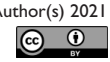
Check for updates

Article

FIRST LANGUAGE

First Language
1–28
© The Author(s) 2021

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0142723721989471
journals.sagepub.com/home/fla
SAGE

Universal strategies for the improvement of expressive language skills in the primary classroom: A systematic review

Keeley L. Dobinson and Julie E. Dockrell
UCL Institute of Education, UK

Abstract

Oral language skills underpin children's educational success and enhance positive life outcomes. Yet, significant numbers of children struggle to develop competence in speaking and listening, especially those from areas of high economic deprivation. A tiered intervention model, graduating the level of provision in line with levels of need, has been posited as most appropriate for supporting children's language development. The first tier, or universal provision, is characterised by high-quality, evidence-informed language teaching for all. To date, our understanding of effective universal language delivery remains limited, particularly in the primary-school age range. This systematic review addresses this gap by identifying and evaluating existing evidence with the aim of informing practice and future research. Following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, a systematic search protocol was used to identify experimental and quasi-experimental studies evaluating universal approaches designed to support children's oracy skills. Thirty-one studies were identified for inclusion and their characteristics and findings are reported and their reliability evaluated. Studies provide indicative evidence for the effectiveness of interactive book reading, structured vocabulary programmes, manualised curricula and approaches involving speech and language therapists. The strengths and weaknesses of our current knowledge are outlined and implications for practice and research are discussed.

Keywords

Primary school, elementary school, tiered provision, universal provision, expressive language, social disadvantage

Corresponding author:
Keeley L. Dobinson, Department for Psychology and Human Development, UCL Institute of Education, 25 Woburn Square, London WC1H 0AA, UK.
Email: keeley.dobinson.15@ucl.ac.uk

Introduction

The need for high-quality universal language provision

Well-developed oral language skills are strongly associated with academic achievement (Roulstone et al., 2011; Spencer et al., 2017), support literacy development (Snow, 2016) and are an important tool for learning across the curriculum (Alexander, 2013). The importance of oral language extends beyond academic success, impacting on social, emotional, and mental health, both at school (Benner et al., 2002) and during later life (Schoon et al., 2010). Oral language is thus a foundation for learning and achievement. However, many children struggle to develop oral language skills. At school entry in the United Kingdom, an estimated 7.58% of children have clinically significant language disorders (Norbury et al., 2016), mirroring levels observed in the United States (Tomblin et al., 1997) a decade earlier. In economically deprived areas, 40% of children are reported to have delayed language (Law et al., 2011), with the most economically deprived experiencing the most marked delays (Law, Todd et al., 2013). Furthermore, school closures arising out of the Covid-19 pandemic may have widened the 'language gap' (Oracy All-Party Parliamentary Group [APPG], 2020).

The number of children requiring support in a context of limited resources has placed strain on the traditional models of individualised identification and treatment of children with language-learning needs by Speech and Language Therapists/Pathologists (SLTs). These high levels of need are reported to outstrip capacity (Bercow, 2008; Law, 2019b; Law, Reilly et al., 2013). Many children requiring support are not being identified (Norbury et al., 2016; Tomblin et al., 1997) and, in England, inconsistencies in access to speech and language services have been identified (Longfield, 2019). These challenges have led to a move towards a continuum of speech and language provision, whereby the degree of support is graduated in line with the level of need of the particular child (Bercow, 2008; Ebbels et al., 2019; Gascoigne, 2006; Lindsay et al., 2012). Tiered approaches to support children's educational needs are well established. For example, in the US response to intervention models (Fuchs & Fuchs, 2006) or multi-tiered systems of support (Burns et al., 2016), where children who do not progress with effective universal treatment are offered enhanced support, are common. This approach is also used in the UK education context to support effective provision for children with special educational needs and disabilities (SEND; see SEND Code of Practice, Department for Education, 2015). In such models, the first tier of provision is offered universally and entails high-quality, evidence-based, language-supporting, or 'quality first', teaching for all children in education settings. Universal provision follows typical pedagogical approaches and may occur as whole-class teaching, small-group activities or individualised tasks. Where the first tier of provision is not sufficient to support an individual's language-learning needs, a second tier of 'targeted' support is offered, again in an education context but more intensively, and often includes withdrawal from regular classroom activities. The third and final tier of provision offers the specialist support of SLTs to those with the most persistent and complex language problems (Ehren & Nelson, 2005; Lindsay et al., 2012).

Tiered approaches are grounded in public health principles (Bercow, 2008; Ehren & Nelson, 2005; Greenwood et al., 2017; Law, 2019b) which seek to prevent difficulties in

the population. This is arguably an appropriate means of addressing language-learning needs given their scale and uneven distribution (Law, 2019a) and the fact that all children need strong oracy skills (Oracy APPG, 2020). Implementing a tiered approach could lead to a number of positive outcomes, including earlier and more accurate identification of those with language-learning needs (Dockrell et al., 2012), more equitable access to appropriate support (Law, Reilly et al., 2013) and the more efficient and cost-effective allocation of specialist resources (Ebbels et al., 2019; Lindsay et al., 2012).

The move to tiered provision necessitates changes to traditional ways of working for both SLTs (Ebbels et al., 2019; Gascoigne, 2006; Lindsay et al., 2012) and teachers, who need a secure grasp of evidence-based, language-supporting practice (Dockrell et al., 2012). However, teachers are reported to lack confidence in supporting oral language (The Communication Trust, 2017), a focus on oral language in schools and classrooms has not been explicit (Jones, 2017), and there is accumulating evidence that further training and support is required to implement effective language-supporting practices in classrooms (Carter, 2015; Dockrell et al., 2017; Dockrell & Lindsay, 2001). Devising effective training is complex as, while there is considerable evidence regarding targeted and specialist interventions (Law et al., 2012), understanding of effective universal provision remains limited to children prior to school entry (see Walker et al., 2020, for a systematic review of support for this age group).

Intervention in the primary years

The focus of research in the early years is to be expected given the pace of language development in the first 5 years of life (Shiel et al., 2012) and the importance of early intervention (Bercow, 2008). Language development is more stable after the age of 5 (Bornstein et al., 2016) and developments are arguably more subtle, making them more difficult to track (Shiel et al., 2012). Accordingly, intervention before primary school may be more effective means of addressing need (Bercow, 2008; Walker et al., 2020). However, the increased stability of language skills in the school years could arise from the more uniform nature of language-learning opportunities in primary schools (Justice et al., 2018), a suggestion supported by findings that teachers place less emphasis on language-supporting practice in favour of the curriculum as children move into formal education (Law et al., 2019). This, combined with the high numbers of children in the early school years experiencing language-learning needs (Norbury et al., 2016) and beyond (Law et al., 2011), highlights the need for a greater understanding of what works in these classroom contexts to support language learning.

The importance of expressive language

Oral language competence is commonly conceptualised as a web of interconnecting skills, each of which must be well developed if language proficiency is to be achieved (Moats, 2010). Interventions are often designed to target component skills and many studies distinguish between receptive and expressive domains of oral language. Recent studies challenge the multi-factorial representation of language, instead proposing models with fewer factors. Indeed, it has been argued that language is best represented as a unitary construct

in which the dimensions of language develop interdependently, reflecting one common factor (Bornstein et al., 2016). By contrast, Lonigan and Milburn (2017) propose a two-factor model consisting of vocabulary and syntax. While it is likely that language factors will vary with development (Tomblin & Zhang, 2006), interventions targeting specific domains remain prevalent in both research and practice (Roulstone et al., 2012) and a focus on the expressive domain may be particularly important in education contexts.

The development of the ability to use high-quality talk in the classroom is increasingly recognised as a key component of children's education (Oracy APPG, 2020). Children are expected to use expressive language for a wide range of functions in the primary classroom (Shiel et al., 2012), where verbal interaction is a key tool for learning across the curriculum (Alexander, 2013). Teachers use spoken language to assess learning (Petersen et al., 2010) and their perceptions of expressive language ability have been significantly correlated with their perceptions of a child's overall development (Vega et al., 2018). There is also evidence that teachers adjust the complexity of their language in response to the language used by their pupils (Justice et al., 2013), whereby children's expressive language may predict the quality of language provision they receive. When considered alongside evidence that expressive language skills, rather than receptive skills, have close links with improved literacy outcomes (National Early Literacy Panel, 2010; Savage et al., 2017), the importance of children developing these skills across the primary phase is clear.

In England, the main focus of the spoken-language curriculum is placed on expressive skills (see curriculum extracts in supplementary materials). Teachers should therefore be familiar with making judgements regarding children's expressive language and, perhaps as a result of this, teachers are more aware when children struggle with expressive language skills than with receptive skills (Dockrell & Lindsay, 2001). Accurate assessment is key to the effective operation of a tiered system of provision, and teaching approaches are more likely to be adopted when they fit with practices already undertaken by teachers (Stoiber & Gettinger, 2016). As such, a focus on expressive language skills is, arguably, the first step in capturing the strengths and weaknesses of our current understanding of universal provision. This approach is further supported by a recent study which found that effects of a language intervention only achieved far transfer in relation to expressive measures and that effects on receptive measures were hard to achieve (Melby-Lervåg et al., 2020). Indeed, these authors argued that their finding supports a focus on developing expressive rather than receptive skills both in future studies and in the classroom.

Universal oral language support

Quality-first teaching in the early school years needs to embed effective language-learning practices and opportunities within classroom activities. Despite an extensive body of research concerning targeted language interventions for vulnerable pupils (Goldstein et al., 2016; Kong et al., 2019; Law et al., 2012, 2017), effective universal approaches have been less frequently considered and randomised controlled trials (RCTs) of universal approaches in schools are uncommon (Law et al., 2012). Evidence is sparse in relation to children in the primary/elementary school age range but some potential strategies emerge from studies that have included younger children.

The efficacy of interactive book reading (IBR) for the development of young children's oral language has been identified in multiple meta-analyses (Blok, 1999). More specifically, in respect of expressive vocabulary, Mol et al.'s (2009) review reported a moderate effect size ($d = 0.62$) and Wasik et al. (2016) emphasised that adult–child interaction is key for the development of vocabulary. This finding accords with social-interactionist theories of language development (Hoff, 2006), which also underpin studies focussing on the quality of interactions between teachers and children. Although no meta-analysis has been undertaken in relation to strategies for the development of narrative skills, in line with research linking the quality of language in the home to improved language (Hirsh-Pasek et al., 2015), the conversational responsivity of teachers can have a positive impact on the amount and complexity of children's language (Girolametto et al., 2003; Justice et al., 2018; Piasta et al., 2012). Furthermore, a meta-analysis of vocabulary interventions suggests such approaches have a strong positive effect on expressive language development ($g = 0.69$; Marulis & Neuman, 2010). However, as this meta-analysis included targeted interventions, its conclusions are not directly applicable to universal classroom practice.

Early language development has been conceptualised as progressing along a developmental continuum of increasing complexity with children progressing at different rates (Shiel et al., 2012). Such a model indicates that effective universal interventions in the primary years are likely to build on the principles established in studies including younger children. However, it is important that approaches are evaluated with the age range they were intended for and in the context in which they will be employed, given that the efficacy of practices has varied depending on the age of participants (Mol et al., 2009) and given the need for ecologically valid interventions (Greenwood et al., 2020). In England, the reception year (age 4–5) is part of compulsory schooling but is guided by the early years play-based curriculum. For this first year of formal education, language-supporting practice may best be guided by the findings of studies focussed on the early years. However, the reception year is taught in the primary school and, to support transition into the more formal curriculum which commences in Key Stage 1 (age 5–7), the degree of formal teaching increases over the course of the year (Aubrey, 2004). Accordingly, effective strategies may differ from those implemented in pre-schools/nurseries and, as prior research has emphasised the importance of embedding oral language work within the school curriculum in reception alongside other key stages (Dockrell et al., 2012; Lindsay et al., 2012), an understanding of effective language-supporting practice from reception through to the end of primary phase (age 11) is key.

Review objectives

To offer effective universal oral language support, it is important to establish principles to guide practice. To further our understanding of which approaches are effective, this systematic review seeks to address the following research questions:

1. What does current evidence tell us about the efficacy and the utility of universal strategies for the development of the expressive language of 4- to 11-year-olds?
2. What are the strengths and limitations of the current evidence base and what are the implications for future research?

Methodology

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement (Moher et al., 2009) guided the methodology of this systematic review.

Selection criteria

The following selection criteria were formulated in order to capture robustly designed studies of interventions described by the first research question:

(a) Participants must be aged 4–11 (which corresponds to the age range for children attending primary schools in England);
(b) Participants represent the pupils who are educated in mainstream inclusive classrooms. Studies which focus exclusively on pupils with SEND or second language learners are excluded;
(c) The intervention/strategy/approach must seek to enhance oral language development;
(d) The intervention should be intended for universal delivery, which is not conditional on the identification of below-average language scores for participants. Delivery of the language programme may occur at class, group or individual level;
(e) In line with the Communication Trust's What Works 'moderate' evidence grading (see http://www.thecommunicationtrust.org.uk/projects/what-works/description-of-evidence-levels/), studies must be either randomised controlled intervention or quasi-experimental intervention studies with a pre-test/post-test design, employing either a business-as-usual control or alternative treatment comparison group;
(f) At least one outcome measure must include expressive language performance.

Study selection

Following an initial review of the literature, key terminology was collated. Trial searches were undertaken to identify the most suitable Boolean search terms. Limiting terms were included to ensure specificity while maximising scope. The final search terms are set out in Table 1. These were entered into the databases listed in the supplementary materials with no date restrictions. Searches returned 44,709 papers in total (see supplementary materials for a breakdown of this total by database). Update functions were used for searches and the cut-off date for inclusion was 17 August 2018. Reference sections of included studies and recent relevant textbooks were screened for relevant studies resulting in the identification of five further papers.

Identified studies were collated using EPPI Reviewer 4 (Thomas et al., 2010). Duplicates were identified and removed. Two stages of screening were then undertaken by the first author with selection criteria applied first to titles and abstracts and subsequently to full texts. A flow chart detailing the selection process is set out in Figure 1.

Table 1. Search terms.

Main search	TI ('oral communication' OR 'oral language' OR 'spoken language' OR oracy OR 'speaking and listening' OR vocabulary OR dialog* OR speech OR talk OR 'verbal communication' OR 'language use' OR 'language usage' OR 'language skills' OR 'language development') AND AB (school* OR class* OR teach*) NOT (aborigin* OR indigenous OR L2 OR EFL OR 'English as a foreign language' OR bilingual* OR 'second language' OR adolescent* OR 'secondary school' OR 'hearing loss' OR 'cochlear implants' OR deaf OR CLIL)
Subsidiary search 1	Dialogic reading OR shared reading OR joint book reading AND teach* OR class* OR school*
Subsidiary search 2	(communication or language) AND (rich or enhancing or friendly or enabling) AND (classroom* or environment* or setting* or school*) NOT (aborigin* OR indigenous OR L2 OR EFL OR 'English as a foreign language' OR bilingual* OR 'second language' OR adolescent* OR 'secondary school' OR 'hearing loss' OR 'cochlear implants' OR deaf OR CLIL)

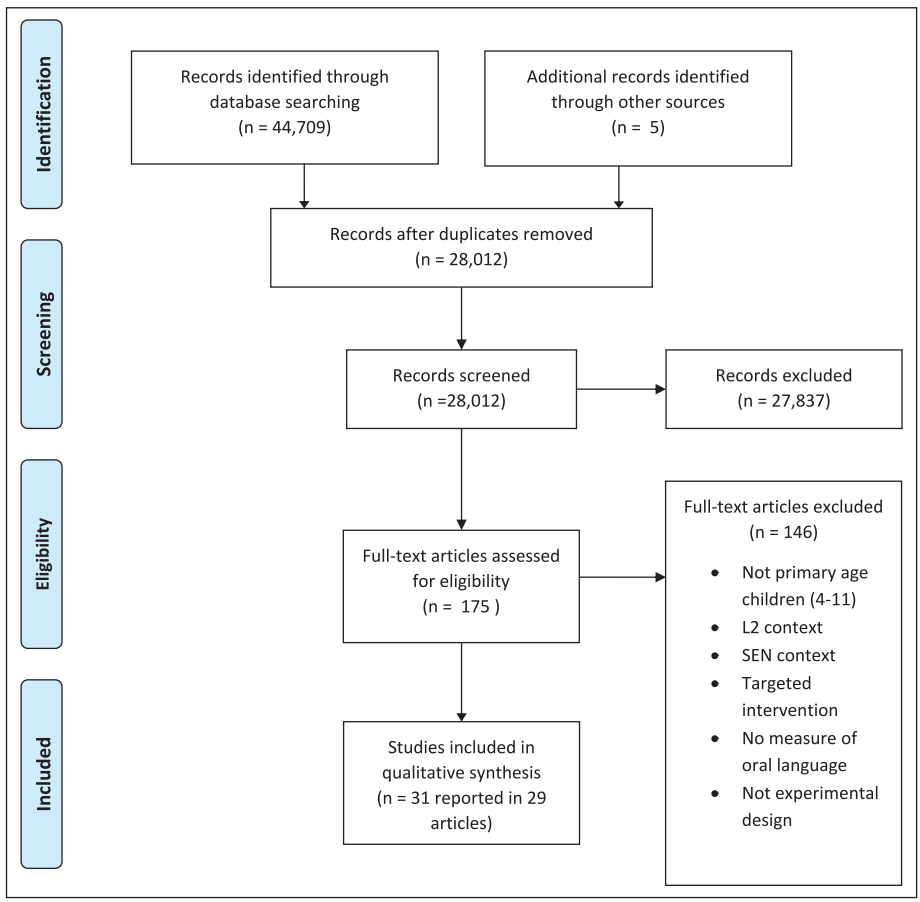


Figure 1. PRISMA flow chart detailing selection process adapted from Moher et al. (2009).

Table 2. Bias assessment tool.

Domain	Questions considered
Selection bias	Was allocation randomised? If not, were groups comparable at baseline?
Detection bias	Was group allocation/identity concealed from data collectors?
Attrition bias	If attrition reported, was impact on baseline comparability addressed?
Reporting bias	Were data reported for all participants? Where data were missing was explanation provided?
Confounding variables	Aside from the intervention, was treatment of participants consistent? Did treatment fidelity checks indicate the intervention was delivered consistently?

Data extraction

A data extraction form based on the EPPI-Centre (2003) Review Guidelines was used to extract key information. Completed data sheets were stored electronically.

Quality appraisal

In line with Cochrane Foundation guidance (Higgins & Green, 2011), a domain-based risk assessment was conducted using an adapted version of the Cochrane Collaboration's tool for assessing risk of bias. The domains evaluated are listed in Table 2. Risk in each domain was categorised as low when the questions listed in Table 2 could be answered affirmatively, high when they could not and unclear when insufficient information was available. Treatment fidelity measures were reviewed to establish whether the content of the intervention, the dose of the intervention and factors which could impact on the level of implementation were considered (Carroll et al., 2007). It was also noted if studies considered levels of treatment fidelity when interpreting results (Moncher & Prinz, 1991).

Data analysis

Extracted data were grouped by variable. Numbers of studies which measured each variable were recorded and, in order to quantify the extent of the evidence base, a percentage of the total number of participants in the included studies was calculated (TP).

Results

Thirty-one intervention studies reported in 29 papers met the inclusion criteria. Included studies are marked * in the reference list. A summary of their characteristics is provided in the supplementary materials and a descriptive summary follows.

Study characteristics

Studies originated from six countries and were conducted in five languages, with the majority from the United States ($n = 22$) and conducted in English ($n = 25$). Sample

sizes ranged from 40 to 1296, with most studies involving between 50 and 100 participants.

Participant characteristics

Details of participant characteristics reported in each study are provided in the supplementary materials and are summarised below.

Demographic information. The studies included 5097 participants aged between 4 and 7. Gender was reported for 73% of participants ($n = 24$), with male and female participants equally represented. Race/ethnicity was reported for 63% of participants ($n = 17$), 47% of whom were African American, 26% White/Caucasian and 17% Hispanic/Latino.

Special educational needs. Twenty studies (40% TP) reported no SEND information, four (16% TP) included only typically developing participants and one (2% TP) reported the exclusion of participants with severe learning difficulties.

Four studies (40% TP) reported that a total of 162 participants had identified SEND requiring additional support, and two studies (2% TP) reported school-level, rather than participant-specific, SEND information. The variation in reporting means that no clear picture of the range of needs of participants emerges.

Language and literacy learning needs. Levels of language difficulties were reported in two studies (3% TP), with nine participants identified as receiving language support, and one study (25% TP) reported the exclusion of participants receiving language support.

Three studies (5% TP) reported that some or all participants were at risk of language or reading problems based on a range of standardised tests. Two studies (5% TP) reported low literacy levels at a school or neighbourhood level.

Home language of participants. Eight studies (35% TP) confirmed that participants received the intervention in their first language or a language in which they were proficient, 9% of participants ($n = 12$) received the intervention in a language other than their first language and 11 studies (27% of participants) provided no information on participants' language status. One study (5% TP) did not report separately for primary-aged children and two (2% TP) reported school-level, rather than participant-specific, language status.

Socio-economic status. Table 3 provides a breakdown of the reported socioeconomic status of participants, reduced or free school meal eligibility being a marker of economic deprivation in both the United States and the United Kingdom. Markers of economic deprivation were confirmed for 38% TP and a further 20% TP attended schools in economically deprived areas.

Intervention characteristics

Included studies focused on two sets of expressive language skills, namely vocabulary and narrative competence, either alone or alongside one another. How studies sought to

Table 3. Breakdown of participant socioeconomic status information.

	No socioeconomic status information	Reduced or free school meals	Other markers of economic deprivation	Low maternal education	Study in school in economically deprived area
Number of studies	11	5	2	2	11
% total participants included in studies	27	33	16	3	21
% total participants confirmed as falling into category	0	31	5	2	0

embed these skills in classroom practice varied by means of implementation (with interventions delivered to pairs and individuals, through small group work, whole-class delivery or combinations thereof), and by the types of activities used. Full details of the interventions can be found in the supplemental materials and a descriptive summary follows.

Approaches to intervention. A range of activities have been used to deliver the targeted skills, some individually and others in combination. Studies of single-strategy approaches ($n = 18$) included play-based teaching ($n = 3$), vocabulary instruction ($n = 2$), the use of technology ($n = 3$), conversation-facilitating approaches ($n = 4$), IBR ($n = 5$) and narrative instruction ($n = 1$). The remaining studies ($n = 13$) employed multiple-strategy approaches, including structured IBR programmes incorporating vocabulary instruction ($n = 6$), the use of manualised curricula incorporating various language stimulation methods ($n = 4$), large-scale professional development training ($n = 1$) and SLT involvement in teaching ($n = 2$).

Mode of delivery. Most interventions ($n = 22$, 91%TP) were delivered in a whole class context, four (4% TP) to small groups and one (1% TP) through a combination of these modes. Three (3% TP) were delivered to individuals and one (1% TP) to pairs of participants.

Intervention duration. Intervention periods ranged from 1 week to 18 months, with most ($n = 18$) lasting between 2 and 6 months, three lasting 1 year or longer and the remainder ($n = 10$) less than 2 months. Twenty studies provided exact dosage information, with frequency ranging from one to six sessions per week, duration ranging from 11 minutes to 1 hour per session and total intervention time ranging from 1 hour to 30 hours ($M = 13.6$ hours). Dosage information was not provided in the majority of studies implementing curriculum-based approaches.

Interventionists. Teachers delivered interventions in 15 studies (80% TP), researchers in 9 studies (11% TP) and a combination thereof in 2 studies (3% TP). Interventions were also provided by SLTs ($n = 2$, 3%TP) and university students ($n = 3$, 3% TP). All

teacher-delivered interventions utilised professional development training and seven of these included support from an expert mentor, although just one study evaluated the impact of mentoring experimentally (Assel et al., 2007). Training periods ranged from 1 hour to 5 days.

Study design

Outcome measures. Studies measured the general expressive vocabulary of participants using standardised measures ($n = 12$), the acquisition of specifically targeted vocabulary using study-specific, researcher-designed measures ($n = 14$) and, in some cases, both ($n = 12$). Eleven studies undertook narrative analysis of language samples, three in combination with study-specific vocabulary measures, one alongside standardised vocabulary measures and two in conjunction with teacher ratings of expressive language skills.

Measures of reliability were reported for 83% of standardised tests of expressive vocabulary, for 71% of study-specific vocabulary measures and for 82% of narrative measures. An assortment of internal consistency, test–retest, inter-rater and split-half reliability were reported and ranged from 0.70 to 0.99 for standardised tests, 0.49–0.98 for study-specific vocabulary measures and 0.63–0.96 for narrative measures. Measures of validity were reported for 33% of standardised tests of expressive vocabulary (range = 0.59–0.90), 14% of study-specific vocabulary measures (range = 0.69–0.84) and 18% of narrative measures (range = 0.61–0.93). Validity was not considered for teacher reports.

Twenty-two studies employed measures of receptive vocabulary, which are detailed along with their results, in the supplemental materials. However, as receptive language was not included in the parameters of this review's search terms, these results are indicative only.

Control groups. Twenty-four studies (90% TP) compared the intervention to business-as-usual classroom practice. One compared an intervention to an alternative treatment programme (1% TP) and six (9% TP) delivered an intervention to both groups making one change to the method used. Three studies (19% TP) carried out observations and checks in the control or comparison conditions to facilitate more valid comparison between conditions.

Treatment fidelity. Treatment fidelity measures were employed in 15 studies (66% TP), of which 13 (63% TP) reported results. All 15 studies utilised study-specific checklists, 14 of which were completed during observations (mean number of observations $n = 5$, range = 1–18, not reported in two studies) and one of which utilised teacher self-report. Observations were conducted by more than one person in seven studies and inter-rater reliability ranged from 76% to 100% based on between 10% and 25% of overall observations. All 15 studies considered adherence to the content of intervention and 4 included measures of dose or exposure. There was considerable variation in the reporting of factors which may have impacted on adherence to the interventions. Four studies evaluated the quality of delivery by teachers, one documented levels of disruptive behaviour to assess the experience of participants, three considered the engagement levels of participants and three

described methods used to promote fidelity, that is, feedback following observations. Just two studies considered the impact of fidelity on outcome measures in the results section (Moncher & Prinz, 1991). Reported intervention fidelity ranged from 30% to 100%, with eight studies (16% TP) reporting average levels of intervention fidelity above 85%.

Study outcomes

Study outcomes grouped by intervention method are detailed in the supplementary materials and are summarised below. Twenty studies reported significant outcomes on all expressive measures employed, six reported significant outcomes in relation to at least one outcome measure employed and five reported non-significant outcomes or were unclear as to findings.

Effect size was reported in 14 studies. For the remaining studies, Cohen's *d* effect sizes were calculated using control and experimental group post-test means and standard deviations where these were provided. Calculated effect sizes are marked [c] and 'effect size incalculable' where insufficient information was provided in the original study.

Single strategy approaches

Vocabulary instruction. Evaluations of vocabulary instruction in isolation were scarce. One study established that the use of sign language did not enhance the general vocabulary of participants (Caron, 2005). Another reported that small-group vocabulary instruction had a significantly positive effect on general vocabulary (effect size incalculable; Benson, 2013), although the comparison group received whole-class instruction so more intensive interaction with adults cannot be excluded as an explanatory factor in this finding.

Conversation-based approaches. Three-way conversations between university students and pairs of participants had no significant impact on general vocabulary or narrative competence (Ruston & Schwanenflugel, 2010), leading the study's authors to suggest that more highly qualified conversation partners may have a greater impact, yet conversations between university students and individual participants significantly improved narrative competence (McCabe et al., 2010; effect size incalculable in both). Researcher-led whole-class decontextualised conversations also had no significant impact on narrative competence nor on specifically targeted vocabulary (Sonmez, 2010).

Play-based teaching. Two studies comparing play-based teaching approaches to 'traditional' teaching approaches reported a positive impact on narrative competence. Baumer et al. (2005) reported that joint adult-child pretence resulted in significant differences in the length ($d = 3.32^{c}$) and linguistic coherence ($d = 4.4^{c}$) of participants' narratives and marginally significant differences in narrative complexity ($d = 0.61^{c}$) compared with a business-as-usual control group, and Stagnitti et al. (2016) reported that the narrative retell abilities of participants attending schools using play-based curricula increased significantly more than those of participants attending schools employing teacher-directed approaches ($d = 0.87$). A further study comparing vocabulary games to discursive follow-up after IBR reported a significant impact on participants' acquisition of target vocabulary (Hassinger-Das et al., 2016; $d = 0.54^{c}$).

IBR. IBR was found to have a significant impact on participants' general vocabulary in studies of both small group (Simsek & Erdogan, 2015; $d = 1.12^{c}$) and whole-class delivery (Okyay & Kandir, 2017; $d = 0.62^{c}$). However, the results from Ergül et al. (2016) were less consistent as whole-class delivery led to statistically significant differences in participants' general vocabulary on just one of two standardised vocabulary measures employed ($d = 0.8^{c}$). Furthermore, in a more intensive condition combining both whole-class and small-group delivery, mean scores in the experimental group reduced ($d = 0.1^{c}$) on one standardised vocabulary measure and significance was not reached on the other. The authors reported that this result was likely to have arisen out of poor implementation by the teacher in this condition, a factor which could not be controlled for due to the small sample size.

Changes in specifically targeted vocabulary were more consistent with significant findings reported in studies of whole-class (Opel et al., 2009; $d = 2.00$) and small group IBR (Lever & Sénéchal, 2011; $d = 0.66$).

One study considered the impact of IBR on participants' narrative competence (Lever & Sénéchal, 2011) and reported a significant positive impact on retelling ($d = 0.28$) and production ($d = 0.38$) but not on measures of language complexity, anaphora or connectives.

Narrative instruction. Spencer et al. (2015) reported that researcher-led whole-class narrative instruction had a significant positive effect on participants' narrative retell skills ($d = 2.87^{c}$), which were sustained through to follow-up at 4 weeks ($d = 3.00^{c}$). Although the intervention had no significant impact on the participants' ability to generate their own story, these skills were not taught directly as part of the intervention.

Use of technology. A small group of studies evaluating the use of technology indicated that the use of e-books in place of traditional books in the classroom supported the acquisition of specifically targeted words (Ihmeideh, 2014, $d = 2.9^{c}$), whereas single viewings of television programmes did not (Silverman, 2013). Repeated exposure to television programmes had a stronger effect on the acquisition of specifically targeted words than single viewings (Silverman, 2013, $d = 0.58$), although, as the author noted, gains were minimal.

Multiple strategy approaches

Combined IBR and vocabulary instruction. Four of the six studies evaluating structured programmes of IBR and vocabulary instruction reported statistically significant and large effects on the acquisition of specifically targeted vocabulary in the experimental groups (Coyne et al., 2010, $d = 1.71$; Gonzalez et al., 2010, $g = 1.01$; Neuman & Dwyer, 2011, $d = 0.64$; Neuman & Kaefer, 2018, $d = 0.48$).

Silverman's (2007) comparison of three different types of IBR instruction reported that 'analytical' instruction, which supported analysis of target words in contexts beyond the participants' own, had a significant and positive impact on specifically targeted vocabulary ($d = 0.58$) compared with 'contextual' instruction, which focused on linking target words to participants' own experiences. 'Anchored' instruction combined analytical instruction with phonological and orthographic consideration of target words and had

a significant impact ($d = 0.94$) on specifically targeted vocabulary compared with contextual instruction.

The final study evaluating the impact of IBR on specifically targeted words only analysed significance in combination with receptive language measures and effect size was incalculable (Vuattoux et al., 2013).

These structured programmes did not have the same impact on participants' general vocabulary as on specifically targeted vocabulary, with very small (d = 0.03) (Neuman & Kaefer, 2018) or non-significant effects (Gonzalez et al., 2010) reported in the two studies evaluating this.

Manualised curricula. Two studies of manualised curricula reported significant differences in respect of participants' general vocabulary. The first set out a programme of daily oral language activities (Kaumans, 1972, $d = 0.28^{c}$), and the second introduced target words using a variety of approaches including explicit vocabulary instruction, IBR and adult–child conversations (Goodson et al., 2010, $d = 0.14$). A further study considered the impact of two structured curricula focussed on the development of different oral language skills across three different forms of early years provision in the United States, namely Head Start settings which serve low-income communities, Title 1 settings which also serve low-income populations but have higher levels of staff qualification, and Universal Pre-kindergarten settings serving children from more affluent backgrounds (Assel et al., 2007). This study reported significant differences between classrooms using a curriculum versus the controls on their overall growth rates in general expressive vocabulary, but effect size was moderated by site (see differential impact below). These findings were not replicated in relation to narrative competence on which a language-focussed curriculum encouraging interaction had no significant impact (Justice et al., 2008).

Professional development. An RCT of a professional development programme to enhance teachers' understanding of effective delivery in various language domains employed a comprehensive battery of oral language and literacy measures, although only the measures of narrative competence are relevant to this review (Snow et al., 2013). Findings were mixed, with a significant positive impact on narrative competence of the experimental group detected by an omnibus story grammar measure ($d = 0.33$) but not by more in-depth narrative story grammar and syntactic complexity measures.

Involvement of SLTs. Speech and Language Therapist–led whole-class narrative and vocabulary instruction had a significant impact on participants' narrative competence ($d = 0.82$) and specifically targeted vocabulary ($d = 1.02$) when compared to business-as-usual provision with equivalent adult support (Gillam et al., 2014), while collaborative teaching by teachers and SLTs significantly improved the general vocabulary of the experimental group compared to a business-as-usual control (Hadley et al., 2000; $d = 0.3^{c}$).

Differential impact. Five studies evaluated whether interventions had a different impact for participants based upon pre-existing language levels, socioeconomic status or the type of setting attended.

In three studies, participants with higher baseline language levels made greater gains on specifically targeted words than those with lower baseline levels. Coyne et al. (2010) reported that effect size varied in line with pre-existing scores on the Peabody Picture Vocabulary Test (85: $d = 1.06$; 100: $d = 1.75$; 115: $d = 2.44$) and Gillam et al. (2014) reported that effect size varied in accordance with the level of risk of language difficulties on the basis of participants' performance on a narrative test (high risk $d = 0.66$ and low risk $d = 2.28$). Gonzalez et al. (2010) reported that pre-test scores predicted post-test scores in general, as well as in specifically targeted, vocabulary.

However, the opposite effect was observed in two studies employing measures of narrative competence. Ruston and Schwanenflugel (2010) reported that participants with lower initial vocabulary scores made significant gains in lexical diversity, whereas those with higher initial scores did not, and Gillam et al. (2014) reported that effect sizes in respect of self-generated narratives were greater for high-risk ($d = 1.00$) than low-risk children ($d = 0.59$). Furthermore, the adoption of a manualised curriculum had a significant impact in classrooms serving 'at-risk' children staffed by less-qualified practitioners (Head Start, $d = 0.68$) but not in settings serving economically deprived families staffed by more highly qualified practitioners (Title 1, $d = 0.04$) or settings serving more affluent communities (Assel et al., 2007, $d = -0.52$).

Quality of studies

Studies were assessed for risk of bias in five domains and the full risk of bias assessment is provided in the supplementary materials. Randomisation of participants was infrequent, and few studies considered baseline comparability, resulting in potential allocation bias in 24 studies. Risk of detection bias given the lack of blinding was high in 7 and unclear in 12 studies. Risk of attrition bias was high in 8 studies and unclear in 11 and risk of reporting bias was high in 6 and unclear in 4 studies. All the studies were considered to be at risk of bias due to the presence of potentially confounding variables although the nature and degree of these varied considerably.

Discussion

The purpose of this review was to identify studies evaluating universal strategies for the enhancement of expressive language in primary schools. Effective universal approaches to enhancing oral language are the first step in developing oracy skills and identifying those children who require greater levels of support (Ebbels et al., 2019). Thirty-one studies met the inclusion criteria, and most ($n = 26$) reported a significant, positive impact on the expressive language skills of participants on the basis of at least one measure. Together, these studies indicate that universal-level strategies have the potential to support these skills. However, the wide range of strategies used across the studies means that it is not possible to recommend a specific approach. Rather, the review highlights potential areas for developing practice and further avenues for research.

The most commonly evaluated strategy was IBR, generally with positive outcomes. Even after short periods of intervention (studies ranged from 8 days to 8 weeks), and across different modes of delivery, studies reported significant positive effects on general

(Okyay & Kandir, 2017, $d = 0.62^{c}$; Simsek & Erdogan, 2015, $d = 1.12^{c}$, respectively) and specifically targeted vocabulary (Lever & Sénéchal, 2011, $d = 0.66$; Opel et al., 2009, $d = 0.66$). However, these were small-scale studies and the mixed findings in Ergül et al. (2016) (which were in part attributed to teacher aptitude) highlight the need for larger scale, randomised studies if more robust conclusions are to be drawn. Furthermore, just one study evaluated the impact of interactive reading on narrative skills (Lever & Sénéchal, 2011) with mixed findings, and the inclusion of narrative measures in future studies would be informative.

Programmes evaluating IBR and vocabulary instruction in combination ($n = 6$) were generally conducted over longer periods than studies IBR alone (range = 6–24 weeks) and, while they reported consistently significant and large effects on children's knowledge of specifically targeted vocabulary (Coyne et al., 2010; Gonzalez et al., 2010; Neuman & Dwyer, 2011; Neuman & Kaefer, 2018) ($d = 1.71$; $g = 1.01$; $d = 0.64$; and $d = 0.48$, respectively), findings were non-significant in respect of general vocabulary (Gonzalez et al., 2010; Neuman & Kaefer, 2018). Authors noted that the degree of impact on general vocabulary was greater for younger children in the sample, as was the case in Mol et al.'s (2009) study. Silverman's (2007) evaluation of different approaches to IBR reported that strategies encouraging abstraction and engagement with the form of target words had a greater impact on their acquisition than those drawing links with children's own experiences. These studies emphasise the importance of research across the age ranges to establish an understanding of relative efficacy of the intervention for different age groups and ability levels and the need for studies to undertake a comparative analysis of different approaches. Studies which address these factors would allow for the identification of the most efficacious methods for different groups of children.

Studies of manualised curricula combining a number of language-supporting strategies reported significant effects on participants' general expressive vocabulary (Assel et al., 2007 (effect size varied by setting as detailed in differential impact above); Goodson et al., 2010 ($d = 0.14$); Kaumans, 1972 ($d = 0.14$), indicating that focused language instruction over extended periods (26, 18 weeks and 90 days, respectively) can impact on expressive vocabulary acquisition, yet a language-focused curriculum encouraging increased and higher quality interactions reported no significant impact on the narrative competence of participants (Justice et al., 2008). Findings in respect of narrative outcomes were similar when the impact of professional development targeting the improvement of oral language teaching was evaluated (Snow et al., 2013). There was a significant impact on narrative skills when measured by an omnibus story grammar measure ($d = 0.33$) but not when measured by more in-depth narrative story grammar and syntactic complexity measures. These findings suggest that narrative skills are more resistant to change than expressive vocabulary but, as both Justice et al. (2008) and Snow et al. (2013) note, when interpreting these findings careful consideration must be given the nature of instruction in the comparison groups. It may be, for example, that instruction in these classrooms was already supportive of narrative development and that the interventions did not add value to the provision already in place. However, it is equally plausible that before impacts on narrative are evident children, need good expressive skills at word and utterance level. Only two included studies (Gonzalez et al., 2010; Neuman & Kaefer, 2018) observed control group practice, and the inclusion of such

observations in future studies could increase both our understanding of business-as-usual classroom practice and our ability to identify why particular effect sizes are observed.

A further study targeting narrative competence through researcher-led whole-class narrative instruction reported a significant effect on participants' narrative retell skills (Spencer et al., 2015; $d = 2.87^{c}$), which were sustained through to follow-up at 4 weeks ($d = 3.00^{c}$). Although the intervention had no significant impact on the participants' ability to generate their own story, these skills were not taught directly as part of the intervention. Overall, the findings provided evidence that whole-group narrative instruction can positively impact on some expressive language skills.

Whole-class narrative and vocabulary instruction led by an SLT also significantly improved narrative competence ($d = 0.82$) as well as knowledge of specifically targeted vocabulary ($d = 1.02$) when compared to business-as-usual provision with equivalent adult support (Gillam et al., 2014), while collaborative teaching by teachers and SLTs significantly improved participants' knowledge of specifically targeted vocabulary compared to a business-as-usual control (Hadley et al., 2000) ($d = 0.3^{c}$). These findings provide a useful backdrop to the ongoing consideration of the nature of the role of SLTs at each tier of provision (Ebbels et al., 2019; Gascoigne, 2006). Larger randomised studies considering both the efficacy, as well as the viability, of such approaches would be beneficial in light of the focus on joint commissioning of services within the tiered framework.

The remaining studies in the review were of single interventions, making generalisation difficult. Similar to findings in respect of a targeted intervention (Kong et al., 2019), positive findings in relation to the use of technology were reported (Ihmeideh, 2014; Silverman, 2013). Given the increased presence of technology in primary classrooms (Levy et al., 2013), this is an important area of development. The mixed findings in respect of conversation-based approaches were surprising given the extensive literature supporting the utility of social-interactionist theories of language development (Hoff, 2006) in the early years, and future studies could usefully consider the impact that the use of more highly qualified interventionists might have on efficacy, taking into account the availability of resources. The positive findings reported by studies evaluating play-based teaching approaches indicate some potential utility, although the study design employed mean that multiple confounding factors cannot be excluded when interpreting the large ES on narrative skills ($d = 0.87$) reported in Stagnitti et al. (2016) and further research is required before clear conclusions can be drawn.

Drawing firm conclusions regarding their relative utility of the studies is not straightforward. The diverse means of implementation in terms of group size, dosage, intervention and levels of training and support provided make comparisons challenging. Some included studies were of brief duration and cannot be compared with longer-term studies as both length and intervention dosage likely impact on efficacy (Goldstein et al., 2016; Melby-Lervåg et al., 2020). Furthermore, in keeping with the findings of Wasik et al. (2016), a number of studies evaluated strategies in combination, making identification of the 'effective ingredients' of intervention approaches (Law et al., 2012) difficult. Comparability would also have been aided by the use of more consistent approaches to reporting intervention fidelity. While all the included studies considered adherence in terms of content, less emphasis was placed on recording dosage or on understanding the

mechanisms supporting implementation, with just one study considering the impact of coaching on outcomes (Assel et al., 2007).

A further challenge to accurate comparison arises from the variety of expressive language measures used in the studies. The majority ($n = 24$) focussed on one domain of expressive language, rather than a combination, limiting generalisations, and potentially failing to capture the nature of expressive language in primary school settings. Furthermore, when testing vocabulary, few studies used a mixture of standardised and study-specific vocabulary measures. Study-specific vocabulary measures can be limited by their lack of standardisation and potential subjectivity, testing retention of what was taught rather than skill development at a general level (Blok, 1999), yet standardised vocabulary measures may not always be sensitive enough to detect change (Marulis & Neuman, 2010). A more robust understanding of intervention effectiveness could be established by using a combination of measures (Goldstein et al., 2016; Wasik et al., 2016) along with the more consistent of reporting of effect sizes. This would facilitate comparison to a range of standardised language benchmarks (Schmitt et al., 2017) deepening our understanding of the educational utility of particular interventions.

When evaluating relative impact, it is also important to consider the longevity of positive effects (O'Connor et al., 2009). Just four studies in the review carried out delayed follow-up testing (Goodson et al., 2010; Okyay & Kandir, 2017; Silverman, 2007; T. D. Spencer et al., 2015) limiting our understanding of longer-term effects of the interventions.

Given the social gradient of language problems (Law, Todd, et al., 2013), it also crucial that we understand how effectiveness varies in line with participants' socioeconomic status and initial language ability. Five studies evaluated whether interventions had a differential impact for participants based upon pre-existing language levels, socioeconomic status or the type of setting attended. In three studies, participants with higher baseline language levels made greater gains on study-specific vocabulary measures than those with lower baseline levels (Coyne et al., 2010; Gillam et al., 2014; Gonzalez et al., 2010). Where these apparent 'Matthew effects' (Stanovich, 1986), with the 'word-rich' benefitting the most, emerge it is possible that application of strategies at the universal level could in fact widen the word gap. The opposite effect was observed in two studies employing measures of narrative competence (Gillam et al., 2014; Ruston & Schwanenflugel, 2010), a more promising pattern when considering the public health aims of tiered provision. A further study (Assel et al., 2007) reported that a manualised curriculum had a significant impact in classrooms serving at-risk children staffed by less-qualified practitioners (Head Start, $d = 0.68$) but not in settings serving economically deprived families staffed by more highly qualified practitioners (Title 1, $d = 0.04$) or settings serving more affluent communities ($d = -0.52$), suggesting that levels of staff qualification could also influence outcomes. Further investigation of these differential effects should be a key feature of future research as this will increase understanding of the elements of language development which can effectively be addressed in whole-class instruction and those for whom targeted support is more appropriate. Analysis of responsiveness as undertaken by Spencer et al. (2015) would aid these considerations as it provides useful information about who universal provision works for and who may require additional support.

Furthermore, conducting studies in schools serving economically deprived areas, as was the case in 20 included studies, is key (Lindsay et al., 2012) as it increases ecological validity, with research findings more likely to be generalisable to settings in areas with the highest levels of need (Greenwood et al., 2020).

Another key element in understanding differential impact is the rigorous and consistent reporting of precise demographic information, something which was lacking in many of the studies included in this review, and which future studies should seek to ensure (Greenwood et al., 2020).

While demonstrating efficacy in research studies is necessary, it is not sufficient to guarantee transfer to practice (Bleses et al., 2018; Greenwood et al., 2020; Justice et al., 2008). Interventions led by researchers, as in Spencer et al.'s (2015) study, do not necessarily have the same impact when transferred into naturalistic classroom practice (Mol et al., 2009). A strength of the included studies is that a number of significant and positive outcomes were achieved by teacher-led interventions (n = 15) and future studies should continue to explore effectiveness in natural contexts (Bleses et al., 2018). It is also essential that educators are willing and able to incorporate the evidence-based strategies into their practice (Dagenais et al., 2012; Law et al., 2019; Lindsay et al., 2012). Stoiber and Gettinger (2016) suggest that the gap between research and practice should decrease when practitioners can easily incorporate strategies into their normal routines. As noted by Law et al. (2019) and Shiel et al. (2012), increased curricula demands, along with a greater focus on preparation for standardised assessment in the primary classroom, may make the incorporation of language-supporting practices more challenging and an understanding of the social validity of interventions (Greenwood et al., 2020) is crucial. Of the included studies, Spencer et al. (2015) considered this by way of teacher questionnaire and Goodson et al. (2010) sought to understand the impact of the intervention on time allocated to other practices in the classroom, but consideration in other studies was absent, making this a key area for development in future research.

Furthermore, although the absence of studies in the 7–11 age range is in keeping with research supporting the stability of language over time (Bornstein et al., 2016), given the persistence of language problems (Law et al., 2011), and the evidence that they are less likely to be identified in the later primary years (Meschi et al., 2012), this represents a gap in the research which future studies should seek to address.

Limitations

This review is subject to various limitations. Only English language papers were included, thereby potentially narrowing its scope. The age-related selection criteria potentially created issues with the interpretation of review findings as distinctions between early-years and more formal education vary in different regions and countries. This impacted on the potential ecological validity of included studies as well as resulting in the exclusion of many studies which cut across age ranges. Studies were assessed to be at risk of bias in multiple domains potentially limiting the reliability of the review's conclusions. A narrow range of study designs were included and given the move towards evidence-informed practice with teachers being encouraged to use a range of evidence

alongside their professional judgement to inform their practice (Coldwell et al., 2017), a broader remit with less stringent inclusion requirements may have been merited. Finally, as with any review, there was a cut-off date for the inclusion of studies and therefore some recent studies which may have enriched the analysis (see, for example, Wasik & Hindman, 2020) were not included.

Conclusion

The current systematic review identified strengths and limitations in our understanding of effective universal intervention in schools. Similar to analyses of preschool interventions (Walker et al., 2020), our current understanding is constrained by gaps in our knowledge about the interventions, the age groups targeted and the measures used. While there were too few studies to draw firm conclusions, studies indicated that whole-class narrative instruction can impact positively on some elements of narrative production, structured curricula employed over extended periods can impact positively on general expressive vocabulary measured by standardised tests, IBR combined with vocabulary instruction can impact positively on specifically targeted vocabulary and the involvement of SLTs in teaching may improve expressive language outcomes across all three measures.

Data indicated that the nature of the setting and the competency of the staff impacted on outcomes, suggesting that less-skilled professionals were more effective when manualised interventions were provided. By corollary children's language levels also impacted on efficacy. These differences raise important questions about the ways in which universal interventions should be conceptualised. All children should receive high-quality oral language support in classrooms; however, the need for improved provision is higher in some areas than others (Lindsay et al., 2012). This may best be reflected in more fine-grained representation of the tiered system which distinguishes between selective universal provision for those at risk on the basis of demographic characteristics and targeted support linked to particular children's skills or outcomes (Greenberg & Abenavoli, 2017; Law, 2019a; Law et al., 2017).

However provision is conceptualised, it must be based on what we know about effective practice and this review highlights that, as yet, data are lacking. If the recommendations of the recent Oracy APPG (2020) report are to be implemented, and teachers are to be well equipped to support the development of the expressive language of their pupils, then future studies should seek to refine and extend our understanding of universal approaches. These studies should employ robust study design and a more 'joined up' approach to facilitate clearer comparisons of utility, while also giving due consideration to how such approaches can be effectively incorporated into practice in schools in the context of available resources in the wider tiered system (Greenwood et al., 2020; Lindsay et al., 2012).

Acknowledgements

We would like to thank Dr Claire Forrest of UCL IOE for her comments on drafts of this paper and Rozz Evans of UCL Library Services for her assistance with formulation of search strategies.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship or publication of this article.

ORCID iD

Keeley L. Dobinson https://orcid.org/0000-0002-7345-9432

Julie E. Dockrell https://orcid.org/0000-0003-3595-6064

Supplemental material

Supplemental material for this article is available online.

References

Studies included in this review are marked *.

Alexander, R. J. (2013). Improving oracy and classroom talk: Achievements and challenges. *Primary First*, *10*, 22–29. http://eprints.whiterose.ac.uk/76321/

*Assel, M. A., Landry, S. H., Swank, P. R., & Gunnewig, S. (2007). An evaluation of curriculum, setting, and mentoring on the performance of children enrolled in pre-kindergarten. *Reading and Writing*, *20*(5), 463–494. https://doi.org/10.1007/s11145-006-9039-5

Aubrey, C. (2004). Implementing the foundation stage in reception classes. *British Educational Research Journal*, *30*(5), 633–656. https://doi-org.libproxy.ucl.ac.uk/10.1080/01411920420 00234629

*Baumer, S., Ferholt, B., & Lecusay, R. (2005). Promoting narrative competence through adult–child joint pretense: Lessons from the Scandinavian educational practice of playworld. *Cognitive Development*, *20*(4), 576–590. https://doi.org/10.1016/j.cogdev.2005.08.003

Benner, G. J., Nelson, J. R., & Epstein, M. H. (2002). Language skills of children with EBD: A literature review. *Journal of Emotional and Behavioral Disorders*, *10*(1), 43–56. https://doi.org/10.1177/106342660201000105

*Benson, K. A. (2013). *An investigation of the impact of small group direct vocabulary instruction on the vocabulary development of kindergarten children living in poverty* (Publication No. 3590408, Doctoral dissertation), Ball State University. https://www.proquest.com/docview/1432222477

Bercow, J. (2008). *The Bercow report: A review of services for children and young people (0-19) with speech, language and communication needs*. Department for Children, Schools and Families. https://dera.ioe.ac.uk/8405/7/7771-dcsf-bercow_Redacted.pdf

Bleses, D., Højen, A., Justice, L. M., Dale, P. S., Dybdal, L., Piasta, S. B., Markussen-Brown, J., Clausen, M., & Haghish, E. F. (2018). The Effectiveness of a large-scale language and preliteracy intervention: The SPELL randomized controlled trial in Denmark. *Child Development*, *89*(4), e342–e363. https://doi.org/10.1111/cdev.12859

Blok, H. (1999). Reading to young children in educational settings: A meta-analysis of recent research. *Language Learning*, *49*(2), 343–371. https://doi.org/10.1111/0023-8333.00091

Bornstein, M. H., Hahn, C. S., & Putnick, D. L. (2016). Stability of core language skill across the first decade of life in children at biological and social risk. *Journal of Child Psychology and Psychiatry*, *57*(12), 1434–1443. https://doi.org/10.1111/jcpp.12632

Burns, M. K., Jimerson, S. R., VanDerHeyden, A. M., & Deno, S. L. (2016). Toward a unified response to intervention model: Multi-tiered systems of support. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), Handbook of response to intervention: The science and practice of multi-tiered systems of support (pp. 121–142). Springer.

*Caron, D. B. (2005). The effect of utilizing American Sign Language as part of a supplementary vocabulary instruction program with hearing kindergarten children (Publication Number: 3174970, Doctoral Dissertation). *University of Florida*. https://search.proquest.com/docview/304995599/fulltextPDF/E1F69FDF8DCA417BPQ/1?accountid=14511

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, *2*(1), 40. https://doi.org/10.1186/1748-5908-2-40

Carter, A. (2015). Carter review of initial teacher training (ITT). *Department for Education*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/399957/Carter_Review.pdf

Coldwell, M., Greany, T., Higgins, S., Brown, C., Maxwell, B., Stiell, B., Stoll, L., Willis, B., & Burns, H. (2017). *Evidence-informed teaching: An evaluation of progress in England* (9781781057698). Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/625007/Evidence-informed_teaching_-_an_evaluation_of_progress_in_England.pdf

The Communication Trust. (2017). *Professional Development in speech, language and communication: Findings from a national survey*. https://www.thecommunicationtrust.org.uk/media/526748/1._tct_workforce_development_report_final_online.pdf

*Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Ruby, M., Crevecoeur, Y. C., & Kapp, S. (2010). Direct and extended vocabulary instruction in kindergarten: Investigating transfer effects. *Journal of Research on Educational Effectiveness*, *3*(2), 93–120. https://doi.org/10.1080/19345741003592410

Dagenais, C., Lysenko, L., Abrami, P. C., Bernard, R. M., Ramde, J., & Janosz, M. (2012). Use of research-based information by school practitioners and determinants of use: A review of empirical research. *Evidence & Policy*, *8*(3), 285–309. https://doi.org/10.1332/174426412X654031

Department for Education. (2015). *Special educational needs and disability code of practice: 0 to 25 years: Statutory guidance for organisations which work with and support children and young people who have special educational needs or disabilities*. Reference: *DFE-00205-2013*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/398815/SEND_Code_of_Practice_January_2015.pdf

Dockrell, J., Bakopoulou, I., Law, J., Spencer, S., & Lindsay, G. (2012). *Developing a communication supporting classrooms observation tool*. Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/219634/DFE-RR247-BCRP8.pdf

Dockrell, J. E., Howell, P., Leung, D., & Fugard, A. J. B. (2017). Children with speech language and communication needs in England: Challenges for practice. *Frontiers in Education*, *2*, 1–14. https://doi.org/10.3389/feduc.2017.00035

Dockrell, J. E., & Lindsay, G. (2001). Children with specific speech and language difficulties: The teachers' perspective. *Oxford Review of Education*, *27*(3), 369–394. https://doi.org/10.1080/03054980120067410

Ebbels, S., McCartney, E., Slonims, V., Dockrell, J., & Norbury, C. F. (2019). Evidence based pathways to intervention for children with language disorders. *International Journal of Language and Communication Disorders*, *54*(1), 3–19. https://doi.org/10.1111/1460-6984.12387

Ehren, B. J., & Nelson, N. W. (2005). The responsiveness to intervention approach and language impairment. *Topics in Language Disorders*, *25*(2), 120–131. https://doi.org/10.1097/00011363-200504000-00005

EPPI-Centre. (2003). Review guidelines for extracting data and quality assessing primary studies in educational research, version 0.9.7. https://eppi.ioe.ac.uk/cms/Resources/Tools/tabid/184/Default.aspx

*Ergül, C., Akoglu, G., Sarica Kesiktaş, A., Karaman, G., Tufan, M., Bahap Kudret, Z., & Deniz, K. (2016). An adapted dialogic reading program for Turkish kindergarteners from low socio-economic backgrounds. *Journal of Education and Training Studies*, *4*(7), 169–184. https://doi.org/10.11114/jets.v4i7.1565

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, *41*(1), 93–99. https://doi.org/10.1598/rrq.41.1.4

Gascoigne, M. (2006). *Supporting children with speech, language and communication needs within integrated children's services: Position paper*. Royal College of Speech and Language Therapists. https://www.mgaconsulting.org.uk/Supporting_children-website.pdf

*Gillam, S. L., Olszewski, A., Fargo, J., & Gillam, R. B. (2014). Classroom-based narrative and vocabulary instruction: Results of an early-stage, nonrandomized comparison study. *Language, Speech & Hearing Services in Schools*, *45*(3), 204–219. https://doi.org/10.1044/2014_LSHSS-13-0008

Girolametto, L., Weitzman, E., & Greenberg, J. (2003). Training day care staff to facilitate children's language. *American Journal of Speech-Language Pathology*, *12*(3), 299–311.

Goldstein, H., Kelley, E., Greenwood, C., McCune, L., Carta, J., Atwater, J., Guerrero, G., McCarthy, T., Schneider, N., & Spencer, T. (2016). Embedded instruction improves vocabulary learning during automated storybook reading among high-risk preschoolers. *Journal of Speech, Language, and Hearing Research*, *59*(3), 484–500. https://doi.org/10.1044/2015_jslhr-l-15-0227

*Gonzalez, J. E., Pollard-Durodola, S., Simmons, D. C., Taylor, A. B., Davis, M. J., Kim, M., & Simmons, L. (2010). Developing low-income preschoolers' social studies and science vocabulary knowledge through content-focused shared book reading. *Journal of Research on Educational Effectiveness*, *4*(1), 25–52. https://doi.org/10.1080/19345747.2010.487927

*Goodson, B., Wolf, A., Bell, S., Turner, H., & Finney, P. B. (2010). *The Effectiveness of a program to accelerate vocabulary development in kindergarten (VOCAB): Kindergarten final evaluation report*. Institute of Education. Sciences. https://ies.ed.gov/ncee/edlabs/regions/southeast/pdf/REL_20104014.pdf

Greenberg, M. T., & Abenavoli, R. (2017). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness*, *10*(1), 40–67. https://doi.org/10.1080/19345747.2016.1246632

Greenwood, C. R., Carta, J. J., Walker, D., Watson-Thompson, J., Gilkerson, J., Larson, A. L., & Schnitz, A. (2017). Conceptualizing a public health prevention intervention for bridging the 30 million word gap. *Clinical Child and Family Psychology Review*, *20*(1), 3–24. https://doi.org/10.1007/s10567-017-0223-8

Greenwood, C. R., Schnitz, A. G., Carta, J. J., Wallisch, A., & Irvin, D. W. (2020). A systematic review of language intervention research with low-income families: A word gap prevention perspective. *Early Childhood Research Quarterly*, *50*(1), 230–245. https://doi.org/10.1016/j.ecresq.2019.04.001

*Hadley, P. A., Simmerman, A., Long, M., & Luna, M. (2000). Facilitating language development for inner-city children: Experimental evaluation of a collaborative, classroom-based intervention. *Language, Speech, and Hearing Services in Schools*, *31*(3), 280–295. https://doi.org/10.1044/0161-1461.3103.280

*Hassinger-Das, B., Ridge, K., Parker, A., Golinkoff, R., Hirsh-Pasek, K., & Dickinson, D. (2016). Building vocabulary knowledge in preschoolers through shared book reading and gameplay. *Mind, Brain, and Education*, *10*(2), 71–80. https://doi.org/10.1111/mbe.12103

Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions, Version 5.1.0*. The Cochrane Collaboration. http://handbook-5-1.cochrane.org/

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, P. K. S., & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*(7), 1071–1083. https://doi.org/10.1177/0956797615581493

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002

*Ihmeideh, F. M. (2014). The effect of electronic books on enhancing emergent literacy skills of pre-school children. *Computers and Education*, 79, 40–48. https://doi.org/10.1016/j.compedu.2014.07.008

Jones, D. (2017). Talking about talk: Reviewing oracy in English primary education. *Early Child Development and Care*, *187*(34), 498–508. https://doi.org/10.1080/03004430.2016.1211125

Justice, L. M., Jiang, H., & Strasser, K. (2018). Linguistic environment of preschool classrooms: What dimensions support children's language growth? *Early Childhood Research Quarterly*, *42*, 79–92. https://doi.org/10.1016/j.ecresq.2017.09.003

*Justice, L. M., Mashburn, A., Pence, K. L., & Wiggins, A. (2008). Experimental evaluation of a preschool language curriculum: Influence on children's expressive language skills. *Journal of Speech, Language, and Hearing Research*, *51*(4), 983–1001. https://doi.org/10.1044/1092-4388(2008/072)

Justice, L. M., McGinty, A. S., Zucker, T., Cabell, S. Q., & Piasta, S. B. (2013). Bi-directional dynamics underlie the complexity of talk in teacher–child play-based conversations in classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, *28*(3), 496–508. https://doi.org/10.1016/j.ecresq.2013.02.005

*Kaumans, D. (1972). *The effect of the Peabody Language Development Program, level 2, on oral language ability, verbal intelligence and school progress with first grade children*. (Publication Number: 1891353860, Doctoral dissertation), University of Iowa. Proquest Dissertations Publishing. https://search-proquest-com.libproxy.ucl.ac.uk/dissertations-theses/effect-peabody-language-development-program-level/docview/1891353860/se-2?accountid=14511

Kong, N. Y., Carta, J. J., & Greenwood, C. R. (2019). Studies in MTSS problem solving: Improving response to a pre-kindergarten supplemental vocabulary Intervention. *Topics in Early Childhood Special Education*. https://doi.org/10.1177/0271121419843995

Law, J. (2019a). The changing perception of communication needs – A litmus test for the Warnock legacy. *Frontiers in Education, 4*. https://doi.org/10.3389/feduc.2019.00042

Law, J. (2019b). Population woods and clinical trees. A commentary on 'Evidence-based pathways to intervention for children with language disorders'. *International Journal of Language & Communication Disorders*, *54*(1), 26–27. https://doi.org/10.1111/1460-6984.12424

Law, J., Charlton, J., Dockrell, J., Gascoigne, M., McKean, C., & Theakston, A. (2017). *Early language development: Needs, provision, and intervention for preschool children from socio-economically disadvantaged backgrounds*. Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Law_et_al_Early_Language_Development_final.pdf

Law, J., Lee, W., Roulstone, S., Wren, Y., Zeng, B., & Lindsay, G. (2012). *'What works': Interventions for children and young people with speech, language and communication needs*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/556912/DFE-RR247-BCRP10.pdf

Law, J., McBean, K., & Rush, R. (2011). Communication skills in a population of primary school-aged children raised in an area of pronounced social disadvantage. *International Journal*

of Language & Communication Disorders, *46*(6), 657–664. https://doi.org/10.1111/j.1460-6984.2011.00036.x

Law, J., Reilly, S., & Snow, P. C. (2013). Child speech, language and communication need re-examined in a public health context: A new direction for the speech and language therapy profession. *International Journal of Language & Communication Disorders*, *48*(5), 486–496. https://doi.org/10.1111/1460-6984.12027

Law, J., Todd, L., Clark, T., Mroz, M., & Carr, J. (2013). *Early Language Delays in the UK*. https://resourcecentre.savethechildren.net/sites/default/files/documents/early_language_delays.pdf

Law, J., Tulip, J., Stringer, H., Cockerill, M., & Dockrell, J. (2019). Teachers observing classroom communication: An application of the Communicating Supporting Classroom Observation Tool for children aged 4–7 years. *Child Language Teaching and Therapy*, *35*(3), 203–220. https://doi.org/10.1177/0265659019869792

*Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology*, *108*(1), 1–24. https://doi.org/10.1016/j.jecp.2010.07.002

Levy, R., Yamada-Rice, D., & Marsh, J. (2013). Digital literacies in the primary classroom. In K. Hall, T. Cremin, B. Comber, & L. C. Moll (Eds.), *International handbook of research on children's literacy, learning and culture* (pp. 333–343). John Wiley & Sons.

Lindsay, G., Dockrell, J., Law, J., & Roulstone, S. (2012). *The Better Communication Research Programme: Improving provision for children and young people with speech, language and communication needs*. *Department for Education*. https://www.bettercommunication.org.uk/BCRP/DFE-RR247-BCRP1.pdf

Longfield, A. (2019). *We need to talk: Access to speech and language therapy.* Children's Commissioner. https://www.thechildrenscommissioner.gov.uk/wp-content/upoloads/2019/06/cco-we-need-to-talk-june-2019.pdf

Lonigan, C. J., & Milburn, T. F. (2017). Identifying the dimensionality of oral language skills of children with typical development in preschool through fifth grade. *Journal of Speech, Language, and Hearing Research*, *60*(8), 2185–2198. https://doi.org/10.1044/2017_jslhr-l-15-0402

Marulis, L. M., & Neuman, S. B. (2010). The Effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, *80*(3), 300–335. https://doi.org/10.3102/0034654310377087

*McCabe, A., Boccia, J., Bennett, M. B., Lyman, N., & Hagen, R. (2010). Improving oral language and literacy skills in preschool children from disadvantaged backgrounds: Remembering, writing, reading (RWR). *Imagination, Cognition and Personality*, *29*(4), 363–390. https://doi.org/10.2190/IC.29.4.f

Melby-Lervåg, M., Hagen, Å. M., & Lervåg, A. (2020). Disentangling the far transfer of language comprehension gains using latent mediation models. *Developmental Science*, *23*(4). https://doi.org/10.1111/desc.12929

Meschi, E., Micklewright, J., Vignoles, A., & Lindsay, G. (2012). *The transitions between categories of special educational needs of pupils with Speech, Language and Communication Needs (SLCN) and Autism Spectrum Disorder (ASD) as they progress through the education system*. Department for Education. https://dera.ioe.ac.uk/16320/1/DFE-RR247-BCRP11.pdf

Moats, L. C. (2010). *Speech to print: Language essentials for teachers* (2nd ed.). Paul H. Brookes.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *British Medical Journal*, *339*(1), b2535. https://doi.org/10.1136/bmj.b2535

Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, *79*(2), 979–1007. https://doi.org/10.3102/0034654309332561

Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, *11*(3), 247–266. https://doi.org/10.1016/0272-7358(91)90103-2

National Early Literacy Panel. (2010). *Developing early literacy report: A scientific synthesis of early literacy development and implications for intervention*. https://lincs.ed.gov/publications/pdf/NELPReport09.pdf

*Neuman, S. B., & Dwyer, J. (2011). Developing vocabulary and conceptual knowledge for low-income preschoolers: A design experiment. *Journal of Literacy Research*, *43*(2), 103–129. https://doi.org/10.1177/1086296X11403089

*Neuman, S. B., & Kaefer, T. (2018). Developing low-income children's vocabulary and content knowledge through a shared book reading program. *Contemporary Educational Psychology*, *52*, 15–24. https://doi.org/10.1016/j.cedpsych.2017.12.001

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*, 1247–1257.

O'Connor, M., Arnott, W., McIntosh, B., & Dodd, B. (2009). Phonological awareness and language intervention in preschoolers from low socio-economic backgrounds: A longitudinal investigation. *British Journal of Developmental Psychology*, *27*(4), 767–782. https://doi.org/10.1348/026151008X372492

*Okyay, O., & Kandir, A. (2017). Impact of the interactive story reading method on receptive and expressive language vocabulary of children. *European Journal of Educational Research*, *6*(3), 395–406. https://doi.org/10.12973/eu-jer.6.3.395

*Opel, A., Ameer, S. S., & Aboud, F. E. (2009). The effect of preschool dialogic reading on vocabulary among rural Bangladeshi children. *International Journal of Educational Research*, *48*(1), 12–20. https://doi.org/10.1016/j.ijer.2009.02.008

Oracy All-Party Parliamentary Group. (2020). *Speak for change: Initial findings and recommendations from the Oracy All-Party Parliamentary Group Inquiry*. All Party Parliamentary Group. https://d5119182-bdac-43d5-be55-e817e7736e5b.filesusr.com/ugd/2c80ff_33e3208ce4dd4764b154682488c53ef7.pdf

Petersen, D. B., Gillam, S. L., Spencer, T., & Gillam, R. B. (2010). The effects of literate narrative intervention on children with neurologically based language impairments: An early stage study. *Journal of Speech, Language, and Hearing Research*, *53*(4), 961–981. https://doi.org/10.1044/1092-4388(2009/09-0001)

Piasta, S. B., Justice, L. M., Cabell, S. Q., Wiggins, A. K., Turnbull, K. P., & Curenton, S. M. (2012). Impact of professional development on preschool teachers' conversational responsivity and children's linguistic productivity and complexity. *Early Childhood Research Quarterly*, *27*(3), 387–400. https://doi.org/10.1016/j.ecresq.2012.01.001

Roulstone, S., Law, J., Rush, R., Clegg, J., & Peters, T. (2011). *Investigating the role of language in children's early educational outcomes*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/181549/DFE-RR134.pdf

Roulstone, S., Wren, Y., Bakopoulou, I., Goodlad, S., & Lindsay, G. (2012). *Exploring interventions for children and young people with speech, language and communication needs: A study of practice*. Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/219627/DFE-RR247-BCRP13.pdf

*Ruston, H. P., & Schwanenflugel, P. J. (2010). Effects of a conversation intervention on the expressive vocabulary development of prekindergarten children. *Language, Speech &*

Hearing Services in Schools, *41*(3), 303–313. https://doi.org/10.1044/0161-1461(2009/08-0100)

Savage, R., Kozakewich, M., Genesee, R., Erdos, C., & Haigh, C. (2017). Predicting writing development in dual language instructional contexts: Exploring cross-linguistic relationships. *Developmental Science*, *20*(1). https://doi-org.libproxy.ucl.ac.uk/10.1111/desc.12406

Schmitt, M. B., Logan, J. A. R., Tambyraja, S. R., Farquharson, K., & Justice, L. M. (2017). Establishing language benchmarks for children with typically developing language and children with language impairment. *Journal of Speech, Language, and Hearing Research*, *60*(2), 364–378. https://doi.org/10.1044/2016_JSLHR-L-15-0273

Schoon, I., Parsons, S., Rush, R., & Law, J. (2010). Childhood language skills and adult literacy: A 29-year follow-up study. *Pediatrics*, *125*(3), e459–e466. https://doi.org/10.1542/peds.2008-2111

Shiel, G., Cregan, Á., McGough, A., & Archer, P. (2012). *Oral language in early childhood and primary education (3-8 years)*. National Council for Curriculum and Assessment. https://ncca.ie/media/2149/oral_language_in_early_childhood_and_primary_education_3-8_years_.pdf

*Silverman, R. (2007). A comparison of three methods of vocabulary instruction during read-alouds in kindergarten. *The Elementary School Journal*, *108*(2), 97–113. https://doi.org/10.1086/525549

*Silverman, R. (2013). Investigating video as a means to promote vocabulary for at-risk children. *Contemporary Educational Psychology*, *38*(3), 170–179. https://doi.org/10.1016/j/cedpsych.2013.03.001

*Simsek, Z. C., & Erdogan, N. I. (2015). Effects of the dialogic and traditional reading techniques on children's language development. *Procedia – Social and Behavioral Sciences*, *197*, 754–758. https://doi.org/10.1016/j.sbspro.2015.07.172

Snow, P. C. (2016). Elizabeth Usher Memorial Lecture: Language is literacy is language – Positioning speech-language pathology in education policy, practice, paradigms and polemics. *International Journal of Speech Language Pathology*, *18*(3), 216–228. https://doi.org/10.3109/17549507.2015.1112837

*Snow, P. C., Eadie, P. A., Connell, J., Dalheim, B., McCusker, H. J., & Munro, J. K. (2013). Oral language supports early literacy: A pilot cluster randomized trial in disadvantaged schools. *International Journal of Speech Language Pathology*, *16*(5), 495–506. https://doi.org/10.3109/17549507.2013.845691

Sonmez, D. (2010). *Supporting kindergarten students' oral language skills: A classroom intervention for literacy development* (Publication Number: 3435844, Doctoral Dissertation), Clark University. https://search.proquest.com/docview/816862501?accountid=14511

Spencer, S., Clegg, J., Stackhouse, J., & Rush, R. (2017). Contribution of spoken language and socio-economic background to adolescents' educational achievement at age 16 years. *International Journal of Language & Communication Disorders*, *52*(2), 184–196. https://doi.org/10.1111/1460-6984.12264

*Spencer, T. D., Petersen, D. B., Slocum, T. A., & Allen, M. M. (2015). Large group narrative intervention in Head Start preschools: Implications for response to intervention. *Journal of Early Childhood Research*, *13*(2), 196–217. https://doi.org/10.1177/1476718X13515419

*Stagnitti, K., Bailey, A., Hudspeth Stevenson, E., Reynolds, E., & Kidd, E. (2016). An investigation into the effect of play-based instruction on the development of play skills and oral language. *Journal of Early Childhood Research*, *14*(4), 389–406. https://doi.org/10.1177/1476718X15579741

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/rrq.21.4.1

Stoiber, K. C., & Gettinger, M. (2016). Multi-tiered systems of support and evidence-based practices. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of multi-tiered systems of support* (pp. 121–142). Springer.

Thomas, J., Brunton, J., & Graziosi, S. (2010). *EPPI reviewer 4: Software for research synthesis*. EPPI-Centre Software, Social Science Research Unit, UCL Institute of Education

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245–1260. https://doi.org/10.1044/jslhr.4006.1245

Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1193–1208. https://doi.org/10.1044/1092-4388(2006/086)

Vega, V., Osman, Y., Katz, B., & Giordano, K. (2018). Hidden expectations: Do teachers use expressive language ability as a proxy for overall development? *Early Child Development and Care*, *190*(11), 1–9. https://doi.org/10.1080/03004430.2018.1546176

*Vuattoux, D., Japel, C., Dion, E., & Dupéré, V. (2013). Targeting the specific vocabulary needs of at-risk preschoolers: A randomized study of the effectiveness of an educator-implemented intervention. *Prevention Science*, *15*(2), 156–164. https://doi.org/10.1007/s11121-013-0379-5

Walker, D., Sepulveda, S. J., Hoff, E., Rowe, M. L., Schwartz, I. S., Dale, P. S., Peterson, C. A., Diamond, K., Goldin-Meadow, S., Levine, S. C., Wasik, B. H., Horm, D. M., & Bigelow, K. M. (2020). Language intervention research in early childhood care and education: A systematic survey of the literature. *Early Childhood Research Quarterly*, *50*(1), 68–85. https://doi.org/10.1016/j.ecresq.2019.02.010

Wasik, B. A., & Hindman, A. H. (2020). Increasing preschoolers' vocabulary development through a streamlined teacher professional development intervention. *Early Childhood Research Quarterly*, *50*(1), 101–113. https://doi.org/10.1016/j.ecresq.2018.11.001

Wasik, B. A., Hindman, A. H., & Snell, E. K. (2016). Book reading and vocabulary development: A systematic review. *Early Childhood Research Quarterly*, *37*(4), 39–57. https://doi.org/10.1016/j.ecresq.2016.04.003