Article

A Phase 2 Exploratory Trial of a Vocabulary Intervention in High Poverty Elementary Education Settings

Maria Cockerill [1],*, Allen Thurston [1], Andy Taylor [2], Joanne O'Keeffe [1] and Tien-Hui Chiang [3]

[1] School of Social Sciences, Education and Social Work, Queen's University Belfast, UK;
E-Mails: maria.cockerill@qub.ac.uk (M.C.), a.thurston@qub.ac.uk (A.T.), j.okeeffe@qub.ac.uk (J.O.K.)
[2] Fischer Family Trust Literacy, UK; E-Mail: literacy@fft.org.uk
[3] School of Educational Science, Anhui Normal University, China; E-Mail: thchiang@ahnu.edu.cn

* Corresponding author

Abstract
This article reports results of a phase 2 exploratory trial of a vocabulary program delivered in elementary schools to improve student's reading ability, including their comprehension. The intervention was tested as a targeted intervention in classrooms with children aged 7–10 across 20 weeks during one school year, with eligible students learning in small groups of four. Teachers and support staff received training in this cooperative learning approach to develop children's vocabulary with particular focus on Tier-2 words. School staff received additional support and resources to equip them to develop and implement the vocabulary instruction sessions to targeted students. The trial was undertaken with a sample of 101 students in seven schools from three English district areas with high levels of socio-economic disadvantage. A standardized reading test was used to measure reading outcomes, with significant gains found in student's overall reading ability, including comprehension. Owing to the positive results found in this trial, including positive feedback about implementation of the technique, next steps should be a larger trial with 48 schools to avoid the risk of sampling error due to limited number of schools.

1. Background

The value of a developed vocabulary for students in lower elementary grades increases as they learn to read (Apthorp et al., 2012), and an extensive vocabulary predicts reading attainment (Biemiller, 2003). Research supports vocabulary instruction and identifies a very high correlation between comprehension and vocabulary intercepts, suggesting that reading comprehension and oral vocabulary knowledge can be understood as reflecting a single higher order language construct (Ricketts et al., 2020).

In 2019, only 51% of disadvantaged students in England achieved the expected reading standards at age 11, compared to 71% of all other students (Department for Education, 2019a). For some years, research has indicated that lower vocabulary expertise in children from high poverty areas adds to attainment failure (Becker, 1977), as they often have smaller vocabularies than their peers, and that this difference in outcomes grows as children age (Beals, 1997; Waldfogel & Washbrook, 2010). Evidence continues to link a significant proportion of children who do not reach their full academic potential with socioeconomic status (Gorard

et al, 2019; Pishghadam, 2011). Although the issue may lie in a lack of literacy experiences outside of school, a form of literacy poverty, schools need to seek effective means to address these inequalities.

There are challenges for any intervention which aims to enhance literacy for students who have not had rich experiences of literacy in their development. One challenge is how to bridge the gap between current linguistic and literacy abilities, and those skills that are needed to interrogate and make sense of text, particularly in a formal and academic setting. The nature of texts used in school often requires the meaning to be interpreted abstractly. This can be problematic if pupils have not had previous experiences in school or in the home where more complex vocabulary is acquired, and abstract conceptualisation is common. Bernstein (1990) saw the essential link between literacy and abstract conceptualisation of meaning as being mediated by familiarity with elaborate codes where children regularly have experiences of being asked to make these links.

Classroom observations from US research suggest that vocabulary teaching in elementary schools is not prevalent (Wanzek, 2014). Yet instruction about word learning has the capacity to be taught well (Duke & Moses, 2003) and have positive impact (Beck & McKeown, 2007), as demonstrated by positive effect sizes (ES+0.91) from a summary of 41 studies (Stahl & Fairbanks, 1986). More recently, research confirms the efficacy of vocabulary teaching to improve word knowledge proficiency (ES+0.29 to 1.21) and text understanding (ES+0.10 to 0.50; Elleman et al., 2009; Marulis & Newman, 2010). Comprehension ability is aided by children's vocabulary knowledge, and instruction benefits children to develop an extensive picture vocabulary enabling them to understand words (Verhoeven et al., 2011). A recent study of struggling readers as they leave elementary settings supports the potential of vocabulary learning with positive assessments of narrative and vocabulary of between +0.15 and +0.26 (Joffe et al., 2019). This builds on earlier studies that showed the benefits of oral vocabulary instruction for both improved vocabulary (+0.02–0.26) and listening comprehension (+0.15–046) from the earliest elementary stage (Fricke et al., 2017).

The study presented here measures the effects of children's reading attainment when exposed to a new vocabulary intervention designed for targeted use in English elementary schools, using a Medical Research Council (2000) phase 2 exploratory randomized controlled trial (RCT).

2. The Intervention

Students aged 7–10 in UK schools receive the intervention implemented by teachers and support staff who have been trained to deliver the program. This includes two days of training and support from delivery experts Fischer Family Trust Literacy (FFTL).

Staff training focuses on the knowledge, skills, and understanding required to implement the vocabulary intervention to the selected students who require it. Staff learn about vocabulary instruction research and about using a multi-strategy methodology in vocabulary instruction. This contains a teaching sequence, varied vocabulary activities, teaching elements including identifying words to teach and using Tier-1 words as an introduction to more complex Tier-2 words (Beck et al., 2013), example session plans, and formats to plan and record sessions. Staff learn about the teaching sequence of revise, teach, practise, and apply, for vocabulary instruction in elementary classrooms. Marzano's six-step cycle for vocabulary instruction developed for use in high schools includes: explain, restate, show, refine, discuss, and play (Marzano & Pickering, 2005). These elements are redistributed in the sequence when adapted for use in elementary settings.

The intervention addresses children's knowledge of words/vocabulary and their ability to use them appropriately, rather than the specific syntax and grammar of standard English which often focuses on sentence construction. Nevertheless, the focus on vocabulary does raise issues which are pertinent to the issue of standard English and grammar. During staff training, therefore, it is emphasized that no vocabulary intervention is likely to work if it involves denigrating, however subtly, the speaker's natural language, their home register. So non-standard forms and word choices are seen as a resource and a starting point for learning new vocabulary. Staff are trained to celebrate the richness and diversity of language in all its forms, whether Tier-1 English words or from another language, whilst expanding the range of registers and the associated vocabulary that will give children access to more conversations and more complex dialogue. In this sense, a non-standard dialect, with its characteristic vocabulary, is an asset and is considered another form of Tier-1 vocabulary, just as the home language of a standard English speaker is likely also to be dominated by Tier-1 forms and choices. The intervention focuses on giving children opportunities to learn a wider range of vocabulary and use more Tier-2 vocabulary, appropriate for more formal and more academic dialogue. Staff training explicitly identifies Tier-1 vocabulary, which would include non-standard forms and dialect words, as a bridge to the more formal register which is opened up by Tier-2 vocabulary choices.

Activities reinforce children's home language as rich resource. These include drama activities and role play to explore how messages need to be conveyed differently according to the formality of the situation, the audience and players, and the purpose, etc. Role play and the word choices are then examined through discussion. Similarly, an activity like "word cline" might include dialectal vocabulary choices which are ordered according to criteria (e.g., best words from a selection to describe something to friends, as compared to an unknown audience; the same words might be discussed but the order

would differ according to the audience and purpose of the conversation). Other activities explore words by linking them to their related words or, as in the "SAME thing" activity, examine synonyms and antonyms.

Other aspects of the vocabulary intervention which are relevant to grammar teaching and non-standard forms are that (1) emphasis is given on idiomatic phrases, which are rarely very formal but often used in formal situations, like news bulletins to give colour and make more vivid a more abstract point—non-standard expressions would potentially be examined as part of this, as the words chosen for consideration can be sourced from newspapers, media and books as well as topics of interest—and that (2) the program highlights the importance of learning about "word families," often moving from a single word to explore how it changes, usually according to its function in the sentence, into a range of other words (e.g., "ladder tracker" activity).

As well as making explicit (but natural) use of grammatical terminology—noun, adjective, adverb, etc.—as a kind of meta-language to be able to talk about words and their function, these activities highlight the importance of morphemes, the smallest units of meaning in words, which transform a noun into an adjective, etc. A prefix and suffix chart is also included in the program resources to facilitate conversations about language and how words derive from each other.

Students are selected by their teachers if they have limited or poor vocabulary, and if their comprehension of texts is less advanced than their decoding skills. The intervention is delivered in small groups (approximately four). Students receive clear induction into the focus words, using familiar Tier-1 words as a bridge to learning less familiar Tier-2 words, using focus words in instances which illuminate meaning. Students repeat pronunciation and meaning, identifying words with comparable meaning, and search for associated words. The group then often undertakes a cooperative learning activity to use and explore the word (e.g., a game, drama, or listening and speaking activity). Students then practice their knowledge application. For example, they may invent phrases using new words in context, and will use the words again at the beginning of the subsequent lesson. The words are consolidated by students and displayed in the classroom to help students use them during subsequent instruction time. Over two terms (approximately 20 weeks with a minimum of 12 weeks), staff in English elementary classrooms deliver vocabulary instruction in small groups, for 15–20 minutes, three times per week (minimum), and daily if possible (optimal). During each week, five new words with a related meaning are introduced.

The FFTL staff training to deliver the intervention includes:

- Tools to diagnose children with less developed word-knowledge.
- The lesson sequence format.

- Training on Tier-2 words as the focus (Tier-1 words used as bridge).
- Guide templates for initial teaching and introducing new words.
- Guidance on teaching approaches which enable students to transit from learning new vocabulary to using it with confidence.
- A resource box of student support materials.

Table 1 summarizes the vocabulary program. Students allocated to the treatment arm of the trial receive the vocabulary program instruction, while students allocated to the control arm do not receive it and continue to receive their usual literacy instruction.

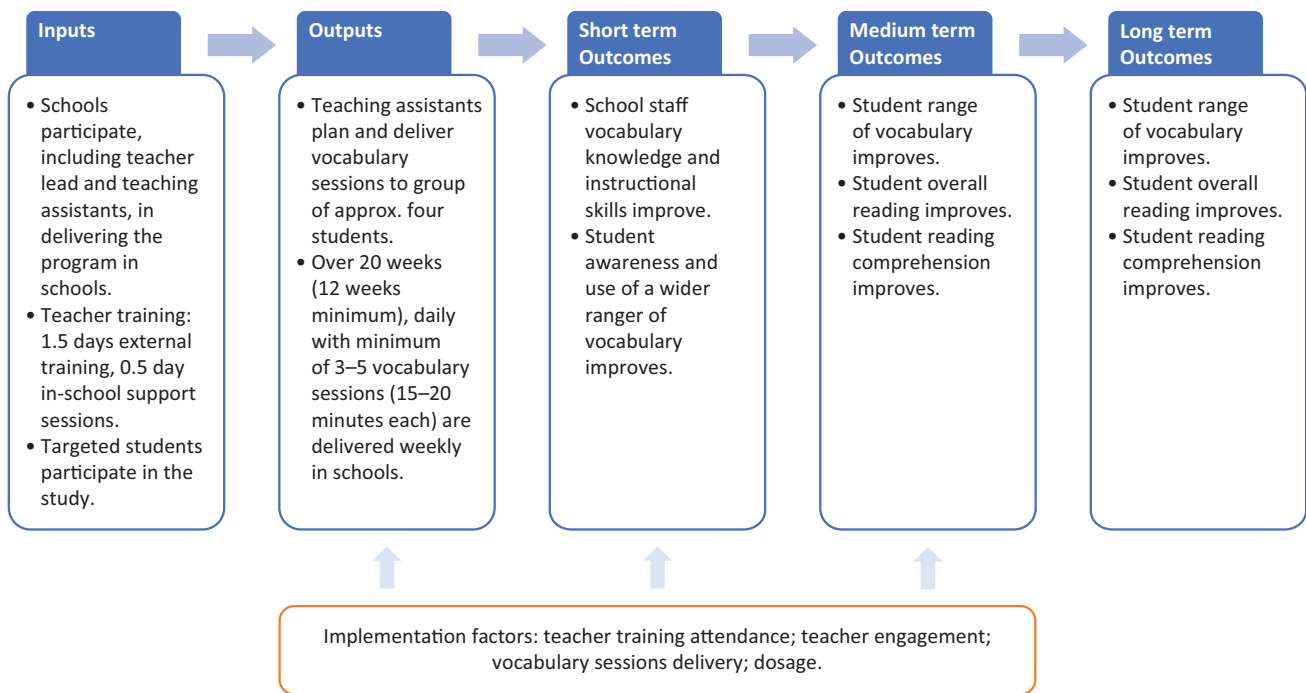## 3. Theory of Change of the Vocabulary Program

Inputs, outputs, and outcomes of the program are described in the logic model (Figure 1). This includes the theory of change, and how delivery factors link up to program outcomes.

3.1. Theory of Intervention and Change

The theory of intervention and change of this program is illustrated in the logic model (Figure 1), which aims to expand student's vocabulary to improve their understanding of text to enhance reading ability.

3.1.1. Theory of Intervention

There is evidence that vocabulary is a predictor of literacy outcomes (Moody et al., 2018; Moore et al, 2014), that vocabulary instruction can enhance word knowledge and reading abilities in children (Beck & McKeown, 2007; Moody et al., 2018), and that vocabulary and background knowledge can contribute to comprehension ability (Apthorp et al., 2012; Cromley & Azevedo, 2007; Moody et al., 2018). Furthermore, researchers recommend teaching flexible word-learning strategies and techniques for self-monitoring to improve reading comprehension and to facilitate word knowledge across a variety of contexts (Wright & Cervetti, 2017). In line with evidence that quality professional training can change teacher practice (Coe et al., 2014), teachers in the present study are trained in the necessary vocabulary strategies and pedagogies to deliver instruction in elementary classrooms. Staff training in the vocabulary program equips teachers with the required knowledge to change the practice of vocabulary teaching, including spaced regular delivery, modelling learning for children, and nurturing cooperative learning structures during sessions whilst providing appropriate scaffolding. Program implementation and student selection are important elements during staff training. Staff from various schools are trained together to implement the program and receive bespoke visits per school to support classroom delivery.

Inputs		Outputs		Short term Outcomes		Medium term Outcomes		Long term Outcomes
• Schools participate, including teacher lead and teaching assistants, in delivering the program in schools. • Teacher training: 1.5 days external training, 0.5 day in-school support sessions. • Targeted students participate in the study.		• Teaching assistants plan and deliver vocabulary sessions to group of approx. four students. • Over 20 weeks (12 weeks minimum), daily with minimum of 3–5 vocabulary sessions (15–20 minutes each) are delivered weekly in schools.		• School staff vocabulary knowledge and instructional skills improve. • Student awareness and use of a wider ranger of vocabulary improves.		• Student range of vocabulary improves. • Student overall reading improves. • Student reading comprehension improves.		• Student range of vocabulary improves. • Student overall reading improves. • Student reading comprehension improves.

Implementation factors: teacher training attendance; teacher engagement; vocabulary sessions delivery; dosage.

Figure 1. Vocabulary program logic model.

The vocabulary program instruction thus includes a multi-varied flexible approach focusing on a teaching sequence, a variety of activities, and teaching elements, including knowing which words to teach, with focus on Tier-2 words (using Tier-1 words as a bridge; see Beck et al., 2013). The focus on Tier-2 words is supported by research as an effective instruction approach (Coyne et al., 2018). Training includes example lesson templates, and teachers learn and are supported in planning and recording template usage.

Research supports sustained vocabulary instruction to make a substantial impact on children's lexical development (Beck & McKeown, 2007; Biemiller, 2003). There is agreement about the benefits of vocabulary instruction to nurture preliterate phonological abilities in young children (Walley et al., 2003). The rationale in respect of the lexical hypothesis is that when younger children learn new vocabulary their depiction of words is refined, leading to phonological awareness and reading fluency (Metsala & Walley, 1998) which becomes embedded through extended literacy opportunities (Walley et al., 2003). In respect of older elementary children, the lexical quality hypothesis suggests exposure to words helps students develop high quality word representations leading to reading proficiency which improves through using new words repeatedly (Perfetti, 2007; Perfetti & Hart, 2002). To improve children's reading and comprehension it is recommended for children to encounter the new words in an active manner (Apthorp et al., 2012). This includes repeatedly providing students with the target word context and meaning (Beck et al., 2002). This research is supported by recent RCT studies with early adolescent struggling readers who found positive effects of +0.15–0.26

(Joffe et al., 2019) and early elementary readers who found positive effects of +0.02–0.26 for improved vocabulary and +0.15–0.46 for listening comprehension (Fricke et al., 2017).

It is important to distribute and scaffold learning by using small groups and cooperative learning structures to practice the cycle of learning new words and embed these in children's usage to self-monitor their understanding. This kind of meta-cognitive development, which is evident from about 5–6 years of age and increases from age eight (Veenman, 2016), can have a positive impact on learning outcomes of ES+0.7. It has the greatest success when learning is scaffolded by teachers and includes cooperative group structures (Higgins et al., 2014). The psychology literature supports the effectiveness of learning using systematic spacing in which learning sessions are implemented with time intervals between them, rather than being delivered in one block of learning time.

Since Ebbinghaus (1885/1964) in the late 19th century, the field of cognitive psychology has found benefits in spacing learning. More recently, meta-analyses estimate that about 75% of 400+ verbal learning studies in the distributed practice literature show a spacing advantage (Cepeda et al., 2006). Verbal learning in elementary school, where spaced or distributed learning is used, appears to be most beneficial for learning simple word recall (Seabrook et al., 2005) and word and fact learning (Sobel et al., 2011), including vocabulary and text comprehension. A recent study identified an effect size of +0.85 for the benefit of spaced learning with verbal contents (Wiseheart et al., 2019). Thus, the vocabulary program is implemented using

a structure of 3–5 weekly sessions of 15–20 minutes each. Scaffolding, identified as an effective approach (van de Pol et al., 2010), is integral to the teacher's modelling process during instruction, and is grounded on the theory developed in the early 20th century by Vygotsky (1978) about learning within the "zone of proximal development." Scaffolding support used in literacy learning continues to be recommended by evidence-driven organizations in England (Education Endowment Foundation [EEF], 2021a). Vocabulary instruction in the present study therefore includes mediation and teacher facilitated support, as children work collectively in small groups of approximately four.

Students learn cooperatively sharing a common goal, including active participation in group activities facilitated by teachers. Research supports cooperative learning structure when activities are correctly aligned to the student's capabilities. Such interaction during vocabulary sessions enables co-construction of word meanings whilst students learn together. Social interdependence theory (Johnson & Johnson, 2012; Johnson et al., 2010) underpins student's cooperative learning and has been tested for its effectiveness over many years, where meta-analyses identify the benefits of implementing this approach in schools (ES +0.19 to +0.91; Johnson et al., 2000). This is supported by more recent meta-analyses with findings of ES+0.44 (Igel, 2010) and ES+0.59 (Capar & Tarim, 2015) for the use of cooperative learning. Evidence is supported by the EEF synthesis of meta-analysis findings of between +0.09 and +0.91 (EEF, 2018), and the EEF Teaching and Learning Toolkit which identifies an average of five months additional progress for students who engage in cooperative learning (EEF, 2021b). Positive evidence of cooperative learning for literacy instruction in elementary schools with effect size of ES+0.20 is found in a meta-analysis including 18 studies in elementary schools (Puzio & Colby, 2013). Most recently, literacy trials using cooperative learning approaches to improve Reading outcomes also support this approach with positive effect sizes of ES+0.13–0.25 (Thurston et al., 2020) and ES+0.24 (Thurston et al., 2021).

The vocabulary program designed with its multi-strategy approach is hypothesized to enhance vocabulary skills and thus improve student's reading ability and comprehension outcomes and is tested in this study using a digital reading measure which is standardized.

3.1.2. Theory of Change

A structured vocabulary program supported by staff training can change the teaching of vocabulary, if training impacts on teacher's knowledge and actions, and their pedagogies change. This in turn impacts on student vocabulary development resulting in improved reading.

Vocabulary instruction dosage, staff questionnaires, and staff training attendance are used to understand the implementation factors and mediators for outcome change.

3.2. Criteria to Recommend the Vocabulary Program for a Phase 3 Definitive Randomized Controlled Trial

To determine the vocabulary program's readiness for a definitive RCT (phase 3) in elementary schools, the criteria outlined below were used:

- Staff training is deliverable as specified.
- The program is deliverable to students as specified.
- Consideration for scale-up will include whether staff evaluate the program positively enough.
- A positive effect size is reported in students who received the vocabulary program, when compared to control group students who did not receive it.

4. Research Questions

Decisions about whether the vocabulary program should be scaled to a phase 3 trial will be informed by addressing the questions below:

1. Can the vocabulary program be implemented in elementary schools?
2. Do students' reading ability improve when they receive the vocabulary program?
3. Does the impact of the vocabulary program differ significantly depending on variations in implementation fidelity (process evaluation)?
4. Should the vocabulary program be scaled?

5. Randomized Controlled Trial and Process Evaluation Design Summary

A logic model was produced (Figure 1) to help guide the process evaluation and enable interpretation of RCT findings for the vocabulary program intervention. To structure the trial, SPIRIT (2015) guidelines were consulted.

In the RCT evaluation, primary outcomes were investigated using ANCOVA. The RCT examined differences in pre- to post-test reading scores on the New Group Reading Test (NGRT). The ANCOVA assessed post-test reading scores for the intervention group who received the program and compared them to control group scores, using pre-test reading scores in the ANCOVA. Results were calculated for each of the main outcome measures and are presented as *effect sizes* (calculated as Cohen's *d*).

The process evaluation was guided by the Medical Research Council guidelines on the development of RCTs for complex health interventions (Moore et al., 2015) and supplements the RCT to assess whether the program was implemented with fidelity. Vocabulary program staff training attendance data (naturally occurring data from the trainer), teacher engagement data, and dosage of program delivery data were collected. A post-program questionnaire was completed by staff, including questions about control group learning during the program.

6. Methods

Pre- and post-intervention tests were completed by all participating students (intervention and control groups). Schools were given guidance to select up-to a maximum of 16 students from years 3 and 5 combined (aged 7–8 or 9–10), with good decoding skills and poor comprehension and vocabulary.

6.1. Reading Outcome Measure

A pre- and post-intervention standardized reading test was undertaken under exam conditions with the selected children. The NGRT (digital version) from GL Assessment was used. This adaptive test with high reliability (Alpha values 0.9; GL Assessment, 2018) assesses children's reading ability, using sentence completion and reading comprehension scales. The pre- and post-test data from the intervention and control groups obtained were used to determine the differential effects of the program on student reading attainment level. The sentence completion and passage comprehension sub-scales in the NGRT combine to provide an overall reading score. Analysis was undertaken on both sub-scales and the combined overall reading score.

6.2. Dosage Record

To assess fidelity of implementation teachers completed a 20-week implementation plan collected at post-test, recording weekly vocabulary instruction information.

6.3. Teacher Questionnaire at Post-Test

School staff involved completed the online questionnaire (Lime Survey) at post-test. The questionnaire included 19 questions: 11 questions with a 4-point scale with the poles only marked from *strongly agree* to *strongly disagree*, five open answer questions and three closed answer questions, each with forced choice options about instruction implementation.

6.4. Training Delivery Naturally Occurring Data

Program trainers collected staff attendance records. Table 1 summarizes the instruments and measures.

6.5. Sample

North East England has the second highest low income and deprivation rate after inner London (Bradshaw, 2020). The funder identified this area for research owing to high levels of disadvantage here and three high poverty district areas within this region were selected. Elementary schools with above 15.8% national average levels of students eligible for Free School Meals (Department for Education, 2019b) from the selected three high poverty district areas in this region were invited to participate in this study. Given the funding restrictions, seven schools participated in the trial with completed Memorandum of Understanding agreement. This group of schools had an average school level of eligible FSM students of 49%, compared to the national average figure of 15.8%.

In total, the study included 104 children from seven schools in North East England from the three high-poverty area districts where recruitment to the trial took place. Students aged 7–10 from two year groups were recruited for the trial, with up to eight students from years 3 and 5 respectively (up to four participated as intervention and four as control from each year-group), giving a total of up to 16 from each school. The demographic nature of the sample was composed as follows: 56 female and 48 males; 38.5% (40 of 104) were entitled to FSM; eight out of 104 reported they were special educational need students under a health and care plan; 25 out of 104 had English as second language; ethnic balance was 77 white British, three white other, 14 Pakistani, two Other Asian, one Afghani, one African, one Chinese, one Indian, one Bangladeshi, three not specified. Schools who agreed to participate were provided with guidance to select eligible students, and as instructed, selected the students for this trial who were

Table 1. Measurement tools.

Outcome	Instrument	Completed by	Alpha values
Reading comprehension	NGRT: Passage comprehension subtest	Pupil	>0.9 (GL Assessment, 2018)
Overall reading	NGRT	Pupil	>0.9 (GL Assessment, 2018)
Reading accuracy	NGRT: Sentence completion subtest	Pupil	>0.9 (GL Assessment, 2018)
Implementation factors			
Dosage	20-week implementation (60–100 minutes weekly)	Teacher	n/a
Teacher engagement	Training attendance	Trainer	n/a
Teacher engagement	Teacher online survey	Teacher	n/a

Table 2. Sample.

	Condition		
	Control	Treatment	Total
Year 3	27	24	51
Year 5	27	26	53
Total	54	50	104

good decoders but had poor comprehension with poor vocabulary (see Table 2).

6.6. Randomization

Randomization to condition was undertaken at the individual level; each student was listed alphabetically within each year group and by school. A whole number was generated between 0 (control) and 1 (vocabulary intervention) using a software program for the generation of random numbers (Random Number Generator for iPhone, version 5.0 by Nicolas Dean). Even numbers of intervention and control participants were ensured in each arm of the trial by assigning the first student to condition in year 3 followed sequentially by the other seven students, and then in year 5 respectively, per school. This was true randomization and no minimization was used, as students were assumed to be evenly distributed in control and intervention groups.

6.7. Sample Size Calculation and Analysis

The sample size for the study was calculated to detect an effect size of +0.33, with ANCOVA, with $p > 0.05$ and 80% power (Soper, 2019). In calculating this sample size, it was assumed that:

1. There was even distribution of sample between control and intervention groups.
2. Loss of sample due to attrition would be <5%. Missing data was to be treated as missing completely at random if levels remained under 5%.

We consider an effect size of +0.33 to be reasonable based on evidence from previous reported effect sizes from experimental studies using vocabulary instruction where the average reported effect sizes range from +0.29 to +1.21 for the impact on vocabulary and +0.10 to +0.50 for the impact on reading comprehension (Elleman et al., 2009; Marulis & Newman, 2010).

6.8. Ethics

Two ethics procedures were undertaken to approve the trial. The participating school Headteachers approved the testing and intervention of the trial. The matching of the process of data, and the merger and analysis of the secondary data set was approved by the Ethics Review Board at the School of Social Sciences, Education and Social Work, Queen's University Belfast.

Before the trial a protocol for the work was developed and accepted for publication (Cockerill et al., 2019).

7. Results

The vocabulary intervention effects on sentence completion, passage comprehension, and overall reading score.

In total, 104 children were randomized to intervention or control group. Missing data were below 5% of the sample. Three students were missing at either pre- or post-test (who left the school). Therefore, the data were assumed to be missing completely at random (due to the N of the sample, it was not possible to explore this statistically). In total, 49 students were left in the intervention and 52 in the control group. The demographic nature of the sample which completed the NGRT at post-test was composed as follows: 56 female and 46 males; 38.24% (39 of 102) were entitled to FSM; eight out of 102 reported they were special educational need students under a health and care plan; 25 out of 102 had English as second language; ethnic balance was 75 white British, three white other, 14 Pakistani, two Other Asian, one Afghani, one African, one Chinese, one Indian, one Bangladeshi, three not specified. Note that in the following analyses, degrees of freedom change as two more students (one from each arm of the trial) did not receive a score on one of the sub-scales at either pre- or post-test. This is because the adaptive test switches to a phonics-based assessment to calculate overall reading scores if the performance of students indicates that a floor effect may occur as they are not able to respond correctly to the easiest of questions as the main test progresses. Therefore, the slight variation that occurred in the numbers of scores are available for analysis in sub-scales. We excluded the two students who did not receive sub-scale scores from the final analysis.

The vocabulary intervention group improved reading significantly compared to the control group. Descriptive statistics for pre- and post-test results in the NGRT reading tests are presented in Table 3. Pre-test reading scores, although slightly higher for the control group, did not differ significantly ($F(1, 102) = 0.26$, $P = 0.871$ not significant). However, due to the slight elevated mean score difference in favour of the intervention, and to enhance power in the analysis, we used ANCOVA, rather than ANOVA, in the analysis of data.

Improvements were observed for the vocabulary group on all NGRT reading sub-scales and the overall reading scale. Overall, the NGRT reading score showed positive effects for the vocabulary intervention over the control. ANCOVA analysis of NGRT scores, indicated that there was a significant gain on overall reading scale ($F_{(1,98)} = 6.11$, $p < 0.05$), sentence completion scale ($F_{(1,98)} = 5.05$, $p < 0.05$), and the passage comprehension scale ($F_{(1, 98)} = 5.05$, $p < 0.05$), with observed power at 68.7%. There was some evidence of clustering effects on mean NGRT overall reading scores within the sample at pre-test at the school level ($F_{(6, 102)} = 6.97$, $p < 0.001$).

7.1. Process Evaluation

Survey responses were received from each school involved in the study, with 11 staff responses received in total (four teachers, six teaching assistants, and one senior manager).

10 respondents indicated that they *strongly agreed* that they had followed the program as instructed in the professional development sessions, and one *agreed*. Eleven respondents *strongly agreed* (6) or *agreed* (5) that the program was easy to follow. Eleven respondents *strongly agreed* (6) or *agreed* (5) that they would keep following the program after the project had finished. All respondents indicated they would recommend it to schools.

Eleven respondents *strongly agreed* (9) or *agreed* (2) that they felt engaged when using the program. Eleven respondents *strongly agreed* (7) or *agreed* (4), that the school received the required support. Comments from respondents indicated that students engaged with the program. Respondent 3 noted: "Children engagement in identifying new words. Children have completed extra work at home linked to the work they have completed as part of the program." Respondent 7 indicated that "the children were eager to engage in lots of conversations about new words and looked forward to hearing the 'new

word' of the day." Respondent 10 reported: "Providing children with the chance to use vocabulary unfamiliar with them but in an informal environment, where the children could have fun and were unthreatened."

Average length of sessions was 18.64 minute (*SD* 2.34 minutes). Average number of weekly sessions was 4.09 (*SD* 0.70). All schools delivered the program as per the specification (15–20 minutes per week, 3–5 session per week). No schools were out of these ranges.

7.2. Cost

The program was implemented with a sample of 101 students in seven schools (including a wait-treatment control group) and cost £1,806 per school. This is equivalent to £125 per child, which is a low cost using the EEF Teaching and Learning Toolkit (EEF, 2021c). These costs included staff training but excluded costs which would not usually be undertaken when engaging with the program, such as teacher cover, evaluation, and testing.

The cost of program implementation included resources and staff time for program delivery. Implementation costs were estimated per student over one-year and included staff training and support (two days), manuals, and resources.

8. Discussion

This RCT of vocabulary instruction in elementary schools had not been tested before and was conducted successfully. The intervention demonstrated that reading improved for students who received the intervention when compared to the control students. Although differences in the sample at pre-test were detected between the control and treatment groups these were not significant. Owing to the slight elevated pre-test mean score difference in favour of the intervention, and to enhance power in the analysis, ANCOVA, rather than ANOVA, was used in the analysis of data. Improvements were

Table 3. Pre- and post-test results in NGRT reading tests.

	Mean NGRT Sentence Completion		Mean NGRT Passage Comprehension		Mean NGRT Overall Reading Score	
	Pre-test score (SD)	Post-test score (SD)	Pre-test score (SD)	Post-test score (SD)	Pre-test (SD)	Post-test (SD)
Control (n = 52, N = 7)	254.79 (54.32)	275.40 (54.56)	255.36 (60.10)	274.69 (54.86)	250.79 (56.68)	269.96 (53.48)
Vocabulary intervention (n = 49, N = 7)	258.64 (49.84)	298.61 (59.50)	246.24 (69.23)	289.80 (72.98)	252.56 (53.47)	292.73 (61.44)
Mean change in raw scores vocabulary intervention vs. control	19.81		24.34		21.36	
Mean change as effect size vocabulary intervention vs. control	+0.36		+0.42		+0.38	

Note: n (number of students) = 101 and N (number of elementary schools) = 7.

observed for the vocabulary group on all NGRT reading sub-scales and the overall reading scale. Overall, the NGRT reading score showed significant positive effects for the vocabulary intervention over the control group, on the overall reading scale (effect size +0.38), sentence completion scale (effect size +0.36), and passage comprehension scale (effect size +0.42). The connection between vocabulary development and improved reading is well established (Blachowicz, 1985; Ricketts et al., 2020). Having effective pedagogies tried and tested for teachers that allow effective vocabulary development is essential (Coyne et al., 2018; Hunt & Beglar, 1998).

All schools implemented the vocabulary intervention within their existing timetable and met target delivery specifications. The staff survey suggested that teachers were able to embed the vocabulary program instruction specifications effectively into their vocabulary teaching. The positive staff responses evidenced that they implemented the key elements of the vocabulary program during instruction. Respondents reported that they both implement the vocabulary program according to specification and would be willing to continue using it and to recommend it. This leads us to conclude that the vocabulary program is suitable for elementary school implementation and can be embedded in schools. Whilst power was at 68.7%, the effect size was large enough to allow significant changes to be detected.

It was demonstrated that the technique can be used in elementary schools. All staff respondents were unanimously positive about the vocabulary program's suitability to their school setting and reported that implementation was straightforward, including negligible effects on workload. Though a sample of only seven schools, those surveyed included a varied spectrum: senior managers, teachers, and teaching assistants. Therefore, beyond the current sample the program demonstrates promise for scalability. In terms of limitations, then it must be acknowledged that despite the randomized design and use of independent measures, there may be a risk of sampling error due to the limited number of schools and students in this efficacy trial. In addition, what remains unknown is potential clustering effects that may occur if implemented at scale. The way to explore these issues is to plan a larger trial capable of taking account of clustering effects. If this were to happen, then to detect an effect size of +0.36, assuming eight students per school, then 48 schools would be required, with alpha at 0.05 at 80% power. This may be necessary as there were clustering effects at the school level at pre-test, despite randomization at the individual level.

The vocabulary program technique had higher effects than simpler reading techniques using cooperative learning, such as peer tutoring in elementary school that detected effect sizes of +0.24 in a cluster randomized trial in 129 elementary settings in Scotland (Tymms et al., 2011). The trial can now be recommended to move to a phase 3 definitive trial, including larger sample. This would give higher power in analysis and allow any poten-

tial clustering effects to be modelled in the absence of any risk of "leakage" or "contamination" to the control group, that is a risk with within school control students. This form of group-work or cooperative learning to improve vocabulary shows promise. This approach, as seen with other forms of cooperative learning, provides a transformative pedagogy with weak framing (Bernstein, 1973), enabling children to develop their key capabilities necessary to flourish in society (Nussbaum, 2000; Sen, 1992).

Promoting the ability to read is important and literacy development is a social entitlement, a key determinant of well-being and a goal of human development (Sen, 1999). The ability to learn and develop (including literacy development) are moulded by the transition periods of student's lives from one stage of competence to another. Students exposed to situations where they can develop a competence and are given freedom to exercise it in cooperation with others, can improve their functioning and form more complex competence sets (Vygotsky, 1978, pp. 85–86). Interactive and participatory approaches to teaching are recommended in early studies supporting the social aspect of learning where students develop their understanding in conjunction and dialogue with peers or adults (Vygotsky, 1978). Children's deepening skills of communication and thinking cooperatively are considered key for their development as full social actors (Biggeri & Karkara, 2014).

Children's participation in economic and social life as adults requires the ability to read with understanding. It has been well documented that linguistic abilities regulate cognitive development (Vygotsky, 1978; Wells, 1985; Wood, 1998). In his analysis of code theory, Bernstein (1982) reviewed the range and type of language structures used by families and within schools. In Bernstein's classification, of codes, an elaborate code requires an environment with rich language and discourse where children can articulate themselves effectively. This theory provides a convincing insight into the linkage between cognitive development and linguistic abilities. Specifically, living in a particular environment requires complying with the requirements of the "code" of the context, as this enables communication between members of that social context/family unit to be understandable and workable (Bernstein, 1990; Bolander & Watts, 2009). Therefore, the close connection between a linguistic code and the environment in which it occurs is evidenced by the differences in linguistic expression between those who grow up in environments with differing approaches to linguistic and literacy development (Bernstein, 1982; Bernstein & Henderson, 1973). Those living in linguistically/literacy stimulated environments develop and extend their language to a greater extent than those who do not (Bernstein, 1990, 1996). This language-rich culture facilitates children's development of an elaborated code that helps them undertake logical reasoning and decode the theoretical, and often abstract, concepts embedded within texts to which they are exposed in school.

A language-rich culture is particularly important when learning more complex vocabulary and comprehension, moving from the words and text to abstract conceptualisation of meaning. Children with poorer vocabulary and comprehension skills may well lack the development of this required "scholastic" code. In this manner, texts can be conceptualized as vertical discourse, and characterised as theoretical, systematic and logical (Bernstein, 1999). By contrast, those growing up in linguistic or literacy poverty where their language may not be stimulated to the same extent as children from more literate backgrounds, may experience fewer opportunities to expand their ability to express abstract terms in verbal discourse, developing the required lexicon and comprehension to fully articulate with school-based texts (Bernstein, 1999). This environment results in the development of a restricted code, including features such as unorganized and unsystematic discourse. Because a restricted code differs greatly from written language, it can then be difficult for students with less well developed language skills to understand the abstract meanings of texts (Wells, 1987). To improve their outcomes when learning, weak social relations such as that which occurs during groupwork/cooperative learning (weak framing) can help to transform vertical discourse (strong classification) into understandable information (Bernstein, 1996; Southgate & Aggleton, 2017). Findings from this study suggest that groupwork/cooperative learning, along with academic learning comprising language rich experiences, are two core elements in developing the linguistic structures and competency for successful comprehension. The language exposure in the processes of vocabulary instruction, by which complex words and discourse about text in a suitably structured environment are shared, provides a good context and medium in which to develop linguistic and literacy skills for children who have been unable to develop these skills previously.

The vocabulary program in this study provides such exposure, including through exposure to Tier-2 words. Providing students with opportunities for vocabulary development can improve outcomes for struggling readers (Coyne et al., 2018; Fricke et al., 2017). Tier-2 words are those which appear infrequently enough that they will probably not be learned incidentally by students and may benefit from direct instruction to promote the use of processes (Beck et al., 2002). Tier-2 teaching and learning benefits from being combined with collaborative work where students have extended opportunities to interact with the vocabulary being learnt together with others (Stahl & Kapinus, 2001). It is this rich process between access to instruction in Tier-2 words and the linkages made with Tier-1 words which are already known, and structure to consolidate learning cooperatively, which provides the opportunity to build the human capabilities for literacy development to which all children and adults must have access to fully participate in society (Nussbaum, 2000; Sen, 1992).

Importantly, the process of vocabulary expansion promoted through the vocabulary program technique used in this study identifies at-risk students to receive a targeted intervention focused on learning Tier-2 words, which are then consolidated using a collaborative process. Building on previous literature, the technique used in this study resulted in improved reading, including decoding and reading comprehension for those children who engaged in the program. Such an approach to vocabulary development could be a significant contributor to reading development and capability development which require literacy skills, for children from socially disadvantaged areas.

9. Conclusion

The study established that the vocabulary program is implementable across a variety of elementary schools from three district areas and was well received by all staff. Using the technique resulted in improved reading ability overall for the intervention group, measured through an independent, standardized reading test not aligned to curriculum materials. The overall reading score, including for comprehension, showed positive effect sizes for the students who received the vocabulary program, compared to those who did not. We therefore cautiously recommend that this program be used by schools, whilst carefully observing effects in their own elementary settings. The criteria to recommend the vocabulary program be scaled were met and the suggested next steps are a larger trial with 48 schools to avoid the risk of sampling error due to limited number of schools, and to take account of clustering effects.

Conflict of Interests

The authors declare no conflict of interests.

References

Apthorp, H., Randel, B., Cherasaro, T., Clark, T., McKeown, M., & Beck, I. (2012). Effects of a supplemental vocabulary programme on word knowledge and passage comprehension. *Journal of Research on Educational Effectiveness*, *5*(2), 160–188. https://doi.org/10.1080/19345747.2012.660240

Beals, D. (1997). Sources of support for learning words in conversation: Evidence from mealtimes. *Child Language*, *24*, 673–694.

Beck, I., & McKeown, M. (2007). Increasing your low income children's oral vocabulary repertoires

through rich and focused instruction. *The Elementary School Journal*, *107*(3), 251–271.

Beck, I., McKeown, M., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. Guilford.

Beck, I., McKeown, M., & Kucan, L. (2013). *Bringing words to life (second edition): Robust vocabulary instruction*. Guilford Press.

Becker, W. C. (1977). Teaching reading and language to the disadvantaged—What we have learned from field research. *Harvard Educational Review*, *47*, 518–543.

Bernstein, B. (1973). On the classification and framing of educational knowledge. In R. Brown (Ed.), *Knowledge, education and cultural change*. Tavistock.

Bernstein, B. (1982). Code, modalities and the process of cultural reproduction: A model. In M. Apple (Ed.), *Cultural and economic reproduction in education: Essays on class, ideology and the state* (pp. 304–355). RKP.

Bernstein, B. (1990). *The structuring of pedagogic discourse: Class, codes and control* (Vol. 4). Routledge.

Bernstein, B. (1996). *Pedagogy, symbolic control and identity*. Taylor & Francis.

Bernstein, B. (1999). Vertical and horizontal discourse: An essay. *British Journal of Sociology of Education*, *20*(2), 157–173.

Bernstein, B., & Henderson, D. (1973). Social class differences in the relevance of language to socialization. In B. Bernstein (Ed.), *Class, codes and control* (Vol. 2). Routledge & Kegan Paul.

Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology*, *24*, 323–335.

Blachowicz, C. L. (1985). Vocabulary development and reading: From research to instruction. *The Reading Teacher*, *38*(9), 876–881.

Bolander, B., & Watts, R. J. (2009). Re-reading and rehabilitating Basil Bernstein. *Multilingua*, *28*, 143–173.

Bradshaw, J. (2020). *Child poverty in the North East region*. Social Policy Research Unit, University of York. https://cpag.org.uk/sites/default/files/files/policy post/Child%20poverty%20in%20the%20NE.pdf

Biggeri M., & Karkara R. (2014). Transforming children's rights into real freedom: A dialogue between children's rights and the capability approach from a life cycle perspective. In D. Stoecklin & J. M. Bonvin (Eds.), *Children's rights and the capability approach. Children's well-being: Indicators and research* (Vol 8, pp. 19–41). Springer. https://doi.org/10.1007/978-94-017-9091-8_2

Capar, G., & Tarim, K. (2015). Efficacy of the cooperative learning method on mathematics achievement and attitude: A meta-analysis research. *Educational Sciences: Theory and Practice*, *15*(2), 553–559. https://doi.org/10.12738/estp.2015.2.2098

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380. https://doi.org/10.1037/0033--2909.132.3.354

Cockerill, M., Thurston, A., & Taylor, A. (2019). Protocol: An efficacy randomized controlled trial of reciprocal reading in high schools. *International Journal of Educational Research*, *97*, 99–106.

Coe, R., Aloisi, C., Higgins, S., & Elliot Major, L. (2014). What makes great teaching? Review of the underpinning research. The Sutton Trust. https://www.suttontrust.com/wp-content/uploads/2014/10/What-Makes-Great-Teaching-REPORT.pdf

Coyne, D. M., McCoach, B., Ware, S., Austin, C. R., Loftus-Rattan, S. M., & Baker, D. L. (2018). Racing against the vocabulary gap: Matthew effects in early vocabulary instruction and intervention. Exceptional Children, 85(2). https://doi.org/10.1177/0014402918789162

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*, 311–325.

Department for Education. (2019a). National curriculum assessments at key stage 2 in England, 2019 (provisional). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/851798/KS2_Revised_publication_text_2019_v3.pdf

Department for Education. (2019b). Schools, pupils and their characteristics: January 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/812539/Schools_Pupils_and_their_Characteristics_2019_Main_Text.pdf

Duke, N., & Moses, A. (2003). *10 research tested ways to build children's vocabulary*. Scholastic Inc.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology.* Dover Publications. (Original work published 1885)

Education Endowment Foundation. (2018). *Collaborative learning: Teaching and learning toolkit*. https://educationendowmentfoundation.org.uk/pdf/generate/?u=https://educationendowmentfoundation.org.uk/pdf/toolkit/?id=152&t=Teaching%20and%20Learning%20Toolkit&e=152&s=

Education Endowment Foundation. (2021a). *Five evidence-based strategies to support high-quality teaching for pupils with SEND*. https://educationendowmentfoundation.org.uk/news/five-evidence-based-strategies-pupils-with-special-educational-needs-send

Education Endowment Foundation. (2021b). *Teaching and learning toolkit*. https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit

Education Endowment Foundation. (2021c). *Cost*. https://educationendowmentfoundation.org.uk/evidence-summaries/about-the-toolkits/cost/?utm_source=site&utm_medium=search&utm_campaign=site_search&search_term=toolkit%20cost

Elleman, A., Lindo, E., Morphy, P., & Compton, D. (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.

Fricke, S., Burgoyne, K., Bower-Crane, C., Kyriacou, M., Zosimidou, A., Maxwell, L., Lervag, A., Snowling, M. J., & Hulme, C. (2017). The efficacy of early language intervention in mainstream school settings: A randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *58*(10), 1141–1151. https://doi.org/10.1111/jcpp.12737

GL Assessment. (2018). *Technical information. New group reading test (NGRT) digital edition*. https://www.gl-assessment.co.uk/media/294737/ngrt-technical-final-proof.pdf

Gorard, S., Siddiqui, N., & See, B. H. (2019). *The difficulties of judging what difference the Pupil Premium has made to school intakes and outcomes in England*. Research Papers in Education. https://doi.org/10.1080/02671522.2019.1677759

Higgins, S., Katsipataki, M., Kokotsaki, D., Coleman, R., Major, L. E., & Coe, R. (2014). *The Sutton Trust Education Endowment Foundation teaching and learning toolkit*. Education Endowment Foundation.

Hunt, A., & Beglar, D. (1998). Current research and practice in teaching vocabulary. *The Language Teacher*, *22*(1). https://jalt-publications.org/old_tlt/articles/1998/01/hunt?y=1998&mon=01&page=hunt

Igel, C. C. (2010). *The effect of cooperative learning instruction on K-12 student learning: A meta-analysis of quantitative studies from 1998 to 2009* [Doctoral dissertation, University of Virginia]. ProQuest Dissertations & Theses Global. https://www.proquest.com/pqdtglobal/docview/816601038/C0D54F22E9FF4311PQ

Joffe, V. L., Rixon, L., & Hulme, C. (2019). Improving storytelling and vocabulary in secondary school students with language disorder: A randomized controlled trial. *International Journal of Language and Communication Disorders*, *54*(4), 656–672. https://doi.org/10.1111/1460-6984.12471

Johnson, D. W., & Johnson, R. T. (2012). Restorative justice in the classroom: Necessary roles of cooperative context, constructive conflict, and civic values. *Negotiation and Conflict Management Research*, *5*(1), 4–28.

Johnson, D. W., Johnson, R. T., & Roseth, C. (2010). Cooperative learning in middle schools: Interrelationship of relationships and achievement. *Middle Grades Research Journal*, *5*(1), 1–18.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis*. University of Minnesota.

Marulis, L. M., & Newman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, *80*, 300–335.

Marzano, R., & Pickering, D. (2005). *Building academic vocabulary: Teacher's manual*. Association for Supervision and Curriculum Development.

Medical Research Council. (2000). *A framework for development and evaluation of RCTs for complex interventions to improve health*. https://mrc.ukri.org/documents/pdf/rcts-for-complex-interventions-to-improve-health

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Erlbaum.

Moody, S., Hu, X., Kuo, L.-J., Jouhar, M., Xu, Z., & Lee, S. (2018). Vocabulary instruction: A critical analysis of theories, research, and practice. *Education Sciences*, *8*(4). https://doi.org/10.3390/educsci8040180

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *The BMJ*. https://doi.org/10.1136/bmj.h1258

Moore, W., Hammond, L., & Fetherston, T. (2014). Strengthening vocabulary for literacy: An analysis of the use of explicit instruction techniques to improve word learning from story book read-alouds. *Australian Journal of Learning Difficulties*, *19*(2), 153–172. https://doi.org/10.1080/19404158.2014.964992

Nussbaum, M. C. (2000). *Women and human development: The capabilities approach*. Cambridge University.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis, in precursors of functional literacy. In L. Verhoeve, C. Elbro, & R. Pieter (Eds.), *Study in written language and literacy* (pp. 189–213). John Benjamins.

Pishghadam, R. (2011). Parental education and social and cultural capital in academic achievement. *International Journal of English Linguistics*, *1*(2), 50–57.

Puzio, K., & Colby, G. T. (2013). Cooperative learning and literacy: A meta-analytic review. *Journal of Research on Educational Effectiveness*, *6*(4), 339–360, https://doi.org/10.1080/19345747.2013.775683

Ricketts, J., Lervåg, A., Dawson, N., Taylor, L. A., & Hulme, C. (2020). Reading and oral vocabulary development in early adolescence. *Scientific Studies of Reading*, *24*(5), 380–396. https://doi.org/10.1080/10888438.2019.1689244

Seabrook, R., Brown, G. D. A., & Solity, J. E. (2005). Distributed and massed practice: From laboratory to classroom. *Applied Cognitive Psychology*, *19*(1), 107–122. https://doi.org/10.1002/acp.1066

Sen, A. (1992). *Inequality re-examined*. Oxford University Press.

Sen, A. (1999). *Development as freedom*. Alfred Knopf.

Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, *25*(5), 763–767. https://doi.org/10.1002/acp.1747

Soper, D. (2019). *A-priori sample size calculator for multiple regression*. Free Statistic Calculators. https://www.danielsoper.com/statcalc/calculator.aspx?id=1

SPIRIT. (2015). *SPIRIT statement*. http://www.spirit-statement.org/spirit-statement

Stahl, S., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model based meta-analysis. *Review of Educational Research*, *56*, 72–110.

Stahl, S. A., & Kapinus, B. (2001). *Word power: What every educator needs to know about teaching vocabulary*. National Education Association.

Thurston, A., Cockerill, M., Chiang, T.-H., Taylor, A., & O'Keeffe, J. (2020). An efficacy randomized controlled trial of reciprocal reading in secondary schools. *International Journal of Educational Research*, *104*. https://doi.org/10.1016/j.ijer.2020.101626

Thurston, A., Cockerill, M., & Chiang, T.-H. (2021). Assessing the differential effects of peer tutoring for tutors and tutee. *Education Sciences*, *11*. https://doi.org/10.3390/educsci11030097

Tymms, P., Merrell, C., Thurston, A., Andor, J., Topping, K., & Miller, D. (2011). Improving attainment across a whole district: School reform through peer tutoring in a randomized controlled trial. *School Effectiveness and School Improvement*, *22*(3), 265–289. https://doi.org/10.1080/09243453.2011.589859

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, *22*, 271–296.

Veenman, M. V. (2016). Metacognition. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 26–40). Routledge.

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*(1), 8–25. https://doi.org/10.1080/10888438.2011.536125

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Waldfogel, J., & Washbrook, E. (2010). *Low income and early cognitive development in the U.K.: A report for the Sutton Trust*. Sutton Trust. https://www.suttontrust.com/wp-content/uploads/2010/02/Sutton_Trust_Cognitive_Report-2.pdf

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing*, *16*, 5–20. https://doi.org/10.1023/a:1021789804977

Wanzek, J. (2014). Building word knowledge: Opportunities for direct vocabulary instruction in general education for students with reading difficulties. *Reading & Writing Quarterly*, *30*, 139–164. https://doi.org/10.1080/10573569.2013.789786

Wells, G. (1985). *Language development in the preschool years*. Cambridge University Press.

Wells, G. (1987). *The meaning makers*. Hodder & Stoughton.

Wood, D. (1998). *How children think and learn: The social contexts of cognitive development* (2nd ed.). Blackwell.

Wiseheart, M., Küpper-Tetzel, C. E. M., Weston, T., Kim, A., Kapler, I., & Foot-Seymour, V. (2019). Enhancing the quality of student learning using distributed practice. In J. Dunlosky & K. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 550–584). Cambridge University Press. https://doi.org/10.1017/9781108235631.023

Wright, T. S., & Cervetti, G. N. (2017). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, *52*(2), 203–226. https://ila.onlinelibrary.wiley.com/toc/19362722/2017/52/2

About the Authors

Maria Cockerill (PhD) is senior research fellow at Queen's University Belfast (UK) and visiting professor at Universidad de Los Andes (Chile) and Pontificia Universidad Javeriana (Colombia). Her research interests comprise of experimental research design for education trials in schools, to improve academic and wellbeing outcomes for students in high-poverty areas aged 3–18 (including current projects in Chile, Colombia, and the UK). She is associate editor of *Education 3–13* (Taylor & Francis) and editorial board member in the *International Journal of Educational Research* (IJER-Elsevier).

Allen Thurston holds research interests in undertaking experimental research design for educational trials and trials of health initiatives in schools. Theoretically, his work focusses on pedagogical perspectives surrounding using cooperative learning, peer tutoring and improving literacy for students living in high poverty communities aged 3–18 (including current projects in Colombia, Chile, Lesotho, and South Africa). He is the editor-in-chief of the *International Journal of Educational Research* (IJER) and is on the editorial board of *Technology, Pedagogy and Education* and *Educational Psychology*.

Andy Taylor is the training and development manager of FFT Literacy, part of the Fischer Family Trust. His role includes developing, and providing training on, effective literacy intervention programmes for schools which improve children's and young people's outcomes in reading and writing. Using the best available research evidence and working closely with colleagues from Queen's University Belfast, FFT Literacy has recently developed programmes on reciprocal reading, reading fluency, and vocabulary development, providing initial CPD and support for implementation.

Joanne O'Keeffe is post-doctoral researcher in the School of Social Sciences, Education and Social Work at Queen's University Belfast. She is interested in research with children and young people in the area of academic attainment, health, and wellbeing, in particular those children and young people from socio-economic disadvantaged backgrounds. Her work mainly involves using randomised controlled trial methodologies and evidence-based programme evaluation in education. She currently supports the University's Widening Participation Unit (WPU) on suitable evaluation strategies of outreach delivery.

Tien-Hui Chiang has produced over 130 articles related to globalization and education policy, sociology of education, and sociology of curriculum. Currently, he is the vice-president of RC04 (Sociology of Education) in UNESCO's International Sociological Association, a constitutional standing committee of UNESCO's World Congress of Comparative Education Societies (WCCES). He is also an associate editor of the *International Journal of Educational Research* (SSCI) and international advisor for a number of international academic bodies, such as IAFOR, SGEM, and ELLTA.