and Aukrust VG (2010) Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry* 51: 612–20.

Marulis LM and Neuman SB (2010) The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research* 80: 300–35.

McKeown MG (1993) Creating effective definitions for young word learners. *Reading Research Quarterly* 28: 17–31.

McKeown M, Beck I, Omanson R, and Perfetti C (1983) The effects of long-term vocabulary instruction on reading comprehension: a replication. *Journal of Reading Behavior* 15: 3–18.

Murphy V and Unthiah A (2015) *A systematic review of intervention research examining English language and literacy development in children with English as an additional language (EAL)*. Oxford: University of Oxford.

Nash H and Snowling M (2006) Teaching new words to children with poor existing vocabulary knowledge a controlled evaluation of the definition and context methods. *International Journal of Language and Communication Disorders* 41: 335–54.

Nation K, Clarke P, Marshall C, and Durand M (2004) Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research* 47: 199–211.

Nation P (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

OECD (2010) Executive summary. In: OECD (ed.) *Closing the gap for immigrant students*. Paris: OECD Publishing, pp. 7–9.

Oxley E and de Cat C (2019) A systematic review of language and literacy interventions in children and adolescents with English as an additional language (EAL). *The Language Learning Journal*. Epub ahead of print 11 April 2019. DOI: 10.1080/09571736.2019.1597146.

Renfrew C (1988) *Action picture test*. Milton Keynes: Speechmark Publishing.

Romaine S (1995) *Bilingualism*. Oxford: Blackwell.

Ross S and Begeny J (2011) Improving Latino, English language learners' reading fluency: The effects of small-group and one-on-one intervention. *Psychology in the Schools* 48: 604–18.

Rupley WH and Nichols WD (2005) Vocabulary instruction for the struggling reader. *Reading and Writing Quarterly* 21: 239–60.

Semel E, Wiig E, and Secord W (2006) *Clinical evaluation of language fundamentals*. 4th UK edition. London: Pearson Assessment.

Shadish W, Cook T, and Campbell D (2002) *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin.

Silverman R (2007) Vocabulary development of English-Language and English-Only learners in kindergarten. *The Elementary School Journal* 107: 365–83.

St John P and Vance M (2014) Evaluation of a principled approach to vocabulary learning in mainstream classes. *Child Language Teaching and Therapy* 30: 255–71.

Stahl SA and Fairbanks MM (1986) The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research* 56: 72–110.

Strand S, Malmberg L, and Hall J (2015) *English as an additional language (EAL) and educational achievement in England: An analysis of the National Pupil Database*. Oxford: University of Oxford.

Swanborn MSL and de Glopper K (1999) Incidental word learning while reading: A meta-analysis. *Review of Educational Research* 69: 261–85.

Verhoeven L and Van Leeuwe J (2008) Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology* 22: 407–23.

Wechsler D (2003) *Wechsler Intelligence Scale for Children*. 4th edition. San Antonio, TX: Harcourt Assessment.

Wilkinson KS and Houston-Price C (2013) Once upon a time, there was a pulchritudinous princess ...: The role of word definitions and multiple story contexts in children's learning of difficult vocabulary. *Applied Psycholinguistics* 34: 591–613.

Appendix 1. Taught and untaught vocabulary.

Travelling and distant lands	Emotions and states	Wrongdoing	Finance and shopping	Accident and emergency
Taught vocabulary (thematic):				
distant	furious	responsible	purchase	rescue
capital	miserable	disagree	bargain	disaster
coast	thrilled	persuade	wealthy	agony
navigate	cautious	fraud	afford	fatal
Untaught vocabulary (non-thematic):				
annual	budget	contagious	donation	genuine
identical	maximum	pilot	starve	tolerate

Appendix 2. Individual scores of intervention participants across all time points (word total score).

Child	Baseline	Pretest	Posttest	Maintenance	Pretest–posttest increase (percentage)
1	25.5	16	51	50	218.8
2	49	53	61	65	15.1
3	25	23	–	51.5	–
4	29	58	103	88	77.6
5	55	54.5	61.5	59	12.8
6	55	54	90	94	66.7
7	64	53	100.5	112.5	89.6
8	65	77	105	97	36.4
9	73.5	50.5	113.5	101	124.8
10	53	57.5	74.5	–	29.6
11	70	92	128	119	39.1
12	49	41	57	–	39.0