A systematic review of intervention research examining English language and literacy development in children with English as an Additional Language (EAL)

Professor Victoria A. Murphy

Adam Unthiah

Executive Summary

Background to the report

This report is part of a larger project commissioned by three charitable groups (The Education Endowment Foundation (EEF), Unbound Philanthropy and The Bell Foundation) to produce a comprehensive review of the National Pupil Database (NPD) and a systematic review on relevant research concerning pupils with English as an Additional Language (EAL). The broad aim of this overall report is to provide a helpful resource for teachers and schools in providing effective support for children with EAL. The first part of this project was carried about by Strand, Malmberg and Hall and reports on a detailed analysis of the National Pupil Database (NPD) and the Longitudinal Study of Young People in England (LSYPE) database with a view to identifying the most at-risk groups of learners with EAL and predictors of low attainment for EAL pupils. In this second part of the overall project we report on the results of a systematic review of the research literature identifying what interventions have been carried out aimed at enhancing the English language and/or literacy skills of children with EAL. The main questions addressed by this second report were:

> • What intervention research has been carried out since the year 2000 which has aimed at improving English language and/or literacy skills in children with EAL?

> • What is the strength of evidence of this research?

Furthermore, this review aims to identify which interventions might be most appropriate to implement in the UK context to better support developing language, literacy, and in turn academic performance, of children with EAL.

Methods used in the review

A number of databases were scanned with key search terms to identify which studies should be reviewed. These databases include the AEI (*Australian Education Index*), BEI (*British Education Index*), LLBA (*Linguistics and Language Behavior Abstracts*), ERIC (*Education Resources Information Center*) and *Web of Science*. These database searches were also supplemented by hand-searching of a range of scholarly journals known to be publications in which such research would appear (e.g., *Reading and Writing; TESOL Quarterly, Elementary School Journal,*

International Journal of Bilingual Education and Bilingualism, Bilingualism Research Journal, Early Education and Development).   This search yielded 975 distinct reports whose abstracts were cross-referenced against inclusion and exclusion criteria to assess their relevance for the specific aims of this review. A number of studies were eliminated because they failed to meet some of the inclusion criteria (e.g., lack of appropriate control group, did not focus on an aspect of EAL provision, did not include English language/literacy measures, targeted children with language and/or learning impairments). Having applied the inclusion and exclusion criteria, 302 articles remained for full-text access. Of these 302 articles, a further 258 were excluded due to failure to meet the inclusion/exclusion criteria. This procedure left 44 studies appropriate for review and which were included in the systematic key word map. A further 15 studies were excluded after having applied the in-depth review criterion that only studies carried out in mainstream education (i.e., not bilingual programmes/education) would be included. This process yielded a total of 29 studies to be included in the in-depth review.

Each study for the in-depth review was then assessed by two independent reviewers, both of whom were experienced researchers in applied linguistics. Each reviewer assigned a rating for four Weight of Evidence (WoE) criteria addressing the overall quality of the research, the relevance of the research for the purposes of the review, the focus of evidence and an overall judgment. Studies were further assessed across four different methodological criteria relating to the study's research design, sample size, level of participant attrition, and fidelity and validity.

Key findings

1. Of the 29 in-depth review studies, only two were carried out in contexts outside of the US (one in the UK and one in Canada). Due to the considerable differences in the demographic, social, and educational infrastructure between the UK and US, it is unwise to assume effective interventions in the US would be equally effective in the UK. This lack of research across different English-speaking nations speaks to the urgent need for carefully conducted intervention studies which examine best practice aimed to improve EAL students' English language and/or literacy within the UK context.

2. The interventions carried out in the review were targeted specifically at instructional activities aimed to improve English language (n=12) and/or literacy skills (n=10), continuing professional development (CPD) activities (n=5) or family literacy practice (n=2). While it is not surprising that 22 of the 29 studies directly involved targeted pedagogical support for English language and/or literacy, it is perhaps somewhat surprising that only 5 studies had a primary focus on CPD and even fewer were directly aimed at enhancing family literacy

practice. The lack of CPD interventions is of particular concern in the UK context given Andrews' (2009) review which identified a significant lack of EAL pedagogy and too much overlap between Special Educational Needs (SEN) provision and the teaching of pupils with EAL. Therefore, more intervention research needs to be carried out to examine CPD that has a direct and effective impact on EAL children's English language and/or literacy outcomes, both internationally and in the UK.

3. Most of the studies in the in-depth review were aimed at primary school pupils, with fewer directed towards early secondary and only one specifically targeted for late secondary pupils. This asymmetry in terms of research on secondary pupils is potentially problematic since it is at this stage when pupils have high stakes formal examinations and need to have exceedingly well-developed English language and/or literacy skills (including academic language) to achieve higher results on national examinations.

4. The interventions that were targeted specifically on language and/or literacy were primarily focussed on enhancing some aspect of vocabulary. Academic vocabulary and word analysis strategies figured prominently in a number of interventions – where explicit teaching of vocabulary was developed through text-based activities and where word analysis strategies that were aimed to enhance phonological and morphological awareness[1] improved students' abilities to understand the relationships between wood roots and derived forms.

5. Interventions that were developed to support struggling readers primarily aimed to improve alphabetic knowledge through phonics training and phonological awareness.

6. While none of the interventions received uniformly high ratings on Weight of Evidence criteria, effect sizes, and/or other methodological criteria, there are a number of interventions aimed at enhancing vocabulary knowledge in children with EAL that could be appropriate for implementation in the UK.

Detailed findings by intervention focus

Language oriented interventions

The significant majority (10/12) of the interventions that were aimed primarily at some aspect of language were focused on developing vocabulary knowledge. Within these interventions, the target was either on academic vocabulary, phonological and/or morphological awareness, or

---

[1] Phonological awareness refers to an ability to identify the phonological characteristics of a word as distinct from the meaning. Morphological awareness refers to the ability to recognize, understand and use different meaningful word parts (i.e., understanding that adding the suffix [-er] on to the verb 'teach' (teach<u>er</u>) changes the word to refer to the agent of the verb).

general vocabulary knowledge. The general strategy across these interventions was to directly teach specific vocabulary through text-based activities (shared book reading, reading tasks, and the like). These interventions varied in terms of the strength of evidence criteria and effect sizes, but there are a number that were rated at appropriately high levels as to suggest it would be worth examining their effectiveness in UK schools.   Those language interventions which were not focused primarily on vocabulary aimed to enhance verbal interaction in classrooms and/or auditory-perceptual and spoken language skills.

Literacy oriented interventions

The interventions primarily focused on literacy either were comprehension oriented and aimed at enhancing listening comprehension or comprehension of specific vocabulary, or aimed to develop lower-level reading skills such as single word reading, decoding, fluency, and phonological awareness.   A number of these literacy-oriented interventions included children who were identified as struggling readers in that they were at the lower end on pre-test measures of reading performance.   The overall ratings for these interventions were mixed, but many of them had appropriately high ratings on strength of evidence criteria to suggest there may be some merit in examining the effectiveness of these interventions in UK classrooms.

Continuing Professional Development interventions

There were five interventions with a primary focus on continuing professional development (CPD) activities. Two of these aimed to support teachers to improve literacy outcomes through teaching academic content (e.g., science).   One was specifically targeted at helping teachers improve reading comprehension and reading achievement in urban areas where there was high teacher turnover (mobility) and one based in the UK was specifically oriented towards helping teachers promote better verbal interaction in classrooms. These CPD interventions tended to receive lower ratings than the language and/or literacy oriented interventions in terms of the strength of evidence. It is somewhat surprising that there were so few CPD interventions – though many of the language/literacy interventions did include a CPD component even if it was not the primary focus. The role of the teacher is critical in supporting children with EAL so it would be useful if more research in the UK could empirically examine different CPD programmes aimed to support EAL children's language and/or literacy outcomes.

Family literacy practice interventions

Only two interventions were identified that aimed to enhance literacy practice in the home. One of these implemented a structured set of parent-child activities to help parents in supporting their child's literacy development outside of school. This intervention was effective for the children with EAL but not the non-EAL children in the study. The other family-oriented intervention aimed to promote literacy across the summer break between school years to mitigate against observed dips in children's reading performance across this time. The intervention was not effective, however.

Recommendations for further research

Only one intervention study in this review was carried out in the UK context. The lack of UK-based experimental or quasi-experimental intervention studies aimed at supporting or enhancing English language and/or literacy development in children with EAL is alarming given the fact that there are such high proportions of children with EAL in UK schools. The first recommendation, therefore, is to develop an intervention research paradigm that explicitly identifies whether and to what extent particular approaches are effective for children with EAL (and ideally, also for non-EAL children).

Few of the interventions in this review proved to be ineffective, and those interventions aimed at enhancing vocabulary (academic and/or general) through text-based activities were particularly successful, both in terms of enhancing target vocabulary but in some cases also through expressive use in literacy-based activities, with promising effect sizes and strength of evidence ratings. It would be very useful, therefore, to examine through detailed research whether and to what extent such vocabulary-based interventions were effective in UK classrooms.

Some children with EAL are struggling readers – that is, they have specific difficulties in single word reading/decoding activities. For those children, interventions that enhance phonics training, phonological awareness, alphabetic knowledge and reading accuracy/fluency have been proven effective in the US context. Further research could be carried out in the UK to support struggling readers with EAL.

The majority of the studies in this review were targeted at primary-level children, and those that included secondary students tended to recruit early secondary pupils. While Strand, Malmberg and Hall (2015) identify that the achievement gap between EAL and non-EAL pupils considerably narrows by later secondary years it would nonetheless be worth examining more specifically what approaches are particularly effective in helping EAL children to make the transition from primary to secondary, and whether there were particular pedagogical approaches that are effective in helping children with EAL in the later secondary years.

Contents   Page number

List of figures   Page Number

List of tables

Introduction

In this report we present the findings of a systematic review implemented to examine intervention studies that have been carried out on the English language and/or literacy development of children with EAL – i.e., where children with a home language that is not English are taught through the medium of English. This issue is of particular importance in considering the development of children with EAL because whereas non-EAL children develop in homes whose linguistic environment mirrors that of their educational context this is not the case for children with EAL. Indeed, a number of children who are EAL can reach school age without having had consistent or sustained exposure to the majority language (English). Research has identified that many children with EAL, even with relatively well-developed oral language proficiency, are likely to have less vocabulary knowledge as non-EAL peers which in turn can have negative consequences on their English language and literacy development (Murphy, 2014). Literacy development is a particular concern since all children begin school faced with the task of learning to read, yet need to complete their primary education with reading skills that enable them to read to learn (Chall, 1983). The literacy skills that are forged in primary school go on to serve as the bedrock upon which academic learning proceeds at secondary education. Many primary school children with EAL, however, tend to struggle in listening and reading comprehension in particular (McKendry & Murphy, 2011; see Murphy, 2014 for a review). Therefore, it is critical that we have a thorough understanding of the development of language and literacy skills in this population in order to be able to develop appropriate and effective pedagogical support.

We carried out a systematic review of the research literature that reports on intervention studies on EAL children's English language and/or literacy development. A systematic review uses formal and explicit methods to describe and synthesize evidence from research. The main focus of the review is to ensure systematicity and hence eliminate, as much as possible, bias in determining which research is included for review, and to carefully examine the quality of the research that is included in the in-depth review. We followed the procedures as outlined by the EPPI Centre (Evidence for Policy and Practice, Information and Co-Ordinating Centre [http://eppi.ioe.ac.uk/cms/]). The EPPI centre has detailed guidance on effective approaches on how to carry out systematic reviews, and has published a number of systematic reviews on a range of topics related to education and social policy together with specific resource materials on best practice in carrying out a review. Consequently, the methods adopted in this review were taken from their guidelines (e.g., Gough, Oliver & Thomas, 2012).

Aims of the Review

The overarching objectives of this review were to:

- identify and review controlled intervention studies (i.e., experimental and quasi-experimental designs) which have focused on and/or included EAL pupils' English language and/or literacy development
- identify the quality of these studies with respect to their contribution to improved understanding of teaching and learning for EAL students
- identify intervention programmes which are most suited to adapting or extending in the UK context to address attainment gaps in pupils with EAL in England.
- identify and review where further research should be carried out should gaps be identified in the literature, both within the UK and internationally.

Methods for selecting studies for the in-depth review

Search Strategy

Databases used for electronic searching were: AEI (Australian Education Index), BEI (British Education Index), LLBA (Linguistics and Language Behavior Abstracts), ERIC (Education Resources Information Center) and Web of Science (including; Science Citation Index, Social Sciences Citation Index, Arts & Humanities Citation Index). Different search strategies were tested to determine which would yield the most appropriate results. The search strategies were refined in subsequent iterations by reviewing the abstracts of different articles and adding/deleting different search terms. Selective manual searching was also carried out on 10 different journals (volumes published between 2000 and May 2014) known to be likely to include relevant papers. These were: *Reading and Writing; TESOL Quarterly, Elementary School Journal, International Journal of Bilingual Education and Bilingualism, Bilingualism Research Journal, Early Education and Development, Journal of Research on Educational Effectiveness, American Educational Research Journal, Research in the Teaching of English and School Psychology Review.* Table 1 identifies which search terms were used as a result of this search strategy. Note that each row represents an 'OR' function where, for example, the terms 'intervention' OR 'treatment' OR 'program(me)' OR 'implementation' were used in conjunction with the other terms. We also used the 'NOT' facility in searching databases where we excluded the terms 'case studies' OR 'disorder', 'autism', 'ethnography'.

Table 1. Keyword Search terms

intervention	**AND**	minority language	**AND**	literacy development
treatment		heritage language		literacy acquisition
program(me)		additional		literacy skill

	language		
implementation	community language		bilingualism
	language minority		literacy
	English language learner		reading development
	ESL		reading skill
	diverse language		reading achievement
	EAL		literacy achievement
	English as a second language		receptive (language/vocabulary)
	English as an Additional language		productive (language/vocabulary)
			expressive (language/vocabulary)
			writing
			phonetic decoding
			phonetic awareness

Inclusion/Exclusion criteria

Once the initial search strategy had been carried out, we identified from abstracts (and for some papers from reading the full articles) the studies to be included in the keyword map, according to the following criteria:

1. Research that has been published in English since 2000
2. Research that examines some aspect of EAL provision
3. Interventions focused on influencing EAL pupils' English language and/or literacy outcomes targeting individuals, small-groups, classroom-based practice, school-wide practice, or parents/families
4. Interventions which include an appropriate control or comparison group.
5. Research which describes peer-reviewed empirical reports in peer-reviewed journal articles
6. Interventions on primary and secondary level pupils
7. Interventions on typically developing children

Studies were excluded if they:

1. did not adhere to either an experimental or quasi-experimental design
2. did not include a control or comparison group
3. did not include outcome measures which addressed EAL children's English language or literacy
4. focussed on children with language and/or learning disabilities
5. published before 2000

6. focused on students in post-secondary education or pre-school (under 4 years)
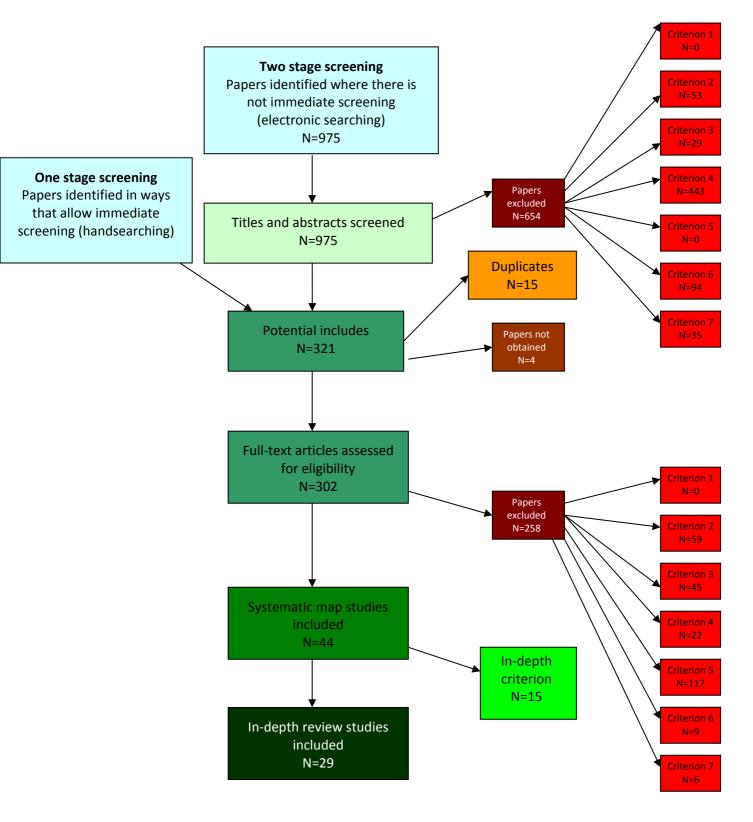
We only included studies from the year 2000 because of the fact that there has been a significant increase in the numbers of children with EAL in UK primary and secondary schools since this time. In the year 2000, 8.7% of the primary school population was EAL compared to 18.1% in 2013. For the secondary school population, 8% were EAL in 2000 compared to 13.6% in 2013 (NALDIC). We also felt that 14 years' worth of research would uncover the main themes in terms of which interventions have been developed and implemented in English-speaking countries.

Applying the criteria above yielded 44 studies for inclusion in the systematic keyword map. For the in-depth review we applied an additional criterion:

1. Interventions on children participating in standard educational provision (i.e., not bilingual programmes where some educational provision is delivered through the medium of the child's L1 and English)

We applied this additional criterion in order to ensure maximum generalizability to the minority language learner context in the UK (and indeed elsewhere internationally) where there is minimal access to bilingual education. The application of this additional criterion yielded a total of 29 studies for the in-depth review. Figure 1 depicts the protocol adopted to identify the articles for the in-depth review.

Figure 1. Systemic Review Protocol

Two stage screening
Papers identified where there is not immediate screening (electronic searching)
N=975

One stage screening
Papers identified in ways that allow immediate screening (handsearching)

Titles and abstracts screened
N=975

Papers excluded
N=654

Criterion 1
N=0

Criterion 2
N=53

Criterion 3
N=29

Criterion 4
N=443

Criterion 5
N=0

Criterion 6
N=94

Criterion 7
N=35

Potential includes
N=321

Duplicates
N=15

Papers not obtained
N=4

Full-text articles assessed for eligibility
N=302

Papers excluded
N=258

Criterion 1
N=0

Criterion 2
N=59

Criterion 3
N=45

Criterion 4
N=22

Criterion 5
N=117

Criterion 6
N=9

Criterion 7
N=6

Systematic map studies included
N=44

In-depth criterion
N=15

In-depth review studies included
N=29

As illustrated in Figure 1, many of the original 975 studies identified from our search were excluded once the abstracts (and/or titles) were cross-referenced against our inclusion and exclusion criteria. Specifically, the significant majority (n=443) were eliminated due to the lack of appropriate control or comparison group. 94 studies were excluded because they were not targeted at primary or secondary school students. These studies predominantly had pre-school students as their focus, and while English language and pre-literacy development of children with EAL in preschool is an important and interesting topic, preschool provision is not the focus of this review. 53 studies were not exclusively targeted on an aspect of EAL provision and 35 studies were focussed exclusively on children who were not typically developing – i.e., had some form of language or learning impairment. Again, the specific challenges faced by children with EAL who have learning difficulties is exceedingly important but not the focus of this review[2]. Finally, 29 studies were excluded because the outcome measures did not include English language and/or literacy assessments.

Applying the inclusion and exclusion criteria in this way left 302 articles which, based on their abstracts, looked like they may be appropriate for the review. We then accessed the full text of each article to carry out a more detailed application of the inclusion and exclusion criteria. This process led us to exclude a further 258 papers. The majority of these (n=117) were excluded because they were not in peer-reviewed journal articles or were not peer-reviewed reports. For example, a number of these were either Master's or DPhil theses and while theses do indeed undergo a process of review, it is not the same kind of rigorous review procedure as scholarly, blind, peer-reviewed articles. Some were excluded because they either did not focus on provision for EAL, English language and/or literacy outcomes and/or did not have a control group. A further 9 were excluded because they were targeted on preschool children and finally, 6 were excluded due to their focus on atypical development. Applying this screening process yielded a final 44 studies in the systematic keyword map (i.e., the process of applying the keywords presented in Table 1 to the databases and hand searching). Appendix A presents the details of each of these 44 studies.

Having identified these 44 studies, we then applied the additional criterion that only those studies which were carried out in traditional educational programmes – i.e., not bilingual education programmes - would be included in the in-depth review. Many countries around the world, and the US in particular where much of this research has been conducted, offer education programmes

---

[2] For a useful discussion addressing the relationship between EAL and Special Educational Needs see: http://www.naldic.org.uk/eal-teaching-and-learning/eal-resources/eal-sen - and the NALDIC Quarterly referenced on this site. For a helpful review of research examining the relationship between dual language development and linguistic disorders see: Paradis, Genesee and Crago (2011).

specifically developed to assist children with EAL in their English language and literacy development and where children with EAL spend a portion of the day educated through the medium of their L1 (see Genesee, Lindholm-Leary, Saunders & Christian, 2006; Murphy, 2014). The significant majority of children with EAL in the UK do not participate in such bilingual education programmes and in order to make this review as applicable as possible to the UK context only studies that were carried out in a setting that was more closely matched with the UK were included. Applying this additional criteria led to 15 studies being excluded resulting in a total of 29 studies for the in-depth review.

Characteristics of the studies included in the in-depth review

This section describes the basic characteristics of the 29 in-depth review studies. One of the most striking of which is that 27 of the 29 studies were carried out within the US as illustrated in Table 2.

Table 2: Country in which study was undertaken (N=29)

Country	Number
USA	27
Canada	1
UK	1

The US has a reasonably high proportion of its population who speak a language other than English where in the 2007 Census the overall figure was 19.7%, with some states (California, New Mexico, Texas) having more than 30% of the overall population speaking a language other than English. Consequently, educators have developed different forms of bilingual education to help support the English language (and sometimes the home language) development of minority language learners (i.e., children with EAL) (García & Kleifgen, 2010; García, 2009). The development of these programmes has contributed to an increase in research examining the needs of children with EAL and their linguistic and academic achievements (Genesee et al, 2006). Nonetheless, there is a stark contrast between the number of studies carried out in the US and other English-speaking nations which underscores the urgent need for more research on these issues in other geopolitical contexts, and the UK in particular.

Table 3 illustrates that of the 29 studies in the in-depth review, the majority of these were focused exclusively on children with EAL while other studies examined non-EAL and EAL children in the same study. Furthermore, some of the interventions in the in-depth review were targeted on teachers (i.e., through continuing professional development activities) and family literacy practices.

Table 3: Population focus (N=29, not mutually exclusive)

Primary Focus of intervention	Number
Non-EAL children	16
English as Additional Language (EAL) learners	29
Teachers	5
Families	2

The studies in the in-depth review reflect research on different age groups (and consequently levels of education). Table 4 illustrates that the majority of studies (n=14) were carried out on children in the early school years – defined as aged 4 to 6 years old which in the UK context would correspond to Early Years/Reception classes. Some studies, however, focused on more than one age band.

Table 4: Breakdown by age group (N=29, not mutually exclusive)

Age band	Number	Articles
Early school years (age 4-6)	14	1.Greenfader & Brouillette, 2013 2. Filippini, et al, 2012 3.Harper, et al, 2011 4. Vadasy & Sanders, 2010 5. Lugo-Neris, et al, 2010 6. Spycher, 2009 7. Solari, & Gerber, 2008 8. Giambo & McKinney, 2004 9. Kotler et al, 2001 10. Crevecoeur, et al, 2013 11. Vadasy, & Sanders, 2013 12. Ehri, et al, 2007 13. Kamps, et al, 2007 14. Tong, et al, 2010
Mid to late primary (aged 7-11)	8	1. Lara-Alecio, et al, 2012 2. Mancilla-Martinez, 2010 3. Matsumura, et al, 2010 4. Kim & Guryan, 2010 5. Kotler et al, 2001 6. Troia, 2004 7. Proctor, et al, 2011 8. Almaguer, 2005 9. Graves, et al, 2011 (6th grade)
Early secondary (aged 12-14)	8	1. Short, et al, 2012 2. Townsend, & Collins, 2009 3. Troia, 2004 4. Lesaux, et al, 2010 5. Snow, et al, 2009 6. Kieffer, & Lesaux, 2012 7. Lawrence, et al, 2012
Mid to late secondary (aged 15-18)	1	1. Kim, et al, 2011

As indicated above, our additional in-depth review inclusion criteria excluded studies carried out within the context of bilingual education programmes. However, a few of the studies (see Table 5)

were carried out within the context of specialist support for English language learners, which sometimes was examined in the context of structured English immersion or sheltered English instruction (SEI) (n = 6). SEI is a pedagogical approach aimed to support rapid English language development in children with EAL. In many examples of Structured English Immersion the majority of the school day is spent in English language activities and children are exited from the programme when they have reached a specific level of fluency in English. In Sheltered English Instruction teachers use clear, direct and simple English along with a range of scaffolding strategies to help make content accessible to students with limited English proficiency. While SEI is not a form of education available in the UK, it nonetheless partially reflects what does sometimes happen in the UK – i.e., all children with EAL are 'mainstreamed' in to classes delivered through the medium of English (i.e., no L1 support), but where children who have specific difficulties with English can receive remedial English Language Arts support outside of the normal classroom sessions. We therefore felt that it would be of interest to include studies which were carried out within this context for the in-depth review.

Table 5: English language support for EAL learners (N=29)

Support	Number	Article
None – mainstream classes only	23	1. Greenfader & Brouillette, 2013 2. Lara-Alecio, et al, 2012 3. Short et al, 2012 4. Filippini, et al, 2012 5. Harper, et al, 2011 6. Kim, et al , 2011 7. Mancilla-Martinez, 2010 8. Matsumura, et al, 2010 9. Lugo-Neris, et al, 2010 10. Kim & Guryan, 2010 11. Townsend, & Collins, 2009 12. Spycher, 2009 13. Solari, & Gerber, 2008 14. Giambo & McKinney, 2004 15. Kotler et al, 2001 16. Troia, 2004 17. Vadasy, & Sanders, 2013 18. Graves, et al, 2011 19. Lesaux, et al, 2010 20. Ehri, et al, 2007 21. Kamps, et al, 2007 22. Almaguer, 2005 23. Kieffer, & Lesaux, 2012
Special or remedial education (SEI) – non mainstream classes	6	1. Vadasy & Sanders, 2010 2. Crevecoeur, et al, 2013 3. Proctor, et al, 2011 4. Snow, et al, 2009 5. Lawrence, et al, 2012 6. Tong, et al, 2010

Given that the significant majority of studies in the in-depth review were carried out in the context of the US, it is not surprising to see from Table 6 that the L1 of most of the children who

participated in the studies came from Spanish-speaking homes. Studies were counted as having mixed L1s if more than one L1 was represented in the sample – a situation most common in UK classrooms. For four studies (Matsumura et al. 2010; Graves et al. 2011; Snow et al. 2009; Lawrence et al. 2012) data on the L1s of the EAL pupils were not specified. Where possible, we recorded the specific L1, meaning that there is some overlap between those studies recorded as mixed and those studies recorded for specific L1s.

Table 6. First Language of pupils (N=29, not mutually exclusive)

Language	Studies including children who speak this language
Spanish	23
English	16
Mixed L1s in the same study	12
Vietnamese	6
Somali	5
Tagalog	4
Arabic	3
Pilipino	3
Hmong	2
Chinese	2
Punjabi	2
Polish	1
Guajarati	1
Japanese	1
Turkish	1
Sylheti	1
Amharic	1
Russian	1
Samoan	1
Cambodian	1
Oromo	1
Tigrigna	1
Lao	1
Albanian	1
Sudanese	1

Types of interventions in the in-depth review studies

The in-depth review studies represent a range of different approaches with respect to enhancing English language or literacy development in children with EAL. Table 7 illustrates that the majority of the interventions were primarily focussed on enhancing some aspect of the children's English language skill and where nearly as many (n=10) targeted some aspect of comprehension and/or literacy development. However, many of these studies focussed on both, where word-level skills were developed through text-based activities. Indeed, the division between 'language' and

'literacy' interventions is somewhat arbitrary in the sense that many interventions included a strong focus on vocabulary due to the importance of developing vocabulary knowledge in developing good reading comprehension skills. Therefore, there are some interventions identified in Table 7 as being focussed on literacy because they included a literacy component, even if the study also had a strong vocabulary element. This distinction between 'language' and 'literacy' should not be considered as rigid, but rather perhaps more as ends on a continuum.

Table 7: Focus of Intervention (N=29)

Primary Focus of Intervention	Number
Language	Snow et al., 2009 Lawrence et al., 2012 Mancilla-Martinez, 2010 Kieffer & Lesaux, 2012 Lesaux et al., 2010 Troia, 2004 Filippini et al, 2012 Townsend & Collins, 2009 Spycher, 2009 Giambo & McKinney, 2004 Crevecoer et al., 2013 Greenfader & Brouillette, 2013
Literacy	Ehri et al., 2007 Proctor et al., 2011 Almaguer, 2005 Solari & Gerber, 2008 Vadasy & Sanders, 2013 Vadasy & Sanders, 2010 Tong et al., 2010 Kamps et al., 2007 Lugo-Neris et al., 2010 Graves et al., 2011
Continuing Professional Development	Lara-Alecio et al., 2012 Short et al., 2012 Kim et al, 2011 Kotler et al, 2001 Matsumura et al, 2010
Family literacy practice	Harper et al, 2011 Kim & Guryan, 2010

A range of different language-oriented features were highlighted in the predominantly language interventions, but as illustrated in Table 8, most focused on some aspect of vocabulary with two exceptions: the one study on verbal interaction (Greenfader & Brouillette, 2011) and the research that evaluated an intervention primarily targeted on auditory-perceptual and spoken skills (Troia, 2004). The majority of the remaining interventions mainly targeted some aspect of vocabulary and many of these focussed on academic vocabulary in particular (n=6).

Table 8: Primary feature(s) targeted by the 'Language' interventions (N=12)

Aspect of language	Number
Academic Vocabulary	6

Verbal Interaction	1
Vocabulary – Phonological awareness	2
Morphological awareness	2
General vocabulary	2
Auditory-perceptual and spoken language skills	1

The majority of those interventions with a primary focus on literacy development targeted reading fluency/accuracy and/or reading comprehension (see Table 9). Indeed, there was only one intervention that was aimed specifically to improve children's writing skills, and this was a continuing professional development (CPD) intervention (Kim et al., 2011). However, as with most of the interventions in the review, more than one aspect of literacy and/or language was included.

Table 9: Primary Feature(s) targeted by the 'Literacy' interventions (N=10)

Primary Aspect of literacy	Number
Shared Reading	2
Reading with Software	1
Reading Rescue (Recovery)	1
Reading comprehension	6
Phonics instruction through texts	2
Reading fluency	5
Reading – fluency (phonetic decoding, phonological awareness)	11

There were only 5 interventions that were chiefly focussed on Continuing Professional Development (CPD) though many of the 29 studies in the in-depth review included some element of CPD even if it was not the primary focus. Of those where the concentration was on CPD, two aimed at supporting teachers to integrate academic content with literacy development, and the remaining 3 aimed to improve writing, reading comprehension and verbal interaction (see Table 10).

Table 10. Primary focus of Continuing Professional Development (CPD) Interventions (N=5)

Primary Aspect of CPD	Article
Integrating literacy skills with content	Lara-Alecio et al (2010) Short et al (2012)
Use of a cognitive strategies approach to text-based analytical writing instruction	Kim et al (2011)
Content-Focused Coaching for reading comprehension and reading achievement in high teacher mobility schools	Matsumura et al (2010)
'Talking Partners' – CPD to promote better interaction in classrooms	Kotler et al (2001

The two family oriented interventions focussed on either enhancing family literacy practice or tried to promote reading during the summer holiday to mitigate against a dip in reading skill during this time (see Table 11).

Table 11. Focus of Family Literacy Interventions (N=2)

Primary focus of Intervention	Article
Parent-child activities to help parents enhance child's literacy skills	Harper et al (2011)
Summer literacy workshops to help prevent summer reading loss	Kim & Guryan (2010)

In-Depth Review of Studies

Having identified the studies for the in-depth review, we then assessed the quality of the evidence of these interventions to help identify whether there may be particular insights or applications from this research to the UK context. We first organised the interventions according to the particular linguistic/literacy feature(s) that were targeted by the intervention to determine whether there were any specific discernable patterns related to which aspects of English language/literacy were investigated in different age groups. Table 12 presents the results of this analysis.

Table 12. Aspects of language and/or literacy targeted by intervention broken down by age group

Focus of intervention	Early Primary	Mid-Late Primary	Early Secondary	Mid-Late Secondary
Word-level skills (e.g. phonological decoding/awareness, morphological awareness)	Filippini et al., 2012 Harper et al., 2011 Vadasy & Sanders, 2010 Vadasy & Sanders, 2013 Solari & Gerber, 2008 Giambo & McKinney, 2004 Troia, 2004 Ehri et al., 2007 Kamps et al., 2007 Tong et al., 2010	Graves et al., 2011 Mancilla-Martinez, 2010 Troia, 2004 Lesaux et al. 2010 Kieffer & Lesaux, 2012		
General vocabulary	Filippini et al., 2012 Lugo-Neris et al., 2010 Spycher, 2009	Mancilla-Martinez, 2010 Kim & Guryan, 2010 Kotler et al,	Snow et al., 2009 Lawrence et al., 2012	

13

	Giambo & McKinney, 2004 Kotler et al, 2001 Crevecoeur et al. 2013 Ehri et al., 2007 Harper et al., 2011 Solari & Gerber, 2008 Crevecoeur et al., 2014	Lesaux et al. 2010 Proctor et al., 2011		
Reading fluency	Vadasy & Sanders, 2010 Ehri et al., 2007 Kamps et al., 2007 Tong et al., 2010	Graves et al., 2011 Lara-Alecio et al., 2012 Mancilla-Martinez, 2010 Almaguer, 2005		
Reading comprehension	Vadasy & Sanders, 2010 Vadasy & Sanders, 2013 Solari, and Gerber, 2008 Ehri et al., 2007 Kamps et al., 2007 Tong et al., 2010	Graves et al., 2011 Mancilla-Martinez, 2010 Matsumura et al., 2010 Kim & Guryan, 2010 Lesaux et al. 2010 Almaguer, 2005 Proctor et al., 2011	Kim et al., 2011	Kim et al., 2011
Academic vocabulary	Spycher, 2009 Tong et al., 2010	Short et al., 2012 Snow et al., 2009 Kieffer & Lesaux, 2012 Mancilla-Martinez, 2010 Lesaux et al., 2010	Lawrence et al., 2010 Townsend & Collins, 2009 Short et al., 2012 Lara-Alecio et al., 2012	
Writing		Short et al., 2012 Mancilla-Martinez, 2010 Lesaux et al., 2010	Kim et al., 2011 Short et al., 2012 Lara-Alecio et al., 2012	Kim et al., 2011
Verbal interaction	Greenfader & Brouillette, 2013 Spycher, 2009 Giambo & McKinney, 2004 Kotler et al, 2001 Troia, 2004 Tong et al., 2010 Kotler, et al., 2001	Short et al., 2012 Kotler et al, 2001 Troia, 2004		
Listening comprehension	Solari & Gerber, 2008	Mancilla-Martinez, 2010		

	Greenfader & Brouillette, 2013 Troia, 2004 Kotler et al., 2001 Crevecoeur et al., 2014	Troia, 2004 Lesaux et al., 2010		
Continuing Professional Development	Greenfader & Brouillette, 2013 Tong et al., 2010 Vadasy & Sanders, 2010 Kotler et al., 2001	Short et al., 2012 Matsumura et al., 2010 Snow et al., 2009 Kim & Guryan, 2010	Kim et al., 2011 Snow et al., 2009 Short et al., 2012 Lara-Alecio et al., 2012	Kim et al., 2011

Table 12 reflects the fact that the majority of the studies focused on primary school students, and their early word-level skills, general vocabulary, and oral language skills including listening comprehension and verbal interaction. Middle primary and early secondary students have been mostly investigated in relation to improving vocabulary and reading comprehension. What is perhaps most striking from Table 12 is the fact that 27 (out of 29) studies examined some aspect of vocabulary, with 10 of these specifically examining academic vocabulary. Reading fluency and reading comprehension are also key areas targeted by these studies. This focus on word-level skills, vocabulary, and reading comprehension reflects current thinking about the importance of oral language (vocabulary) skills in supporting reading comprehension (Genesee et al., 2006; Murphy, 2014).

We adopted a two tier process in examining the quality of the evidence of these studies. In the first phase of our evaluation we assembled a reviewing team consisting of applied linguistics scholars at the University of Oxford. Each paper was read independently by two members of the review team who rated each study according to Weight of Evidence (WoE) criteria. WoE criteria enable reviewers to examine quality and relevance of individual research papers. Despite meeting our specified inclusion/exclusion criteria, individual studies may not meet quality and relevance standards (Gough, 2007). However, as a construct, the notion of 'quality' is complex and not always objectively straightforward to identify. Quality can be measured in terms of generic or intrinsic judgements or against specific criteria. Importantly, a piece of research might meet expected standards in terms of generic assessments but not be particularly relevant or useful for the specific purposes of a review. Consequently, Gough (2007) proposes a Weight of Evidence framework which includes a generic judgement (WoE A), a review-specific judgement relating to the research design (WoE B), a review-specific judgement about the focus of the evidence (WoE C) and finally an overall judgement (WoE D). These four WoE criteria are thus:

Weight of Evidence A: Taking account of all quality assessment issues, can the study findings be trusted in addressing the study aim(s)? [High, Medium, Low]

Weight of Evidence B: What is the appropriateness of the research design and analysis for addressing the aims of the specific review? [High, Medium, Low]

Weight of Evidence C: What is the relevance of particular focus of the study for addressing the aims of the review? [High, Medium, Low]

Weight of Evidence D: Taking into account the quality of execution, appropriateness of the design and relevance of focus, what is the overall weight of evidence this study provides to answer the question of this review? [High, Medium, Low]

Each reviewer applied this WoE framework to their assessment of each study. In doing so, we also asked them to consider specific features of research studies to help them make their assessment. These features relate to the rationale for carrying out the study, level of detail provided, the clarity of the research questions, the specific methodology, how participants were assigned to their respective groups, clarity of the variables, the sampling frame, sample size, appropriateness of the procedure with relevant safeguards, assessment of the analytical framework and conclusions/implications. This procedure meant that for each study there were two ratings and where there was a discrepancy of only one rank between the two independent (blind) reviewers we report the lowest rating in order to be most conservative. For example, if reviewer A ranked a paper as 'low' and reviewer B ranked the same paper as 'medium' we reported the 'low' ranking to adopt a more stringent criterion. If the two reviewers' ratings were discrepant by more than one rank, (i.e., one reviewer rated the paper as 'low' and the other rated the paper as 'high') we opted for the middle ranking 'medium' to reflect both reviewers' assessments. Clearly this procedure has some flaws in that these are subjective assessments. However, as indicated above, each reviewer was him/herself an experienced researcher in the area of applied linguistics, three of whom had themselves carried out systematic reviews of this nature. Furthermore, we adopted a second tier of review (described below) to supplement the review protocol.

Review of interventions with a primary focus on Language

Table 13 presents the reviewers' ratings for those studies whose primary focus was on some aspect of students' English language.

Table 13. Weight of Evidence assessment for studies with a primary focus on Language.

Study	WoE A Trustwothiness of evidence	WoE B Appropriateness of design for this	WoE C Relevance of Focus for this	WoE D Overall strength of evidence for

		review	review	this review
Snow et al., 2009	Low	Medium	High	Medium
Lawrence et al., 2012	High	High	High	High
Mancilla-Martinez, 2010	Medium	Medium	Medium	Medium
Kieffer & Lesaux, 2012	Medium	Medium	High	Medium
Lesaux et al., 2010	High	High	Medium	High
Filippini et al., 2012	Medium	Medium	Medium	Medium
Townsend & Collins, 2009	Medium	Medium	High	Medium
Spycher, 2009	Medium	Medium	High	Medium
Giambo & McKinney, 2004	Medium	Medium	Medium	Medium
Crevecoeur et al., 2014	Medium	Medium	High	Medium
Greenfader & Brouillette, 2013	Low	Medium	Medium	Low
Troia, 2004	Medium	Medium	High	High

Lawrence et al (2012) is a follow up study of the Snow et al (2009) study, both of which focused on a specific intervention called 'Word Generation'. Indeed, a third paper in this category, Mancilla-Martinez (2010), also examined the 'Word Generation' intervention. The Word Generation intervention incorporates research-based principles of vocabulary learning (Snow et al., 2009) where academic vocabulary is explicitly taught (5 words each session) and then included in different activities throughout a week across different curricular topics (e.g., maths, social science and science). Therefore, it incorporates many opportunities to use the target words in different ways (texts, debates, discussion, writing, etc). In this way, the intervention helps promote oral vocabulary as well as targeting the development of key specific academic vocabulary. In the original Snow et al (2009) report, students in the Word Generation intervention learned more target academic vocabulary words than students in control schools and where children with EAL were identified to benefit more from participating in the intervention than non-EAL children. In the follow up Lawrence et al (2012) study, the long and short-term effects of participating in the Word Generation programme were compared for 3 groups of students i) proficient English speakers from non-EAL homes, ii) proficient English speakers from EAL homes and iii) limited English proficient students. The specific focus was to determine if participation in Word Generation benefited all students irrespective of home language status and proficiency, and whether all students maintained knowledge of target words relative to comparison groups. English proficient students from EAL homes made strong gains and maintained them compared to control students even a year later. English proficient students from non-EAL homes also made gains relative to the comparison group and maintained these gains across the course of the study. Limited English proficient students, however, did not show short-term or long-term benefits from the Word

Generation Program. The third paper in this category that reported on the Word Generation intervention was Mancilla-Martinez (2010) who investigated whether the Word Generation programme would improve fifth grade (age 10-11) EAL students' literacy outcomes. Mancilla-Martinez (2010) reported that the Word Generation programme helped students gain knowledge of more (target) vocabulary than children in the control group and they also had heightened word awareness. This finding replicates the findings of the Snow et al (2009) and Lawrence et al (2012) studies. Perhaps most importantly, students in the Word Generation programme in Mancilla-Martinez' (2010) study also used more of the target words in their writing suggesting possible transfer of this (academic) vocabulary intervention on to literacy skills. Each of the three papers in this category depicted in Table 13 report facilitative effects of participating in the Word Generation programme. However, of these three, only the Lawrence et al (2012) received an overall rating of 'high' perhaps due in part to methodological variability across the three studies, and/or reviewer subjectivity.

Four other papers rated in Table 13 also reported on interventions aimed at enhancing academic vocabulary. Kieffer and Lesaux (2012) and Lesaux et al (2010) report on an intervention called 'Academic Language Instruction for All Students' (ALIAS) which focused on building word-level skills of key target academic vocabulary, including relational and syntactic aspects of morphological awareness (e.g., understanding that 'complex' is the root of the word 'complexity'). Both of these studies report facilitative effects on 6th graders' word knowledge having participated in the ALIAS intervention. Townsend and Collins (2009) and Spycher (2009) each respectively report on similar interventions in that they also include direct, explicit vocabulary instruction of target academic vocabulary. Townsend and Collins' (2009) intervention was carried out on early secondary (12 year old) children with EAL through book reading activities, and Spycher's (2009) aimed at improving kindergarten children's academic language through comparing implicit vs. intentional approaches to vocabulary learning through science instruction.

Seven out of the twelve studies identified in Table 13 report on interventions targeting academic vocabulary. All of these interventions report facilitative effects for those children who participated in the interventions on developing their academic vocabulary, general vocabulary, and word awareness skills. However, of these, only Lawrence et al (2012) and Lesaux et al (2010) were ranked as 'high' overall.

The remaining interventions identified in Table 13 focused on direct, explicit instruction of vocabulary through story reading (Crevecoeur et al., 20140), or direct instruction of vocabulary targeting word-level skills such as phonological and morphological awareness (Filippini et al 2012;

Giambo & McKinney, 2004).   Troia (2004) is the only example of a software-based intervention aimed to develop aspects of EAL children's oral interaction and received a high rating on the WoE criteria due to its relevance and methodological rigour.   Unfortunately, it is one of the few studies in this in-depth review which reports no facilitative effects of the intervention. Finally, Greenfader and Brouillette (2013) report on an intervention where drama activities were introduced aimed at developing EAL children's oral interaction.   Unfortunately this study suffered from a number of methodological flaws, where many key features of the design and analyses were not explained in sufficient detail leading to its comparatively low WoE ratings.

In addition to the WoE criteria we examined each study for whether an effect size was reported, for which outcome measures, and the strength of the effect.   In some cases, a range of effect sizes is provided when the authors reported effect size by schools within treatment clusters. The strength of an effect is somewhat determined by which specific statistic the researchers calculated to determine the effect size (e.g., Cohen's d vs. η^2 (eta squared), etc.).   Many studies tend to report Cohen's d as an effect size, therefore, unless otherwise specified, Table 14 reports Cohen's d. An indication of the general strength of the effect is also reported. It is important to note that many studies report a wide range of effect sizes for both non-significant and significant findings, across a range of measures, groups, and conditions.   For the sake of brevity, in Table 14 we report either the range of effect sizes across different measures, groups, and conditions OR we present a sample of effect sizes for the main comparison of the intervention.   Some of the interventions carried out complex hierarchical multilevel modeling which is an appropriate analytical strategy for many designs within educational development paradigms such as the ones under discussion in this report and consequently did not report an effect size.   Effect size reporting in multilevel modeling is more complex than calculating Cohen's d or reporting (partial) η^2 and there is much less consensus on how effect sizes should reported in these statistical approaches (Peugh, 2010).   While there are ways to report effect sizes within the multilevel modeling paradigm these were not always reported.

Table 14. Treatment effect sizes on outcomes measures for children with EAL in interventions on language

Study	Effect Size Reported?	Outcome Measure and Effect Size	Strength of Evidence
Snow et al., 2009	Yes	Across academic vocabulary the range = 0.33-0.65	Medium-High
Lawrence et al., 2012	No	Used multi-level growth modeling	
Mancilla-Martinez, 2010	Yes	Word Generation multiple choice = 1.24 Word Generation real words self-check = 0.58	High Medium

		Word Generation nonsense words self-check = 0.80	High
Kieffer & Lesaux, 2012	Yes	Real Word Decomposition = 0.16 NonWord Derivation = 0.25	Low Medium
Lesaux et al., 2010	Yes (EAL-non-EAL difference)	Target Word Mastery = 0.51 Morphological Decomposition = 0.41 Target Word Association = 0.60 Word Meanings in Context = 0.39 Gates-MacGinitie = 0.42 SAT-10 Reading Vocabulary = 0.52	Medium Medium Medium-High Low Medium Medium
Filippini et al., 2012	Yes	Vocabulary (MA Group) = 0.56 Vocabulary (SR Group) = 0.28 Vocabulary (PA Group) = 0.04 Phonological Decoding (MA Group) = 0.33 Phon Decoding (SR Group) = 0.49 Phon. Decoding (PA Group) = .17	Medium Low Low Medium Medium Low
Townsend & Collins, 2009	Yes(partial η^2)	Academic Vocabulary = 0.15-0.36	Medium-High
Spycher, 2009	Yes	Science vocabulary = 0.98	High
Giambo & McKinney, 2004	Yes	Receptive vocabulary (PA Group) = 0.74 Receptive vocabulary (story reading group) = 0.59 Oral Proficiency (PA Group) = 0.58	High
Crevecoeur et al., 2014	Yes	Target word knowledge = 1.08 PPVT (receptive vocab) = 0.29 Listening Comprehension = -0.05	High Low Low
Greenfader & Brouillette, 2013	Yes (R^2)	Listening = 0.43 Speaking = 0.47	Medium Medium
Troia, 2004	Yes	Letter-Word Identification = 0.11 Word Attack = 0.01	Low Low

Table 14 illustrates that there was a range of magnitudes with respect to effect sizes of statistically significant findings from studies focused on language. However, the only study which reported uniformly low effect sizes was Troia (2004) who also reported very few significant effects for the treatment. Furthermore, Crevecoeur et al (2014) reported high effect sizes for their outcome measure on target word knowledge despite lower effect sizes on other measures. Across the studies in the category, the effect size most commonly reported ranges from medium to medium-high. These findings are encouraging with respect to the extent to which these interventions might be applicable to other contexts such as the UK.

As a result of the problem of inherent subjectivity of the reviewers' ratings, as well as the need to specifically evaluate the methodologies of the studies to assess their impact, we carried out a second phase of reviewing where we applied the criteria identified in Table 15 to each study. This

additional, second tier of review was carried out in an effort to be as rigorous as possible with respect to assessing the overall quality of the studies. The criteria in Table 15 do overlap somewhat with the WoE criteria, however, they also provide a more explicit rating system for the key methodological features under consideration in this review.

Table 15. Criteria for rating the strength of evidence provided in an individual study (does not take into account the size of the effect)

Strength of evidence	Explanation	1.Research design	2. Number of cases	3. Attrition	4. Outcome measure	5. Fidelity & validity
High	Findings are highly secure and makes a substantial contribution to the existing evidence	Fair and clear experimental design (RCT)	A good number of cases or clusters, a 'well-powered study' (e.g. 100 cases or 50 clusters or more per arm)	Balanced groups with minimal attrition	Robust, valid outcome measure, standardised or widely acceptable, not intervention-specific, delivered blind	Clearly defined intervention, no evidence of threats to validity, or experimenter effect (independent evaluation)
Medium	Findings are moderately secure and makes a contribution to the existing evidence	Well-matched comparison group (quasi-experiment)	A medium number of cases (e.g. 50 cases or 20 clusters per arm)	Some imbalance or moderate attrition	Robust, valid outcome measure	Reasonably clear intervention, some threats to validity or concerns about experimenter effect.
Low	Findings are insecure and add little to the existing evidence	Comparison group with poor or no matching (eg. Volunteer v other)	A small study or 'low-powered study' (e.g. less than 40 cases or 10 clusters per arm)	Substantial imbalance or high attrition	Concerns about validity, reliability and that the outcome measure is inherent to the treatment	Poorly specified intervention, serious threats to validity and strong indication of experimenter effect (lack of independence)

These criteria were applied to the twelve studies with a focus on language to yield the following ratings presented in Table 16.

Table 16. Methodological strength of evidence criteria for interventions primarily focussed on Language

Papers	Research Design	Number of Cases	Attrition	Outcome Measure	Fidelity & Validity
Snow et al., 2009	Medium	High	Low	Medium-Low	Medium
Lawrence et al., 2012	Medium	High	Low	Medium-Low	Medium
Mancilla-Martinez, 2010	Medium	Low	Low	Medium-Low	Medium
Kieffer & Lesaux, 2012	Medium-Low	High	Low	High	Medium-High
Lesaux et al., 2010	Medium-Low	High	Medium-Low	Medium-high	Medium-High
Filippini et al.,	Medium	Low	Medium	Medium-high	Medium-High

2012					
Townsend & Collins, 2009	Medium-Low	Low	Medium-Low	High	Medium
Spycher, 2009	Medium	Low	Low	Low	Medium
Giambo & McKinney, 2004	Medium	Low	Low	High	Medium
Crevecoeur et al., 2014	Medium-Low	Low	Low	Low	Medium-High
Greenfader & Brouillette, 2013	Low	Medium-High	Low	Medium-High	Low
Troia, 2004	Medium	Medium	Low	High	Medium-High

Similar to the results of the application of the WoE criteria in Table 13, Table 16 presents a rather mixed picture with respect to the strength of the evidence from the intervention studies with a primary focus on features of English language. It is worth noting, however, that some of the differences between Table 13 and 16 are reflected in the criteria used to make the ratings. The criteria in Table 15 preclude rating a study as 'high' for research design unless it is a pure randomized control trial (RCT) which is relatively rare in educational studies of the type under review here. A pure RCT is one where participants are randomly assigned to groups in the design and where different groups are carefully matched on key variables. Ideally, participants would also be randomly sampled from the population. None of the interventions could be conceived of as a pure RCT either due to lack of (or weak) matching between comparison groups, or due to a lack of random allocation to treatment conditions or both. Consequently, none of them achieved a 'high' rating in the research design category. However, many studies did randomly assign to groups and made significant efforts to match the participants on relevant variables hence the almost uniform 'medium' rating on this category. Similarly, in order to receive a rating of 'high' on the 'Fidelity and Validity' criterion, the intervention needed to be evaluated (at post test) by an independent group of researchers. This was not the case for any of the studies in this review, despite the fact that many of them included a number of appropriate and relevant methodological controls to ensure fidelity to treatment and to eliminate researcher bias as much as possible. This criterion of having a different (i.e., independent) set of evaluators for the intervention to assess validity was not included in the WoE assessment rankings, hence the discrepancy across these ratings.

Interestingly, very few of the studies had well-balanced groups and reported careful and detailed information about participant attrition, leading to an almost uniform 'low' rating on this category. There is quite a lot of variability in terms of the numbers of participants with some studies having a robust number suggesting a well-powered study (Snow et al., 2009; Lawrence et al., 2012; Kieffer & Lesaux, 2012; Lesaux et al., 2010; Greenfader & Brouillette, 2013) whereas others had a very low number of participants in each group. Similarly, most studies used some combination of

researcher-designed or adapted outcome measures which resulted in a 'low' rating in addition to well-validated standardised assessments, resulting in a higher rating. On a positive note, there was only one study in this category that could be considered a poorly specified intervention with serious threats to the internal validity (Greenfader & Brouilette, 2013) and this low rating could be due in large part to the way in which the intervention is described in this paper. The focus in this study is on the use of dramatization in the classroom with very little detail on the methodological features of the study itself, hence it was not possible to assign higher ratings to this study. Furthermore, a high rating does not mean that the intervention is successful. Troia (2004), for example, is an example of a reasonably carefully designed study aimed at evaluating a specific software programme that has been widely acclaimed to be successful in improving language outcomes in EAL and non-EAL children. Troia's intervention, however, suggests otherwise in that there were very few significant differences between the treatment and control participants in his study.

This section has reviewed the twelve papers that had a primary focus on improving some aspect of EAL children's English language. Most of these focused on some aspect of vocabulary, mostly academic vocabulary and word-level skills/knowledge such as phonological and morphological awareness, and alphabetic knowledge. Many of these interventions were carried out through text-based activities so it is worth reiterating that it is a somewhat arbitrary categorisation to argue they focussed predominantly on language – except that the chief focus of the tasks in which the participants engaged in the intervention were more on vocabulary and word-level activities relative to literacy. Tables 13 and 16 together show a somewhat mixed set of reviews with respect to both the WoE criteria and the methodological ratings in Table 16. However, few of these interventions could be considered to be of very poor quality, and with the exception of Troia (2004) show some facilitative effects of participation in the treatment conditions. These findings suggest that there is significant scope for considering implementing interventions like these in the UK context.

Review of interventions with a primary focus on Literacy

In this section, we report on the review of the studies with a primary focus on literacy – though note again that this does not preclude language oriented tasks, rather that the main focus of the interventions was weighted more on developing literacy and/or comprehension skills. Table 17 illustrates the WoE ratings for these studies.

Table 17. Weight of Evidence assessments for studies with a primary focus on Literacy

Study	WoE A Trustwothiness of evidence	WoE B Appropriateness of design for this review	WoE C Relevance of Focus for this review	WoE D Overall strength of evidence for this review
Vadasy & Sanders, 2010	High	Medium	High	High
Vadasy & Sanders, 2013	Medium	Medium	Medium	High
Ehri et al., 2007	High	Medium	High	High
Proctor et al., 2011	High	High	Medium	Medium
Almaguer, 2005	Low	Low	High	Low
Solari & Gerber, 2008	High	High	Medium	High
Tong et al., 2010	Low	High	High	Medium
Kamps et al., 2007	Medium	Medium	High	High
Lugo-Neris et al., 2010	High	Medium	High	High
Graves et al., 2011	Low	Medium	Medium	Low

Vadasy and Sanders (2010, 2013) is an example of an intervention focused on phonics training and decoding where the 2010 study is the report of the implementation of the intervention and the 2013 paper reports on a two-year follow up of the same intervention and participants. Both of these papers were rated as 'high' overall though there is some variability in terms of their ratings for each of the WoE criteria. The students in the intervention received individual systematic and explicit phonics instruction where children participated in sessions aimed to improve their decoding skills and oral reading practice. The overall results indicated that the intervention students significantly outperformed control students on measures of alphabetic knowledge, word reading, spelling, passage reading fluency and reading comprehension and in general indicated significant positive treatment effects for kindergarten students who averaged in the lower quartile in language and literacy skills at pretest. In Vadasy and Sanders (2013), follow up analyses on data from 96% of the original sample were available where there was little loss reported over time, and even small gains noted. They reported strong positive relationships with each predictor variable and grade 3 outcomes (i.e., the two year follow up). Taken together, therefore, these studies illustrate benefits of introducing these word-level analyses/activities for students with EAL who are struggling with reading.

Solari and Gerber (2008) also received an overall high rating on WoE D and like the two Vadasy and Sanders (2010; 2013) studies were focused in part on phonological features of word reading. However, Solari and Gerber also investigated listening comprehension and general vocabulary. Interestingly, Solari and Gerber (2008) was one of the few studies which included listening comprehension skills in the intervention and directly compared the results of three different

treatment conditions which varied in terms of how much time was spent on either phonological awareness (PA) activities, listening comprehension and vocabulary activities (LC) and alphabetic knowledge activities. Their results indicated that, not surprisingly, those children who were in the groups where more time was spent on listening comprehension and vocabulary outperformed the other children on measures of listening comprehension. However, the listening comprehension treatment group also had higher scores on outcome measures of phonological awareness. Their study also included children who were 'at risk' for reading difficulties as measured by pre-test measures of PA and vocabulary and their results indicated that both 'at risk' and non at risk students who received the intervention where LC was emphasized over PA performed equally as well as those who received only PA or had PA plus alphabetic training. This study is one of the very few to identify that focusing on listening comprehension skills can not only improve listening comprehension but at no cost to word level skills (e.g., phonological awareness) and emphasizes the importance of developing the 'listening comprehension' contribution to models of reading comprehension (e.g., the Simple view of Reading – Gough & Tunmer, 1986).

Ehri et al (2007) and Kamps et al (2007) each received high overall ratings on WoE D and are both examples of implementing pre-developed (i.e., not researcher designed) interventions aimed at improving EAL children's reading performance. Ehri et al (2007) specifically examined the effectiveness of a version of *Reading Recovery,* which they refer to as *Reading Rescue* which (like Vadasy & Sanders, 2010; 2013) included systematic, sequential phonics instruction and decoding where children read short books which contained both high frequency and decodable words and included text-implicit and text-explicit questioning predicted to foster reading comprehension. Kamps et al (2007) implemented three different curricula: 'Reading Master', 'Early Interventions in Reading' and 'Read Well', each of which is considered an 'integrated curriculum' using direct instruction strategies, teacher modeling and multiple activities and repeated practice to teach and reinforce literacy skills through guided reading activities. Both of these studies report facilitative effects of their respective interventions and the *Reading Recovery (Rescue)* study in particular has been shown to be effective for struggling readers at the beginning stages of learning how to read (in either EAL or non-EAL children). The effectiveness of *Reading Recovery* for EAL children has been demonstrated in Clancy (2009; 2010) in the UK context.

Lugo-Neris et al (2010) and Almaguer (2005) report on interventions involving some shared reading activities where Almaguer reports on the effectiveness of reading in pair groups, where one member of the pair is of a higher reading ability than the other. Lugo-Neris et al report on an intervention implementing a shared storybook reading programme which included direct and explicit instruction of vocabulary, as well as through links to the L1, which was Spanish. Both

studies report effective results from their interventions. Almaguer (2005) concluded that dyad (pair) book reading with mixed ability pairs can be effective for both the 'lead' and 'assistant' reader for both word reading fluency and comprehension and that pair book reading activity is potentially particularly beneficial for children with EAL who may be reluctant to speak out in class. Lugo-Neris et al (2010) argued for positive transfer of knowledge between L1 and L2 in word learning where Spanish expansions of novel vocabulary words during English storybook reading resulted in comparable or greater growth in children's expressive knowledge of the target vocabulary. However this advantage was shown only for 1 out of 3 of the dependent measures (expressive definitions) used in this study. Graves et al (2011) also report on an intervention which included guided book reading activities for 6th grade struggling readers, both non-EAL and EAL. However, while the EAL results were reported separately and suggested benefits of the intervention for EAL children, the EAL results in particular were somewhat anecdotally reported, no doubt contributing to the low WoE rating for this study.

Proctor et al (2011) was the only paper in this predominantly 'literacy' category which included the use of software. Specifically they used a Strategic Digital Reading (SDG) prototype called 'ICON' (Improving Comprehension Online) with Spanish translation of all texts and human voice capability. Therefore, like the Lugo-Neris et al (2010) study, students were able to tap in to the L1 in completing the different vocabulary and reading activities with ICON. Proctor et al (2011) report strong treatment effects for vocabulary - but no effects for comprehension - which the authors argue may be due to the fact that the ICON intervention did not target the 'right' strategies to show effects on a standardized assessment of reading comprehension. The main treatment effects only emerged on the researcher-designed outcome measures and not on the standardized measures used in this study, which undoubtedly contributed to the 'medium' rating on the WoE criteria D.

Tong et al (2010) is similar to many of the other studies in this category in that the focus of the intervention was on word-level skills, word reading fluency and comprehension, however, the kindergarten students in this study were participating in a Structured English Immersion (SEI) setting. Schools were randomly assigned to either the intervention condition which implemented an enhanced structured English immersion model or control condition which was the typical SEI practice. Tong et al (2010) report that their intervention was effective because EAL students in the enhanced SEI group responded favourably on 5 of their outcome measures: phonological awareness, knowledge of phonology and syntax, receptive oral language, letter and word recognition and reading comprehension. The overall WoE criterion for this study was medium, and it received a particularly damning WoE rating on the trustworthiness of the evidence criterion.

These studies were also analysed with respect to their effect size as indicated in Table 18.

Table 18. Treatment effect sizes on outcomes measures for children with EAL in interventions on literacy

Study	Effect Size reported?	Outcome measure and effect size	Strength of evidence
Vadasy & Sanders, 2010	Yes	Across measures, approximate Cohen's d = 0.83 for EAL and NS combined	High
		Across measures, approximate Cohen's d for EAL only = -0.030	Medium
Vadasy & Sanders, 2013	Yes	Word reading = 0.45 Spelling = 0.36 Reading Comprehension = 0.24	Medium Medium Low
Ehri et al., 2007	Yes	Decoding = 0.53 Comprehension = 0.43 Word Reading = 0.72 Word Attack = 0.65	Medium Medium High Medium-High
Proctor et al., 2011	Yes	Vocabulary Breadth = 0.84 Vocabulary Depth (Definitions) = 1.26 Vocabulary Depth (Picture) = 1.12	High High High
Almaguer, 2005	Yes	Reading Fluency = 0.74 Reading Comprehension = 0.60	High High
Solari & Gerber, 2008	Yes - η^2	Range across measures = 0.2 – 0.61	Medium-High
Tong et al., 2010	No	Used Multilevel Modeling	
Kamps et al., 2007	Yes	Nonword Fluency = 0.70 Oral Reading Fluency = 0.58	High Medium
Lugo-Neris et al., 2010	Yes - η^2	Naming = 0.45 Receptive Vocabulary = 0.85 Expressive Definitions = 0.70	High High High
Graves et al., 2011	Yes	Oral Reading Fluency = 0.57	Medium

The effect sizes reported in the studies in this category overall seem higher than those studies with a focus predominantly on language where there is only one low effect size on reading comprehension in Vadasy and Sanders (2013). The remaining are generally medium to high suggesting the treatment effects in these various studies were both statistically significant and educationally meaningful for the participants.

These studies were then reviewed according to the criteria presented in Table 15, the results of which are presented in Table 19.

Table 19. Methodological strength of evidence criteria for interventions primarily focussed on Literacy

Study	Research Design	Number of Cases	Attrition	Outcome Measure	Fidelity & Validity
Vadasy & Sanders, 2010	Medium-High	High	Medium-High	High	Medium-High
Vadasy & Sanders, 2013	Medium-High	Medium	Medium-High	High	Medium-High
Ehri et al., 2007	Medium	Medium	Medium	High	Medium-High
Proctor et al.,	Medium	High	Low	Medium	Medium

2011					
Almaguer, 2005	Medium-Low	Low	Low	Low	Medium-Low
Solari & Gerber, 2008	Medium	Low	Low	Medium-High	Medium
Tong et al., 2010	Medium	Medium	Medium	High	Medium-High
Kamps et al., 2007	Medium	High	Low	High	Medium
Lugo-Neris et al., 2010	Medium	Low	Medium	Low	Medium
Graves et al., 2011	Medium-Low	Medium-Low	Medium	Medium	Medium-Low

As with the interventions with a predominant focus on language, there are a few discrepancies between the WoE ratings in Table 17 and the methodological rankings in Table 19. However, as indicated above, unlike with the WoE criteria, this is in part due to the fact that to obtain a 'high' on 'research design' it has to be a pure RCT, and to obtain a 'high' on the 'fidelity and validity' category the evaluators of the intervention have to be independent of the researchers. Neither was the case for any of the studies reviewed in this in-depth review, hence the lower rankings for some studies. However, many studies that fell short of having independent evaluators, or a pure RCT design still nonetheless implemented many careful methodological controls to ensure as much as possible that threats to the reliability and validity of the intervention were reduced (e.g., Vadasy & Sanders, 2010; 2013, Ehri et al, 2007; Tong et al., 2010). Many of these studies did assign randomly to groups (a key characteristic of an RCT) even if they did not randomly sample from the population and/or matching procedures between groups were not implemented. As with the language-oriented interventions, there was some variability across the studies in terms of the number of participants, the validity of the outcome measures, and whether and to what extent the interventions had balanced groups and monitored participant attrition. However, none of the interventions were ranked as exclusively low which would have suggested serious threats to the design and implementation of the study. Note, however, that the Almaguer (2005) and Graves et al (2005) studies were both on the low-end of medium suggesting some significant concerns.

The majority of the papers with a primary focus on literacy either focus on explicit instruction of word-level skills (e.g., phonics, word-level skills (e.g., phonological awareness), alphabetic knowledge, and reading fluency) through text-based and/or comprehension-based activities. Many of these studies also examined these variables within the context of shared book or group reading activities. Both the WoE criteria and the methodological strength of evidence criteria in Tables 17 and 19 respectively, illustrate a mixed group of studies, with some coming out fairly highly rated in terms of their strength of evidence and relevance to this review. It is clear, therefore, that there are some potentially promising possibilities to explore within the UK context.

Review of interventions with a primary focus on Continuing Professional Development

Table 20 presents the WoE criteria ratings for the five studies that had a primary focus on continuing professional development as a means to improve EAL children's English language and/or literacy outcomes.

Table 20. Weight of Evidence assessments for studies with a primary focus on Continuing Professional Development

Study	WoE A Trustwothiness of evidence	WoE B Appropriateness of design for this review	WoE C Relevance of Focus for this review	WoE D Overall strength of evidence for this review
Lara-Alecio et al, 2012	Medium	Medium	High	High
Short et al., 2012	Medium	High	High	Medium
Kim et al., 2011	Medium	Low	Medium	Medium
Kotler et al., 2001	Medium	Medium	High	Medium
Matsumura et al., 2010	Medium	Medium	Low	Low

The highest rated study in this category is that by Lara-Alecio et al (2012) which reports a CPD intervention that included biweekly staff sessions and ongoing workshops with teachers to help promote the integration of language and literacy development within the science curriculum. Instructional activities included Daily Oral and Written Language in Science (DOWLS). These activities were supplemented with CRISELLA (Content area Reading in Science for English Literacy and Language Acquisition) tasks that focused on vocabulary development and extension through science-related expository texts. The purpose of these activities was to improve students' understanding of science concepts but given the inclusion of direct instruction of vocabulary, pronunciation, definitions and the like, to also improve the children's English vocabulary and literacy. This intervention also included the integration of writing in a unit called WAVES (Written and Academic oral language Vocabulary development in English and Science).

Short et al (2012) describe an intervention focused on a system for lesson planning and delivery that incorporates what are argued to be best practices for teaching academic English and provides teachers with a coherent method for improving student achievement. The intervention aimed to evaluate whether children with EAL who received the CPD through the SIOP (Sheltered Instruction Observation Protocol) model had higher achievements in reading, writing, and oral proficiency in English than teachers who do not receive the SIOP intervention. As with Lara-Alecio

et al (2012) the Short et al (2012) received uniformly good (medium-high) ratings on the WoE criteria.

Kim et al (2011), however, was not deemed by reviewers to be as relevant for the purpose of this specific review. They describe a study examining the 'Pathway Project' which is a CPD programme that argues for the effectiveness of teaching students cognitive and metacognitive processes, including modelling, scaffolding, guided practice, and independent use of strategies so that students develop the ability to select and implement appropriate strategies independently and to monitor and regulate their use. Teachers used the 'Reader's and Writer's Tool Kit' where teachers learned to use pretest, on-demand writing results and Pathway materials to teach a cognitive strategies approach to text-based analytical writing. Coaches then help teachers integrate analytical writing strategies into the EAL curriculum which in turn results in the teachers' use of cognitive strategies in reading and writing activities in their classrooms. This strategy use is meant to enhance student performance on the on-demand writing assessment which in turn is meant to result in enhanced student performance on the standardised assessments of English language arts. This is one of the few studies that included older students (upper secondary level).

Kotler et al (2001) is the only UK study in this review. It describes a CPD programme called 'Talking Partners' which presents a model for classroom interaction to help students' develop oral language skills aimed at helping teachers establish better interaction in classrooms. The CPD includes suggesting different types of verbal problem-solving activities to the teachers which are aimed to stretch the learner and which require interaction with both teacher and other pupils in small groups.

Matsumura et al (2010) describes research on a CPD programme aimed to support teachers where there is high teacher turnover (mobility) in urban districts, which also is an area where many children with EAL live and attend school. The implication is that the high teacher mobility can have a negative impact on the quality of teaching, hence the development of the "Content-Focused Coaching" (CFC) - aimed at providing teachers with a coach to help support them in their teaching of reading comprehension. The purpose of the study was to examine the CFC programme's effectiveness in primary schools with high numbers of low income, minority, EAL students.

The effect sizes reported for these five studies are presented in Table 21 and unlike the language and literacy categories of studies, the effect sizes, when reported, are comparatively low. While most of these studies reported reasonably good measures for ensuring fidelity to treatment (see

Table 22), the general impact on the students' performance nonetheless seems to be less educationally significant (i.e., smaller effect sizes) in this group of studies on CPD activities.

Table 21. Treatment effect sizes on outcomes measures for children with EAL in interventions on Continuing Professional Development (CPD)

Study	Effect Size reported?	Outcome measure and effect size	Strength of evidence
Lara-Alecio et al, 2012	Yes (Cramer's V and partial η^2)	District reading test 1 through 6 = 0.103-0.238 TAKS (Texas Assessment of Knowledge and Skills) in reading 0.11 DIBELS (Dynamic Indicators of Basic Literacy Skills) = 0.134 (partial η^2)	Low-Medium Low-Medium Low
Short et al., 2012	Yes	Writing = 0.31 Reading = 0.16 Oral = 0.29	Low-Medium Low Low-Medium
Kim et al., 2011	No	Used Multilevel modeling	
Kotler et al., 2001	No		
Matsumura et al., 2010	No	Used multilevel modeling	

We then applied the criteria in Table 15 on the key methodological variables to these five studies, the results of which are presented in Table 22.

Table 22. Methodological strength of evidence criteria for interventions primarily focussed on Continuing Professional Development (CPD)

Study	Research Design	Number of Cases	Attrition	Outcome Measure	Fidelity & Validity
Lara-Alecio et al, 2012	Medium	Medium	Medium-Low	Medium-High	Medium
Short et al., 2012	Medium	Low-Medium	Medium-Low	Medium	Medium
Kim et al., 2011	Medium	Medium	Medium	Medium-High	Medium
Kotler et al., 2001	Low	Low	Low	Medium	Low
Matsumura et al., 2010	Medium-Low	Medium	Medium-Low	Medium-Low	Medium

As with the other studies in this review, none of them achieved a rating of 'high' on the 'fidelity and validity' category or the 'research design' category. While two of the studies were on the high end of medium on outcome measures (Lara-Alecio et al, 2012; Kim et al., 2011) none of these studies were rated as high on any of these criteria. However, only one study had a more uniform 'low' rating (Kotler et al, 2001), ratings which are consistently lower than the WoE criteria for this study in Table 20. This is due most likely because while the focus of the Kotler et al (2001) study is highly relevant for this review – i.e., a CPD programme aimed to enhance verbal interaction in classrooms aimed at improving EAL children's English language – from a methodological point of

view it fell short on a number of variables. Short et al (2012) included a small number of teachers (23 in treatment and 19 in control), but a reasonable number of students (278 EAL in treatment group and 169 in control) and while the low number of teachers can influence the intervention, the higher number of students who were yielding the data for the outcome measures was more appropriate resulting in a mixed rating of Low-Medium for 'number of cases'. Unfortunately, the CPD interventions generally receive a lower rating on methodological characteristics.

Review of interventions with a primary focus on family literacy practice

This final sub-category of studies in this review were those two studies that focused on some aspect of family literacy practice. The WoE criteria for these two studies are presented in Table 23 and illustrate that the Harper et al (2011) study is quite highly rated in comparison to the Kim and Guryan (2010) study.

Table 23. Weight of Evidence assessments for studies with a primary focus on family literacy practice.

Study	WoE A Trustwothiness of evidence	WoE B Appropriateness of design for this review	WoE C Relevance of Focus for this review	WoE D Overall strength of evidence for this review
Harper et al, 2011	High	High	Medium	High
Kim & Guryan, 2010	Medium	Medium	High	Medium

The aim of Harper et al (2011) - a study carried out in Canada - was to evaluate a family literacy program on EAL children's early reading development by looking at EAL and non-EAL children's scores in 3 areas of early reading before and after families' participation in the program compared against a control group of ELL and non-EAL children. Kindergarten children (aged 4-6) and their parents with a range of L1s participated in a 9-week family literacy programme featuring joint parent-child activities related to language and literacy development. The programme provided parents with information and ideas for creative and meaningful ways to enhance their children's emergent literacy skills. The results indicate a significant program effect for EAL children but not for EL1 children where children with EAL demonstrated significantly greater gains in their ability to infer meaning from print against NS children and EAL children who did not participate in the intervention.

Kim and Guryan (2010) aimed to prevent summer reading loss by sending books to children's homes and to encourage independent reading during summer holidays. Children were randomly assigned to one of 3 conditions: i) children received 10 self selected books in the mail, ii) a family

literacy group where children received 10 self selected books *and* were invited to attend three 2-hour family literacy events and, iii) a control group where children received 10 self selected books *after* post test. The results of their study suggested that children in the treatment and family literacy groups (groups 1 and 2) reported reading more books over the summer than children in the control group however, there was no effect on reading achievement - neither treatment group, with or without the parent literacy training - had an effect on the outcome measures. Therefore, this study, like Troia (2004) is one of the few in this review that does not report statistically significant facilitative effects for participating in the intervention.

The effect sizes for the interventions on family literacy practice are presented in Table 24.

Table 24. Treatment effect sizes on outcomes measures for children with EAL in interventions on family literacy practice

Study	Effect Size reported?	Outcome measure and effect size	Strength of evidence
Harper et al, 2011	Yes (partial η^2)	Test of Early Reading Ability (TERA-3): range across effects = 0.02 – 0.57	Low-Medium
		Alphabet Knowledge (range across effects) = 0.04 – 0.46	Low-Medium
		Conventions of Print (range across effects) = 0.01 – 0.20	Low
		Meaning (range across effects) = 0.03 – 0.24	Low
Kim & Guryan, 2010	No		

Kim and Guryan (2010) did not report effect sizes in their analyses. Harper et al, however, report low to medium level effects. As with the interventions on CPD therefore, the interventions on family literacy seemed to have less of an impact on students' outcomes, even if still statistically significant.

Table 25 presents the additional methodological strength of evidence for these two studies.

Table 25. Methodological strength of evidence criteria for interventions primarily focussed on family literacy practice.

Study	Research Design	Number of Cases	Attrition	Outcome Measure	Fidelity & Validity
Harper et al, 2011	Medium	Medium	Low-Medium	Low-Medium	Medium
Kim & Guryan, 2010	Medium	Medium-High	Medium	Low-Medium	Medium

Harper et al (2011) included data from 132 children and their families and came from a range of different L1 backgrounds but as is the case with the UK, a higher proportion came from South Asian language families (e.g., Punjabi, Urdu, Hindi). There was a degree of attrition in that not all families participated in each of the 9 sessions that formed this intervention, however, this attrition

was thoroughly monitored and reported and it did not result in substantial imbalance between the groups, hence the low end of Medium rating on 'attrition'. The study included one standardised assessment (TERA-3) which measures aspects of English reading but also a researcher-designed parent questionnaire aimed providing general information about the children in the study. Kim and Guryan (2010) were able to randomly assign children and their families to the different groups in their study and collected data from a range of background variables that were used to ensure matched groups. They also report the attrition from pre-test as being at 12% and carried out a chi-square analysis to show that there was no statistical relationship between experimental conditions and attrition rates. Different implementation checks were applied to try and gauge the extent to which the books being sent home were at appropriate levels for the children and the extent to which families participated in the summer literacy events. While Harper et al (2011) received generally higher WoE ratings than the Kim and Guryan (2010) study, in examining the specific methodological variables in Table 25, the two family literacy based studies are fairly similarly ranked.

Conclusions and Implications

In the previous sections we outlined the protocol followed to identify the 29 studies included in an in-depth review of intervention studies aimed to improve some aspect of EAL children's English language and/or literacy. The results of the review indicated that the significant majority of the studies were carried out in the context of the US. This finding demonstrates the urgent need to carry out controlled intervention studies within the UK context. The educational and social infrastructure of the US is very different from the UK, hence making generalisations across US-based studies to the UK is somewhat reckless. Furthermore, the population of EAL students in the US is quite different in that a large proportion of minority language learners in the US are Spanish-speaking, hence drawing on the L1 is feasible and indeed appropriate where there is a degree of homogeneity in the L1 of EAL students. In most parts of the UK, there is far less consistency with respect to EAL children's L1 background which makes it more challenging to explicitly use the L1 in targeted interventions aimed at improving children's English language and/or literacy.

The review also indicated that the majority of the studies focused on either some aspect of vocabulary or word-level skills such as phonological awareness, decoding, alphabetic knowledge and morphological awareness. Within those studies that focused on vocabulary, a large number targeted academic vocabulary specifically. Undoubtedly this focus on vocabulary is due in part to the considerable number of research studies that have clearly identified strong predictive relationships between vocabulary knowledge and reading comprehension (e.g., Nation &

Snowling, 1998; 2004). Furthermore, a number of studies have identified that children with EAL tend to have less vocabulary knowledge than non-EAL peers (e.g., Bialystok, Luck, Peets & Yang, 2010; Cameron, 2002). It is not surprising therefore, to see that an area of focus in intervention studies is to enhance vocabulary knowledge in children with EAL. The specific focus on academic vocabulary is again reflected in the numerous studies that have argued for the importance of academic language in underpinning curricular achievement (e.g., Cummins, 2000; 2012). More research aimed at enhancing vocabulary knowledge needs to be carried out in the UK context.

It is worth considering why there has been such a focus on vocabulary, phonics, decoding and alphabetic knowledge in these interventions. As indicated earlier, current thinking about the potential reasons underpinning the reported lower levels of academic achievement in many students with EAL suggest weaker reading comprehension skills may in part be responsible. Indeed, some children with EAL in the UK have been shown to lag behind their non-EAL peers on measures of reading comprehension, despite having comparable skills in single word reading (decoding) (Bourgoyne, Whitely & Hutchinson, 2011; Bourgoyne, Kelly, Whitely & Spooner, 2009; Hutchinson, Whitely, Smith & Connors, 2003). This focus on vocabulary is particularly important since the EAL children in these studies did not have a problem with single word reading, decoding and alphabetic knowledge, yet they nonetheless were significantly behind their non-EAL peers in comprehension. This is presumably due to the fact that their general vocabulary knowledge was less well elaborated relative to non-EAL peers. Within the Simple View of Reading model (Gough & Tunmer, 1986), reading comprehension is conceived of as being made up of comprehension skills (vocabulary and listening comprehension) as well as decoding/phonics skills. If children with EAL typically have good decoding skills, as has been demonstrated in numerous studies, yet nonetheless lag behind non-EAL peers in reading comprehension, then vocabulary is a good candidate for intervention given the number of studies that have demonstrated a general lack of vocabulary knowledge in EAL children relative to non-EAL peers. The studies in this in-depth review which focussed on these types of vocabulary certainly proved to be successful, either in terms of the students' learning of the target vocabulary of the study or in terms of wider more literacy-based measures (e.g., Mancilla-Martinez, 2010).

These issues are particularly important when considering approximately 95%-98% of the vocabulary in a text needs to be understood in order to be able to derive a general meaning of the text (Schmitt, Jiang & Grabe, 2011), though this percentage does vary somewhat depending on the text. Given the importance of knowing almost all of the words in a text to extract meaning from it, expanding all pupils' vocabulary, but particularly EAL pupils, is likely to be effective as has been demonstrated by many of the studies in this review. Academic vocabulary is of particular

importance because there are many words which are unlikely to occur in conversation but which are likely to turn up in a wide range of academic texts. It is estimated that there are about 7,000 word families which fall into this category. Teachers cannot possibly explicitly teach 7,000 word families but they can teach some, and they can also teach strategies for learning new words based on morphemes (i.e., the smallest meaningful unit of words) from which words are constructed. This was an effective approach adopted in a number of studies (e.g., Kieffer & Lesaux, 2012; Lesaux et al, 2010).

Some children struggle with reading (both non-EAL and EAL) and those readers who have lower scores on measures of reading and reading comprehension can have a range of different difficulties which lead to their weaker performance. Some readers struggle with the decoding aspect of reading - mapping phonemes on to graphemes. Whereas generally speaking children with EAL do not have this problem, we would still expect to find a proportion of the EAL population who have difficulty with decoding. A number of the studies in this review reported on evidence where children were 'at risk' or were struggling readers as identified by their low performance on pre-test measures of word knowledge and single word reading performance. In these cases, focussing on word-level skills such as phonics, decoding, and alphabetic knowledge, is important and as the studies in this review attest, can be effective. For example, Vadasy and Sanders (2010; 2013) and Ehri et al (2007) were some of the examples of effective interventions aimed at improving word reading performance in children who struggle with these aspects of reading.

Therefore, interventions which focus on vocabulary, and specific aspects of word reading (e.g., decoding) can be effective for a number of different reasons, but clearly must be aimed at the needs of the students. Students who have particular difficulties with single word reading, and reading accuracy/fluency can benefit from interventions which improve their alphabetic knowledge and decoding skills (e.g., Vadasy & Sanders, 2010; 2013; Ehri et al, 2007). Students who have good decoding skills, yet who nonetheless have lower reading comprehension scores than non-EAL children can benefit primarily from vocabulary interventions that improve children's word knowledge, and specifically with academic vocabulary which may not be encountered as frequently as other more all-purpose words.

There were five examples of Continuing Professional Development (CPD) programmes in this review. This small number is somewhat surprising in that pedagogical approaches are the most obvious starting point for considering improving EAL students' academic achievement. However, a number of the interventions included some form of CPD even if it was not the main focus (see Table 12). One of the main themes that emerged from a review of research on EAL provision

commissioned by the then *Training and Development Agency (TDA)* indicated that there was a lack of specialised staff, that management of EAL provision was ill-defined, that there was too much crossover with Special Educational Needs (SEN) provision, and that there was a need for more EAL-focused training for teachers due to a striking absence of EAL pedagogy (Andrews, 2009). Andrews (2009) was carried out when EMAS (Ethnic Minority Achievement Service) teams were still largely in place, whereas currently funding for EMAS has been largely withdrawn and hence there is arguably even less support for the development of effective EAL pedagogy. Andrews's (2009) review was carried out in the UK context and of the 5 CPD studies that were included in this review, four of these were carried out in the context of the US.   Therefore, there is a considerable need for further research that examines more precisely how CPD can support teachers of children with EAL through carefully controlled intervention studies.

There were only two intervention studies that examined the role of parents/family in enhancing EAL children's language and/or literacy performance, and of these, only one (Harper et al 2011) reported beneficial effects for the EAL pupils in their sample. This lack of focus on parent-oriented interventions and mixed level of success may reflect Gorard and See's (2013) findings that primary-age interventions which enhance parental involvement are generally ineffective in increasing children's attainment. This is a somewhat surprising finding given the number of studies suggesting that the environment in the home is such a powerful predictor of children's levels of academic success (Murphy, 2014) and is clearly an area that would benefit from further attention. Parental involvement in the context of children with EAL might be particularly amenable to intervention if it helped those parents with weaker English skills to bridge the home language – school language barrier.   In other words, outcomes which were not included in this review such as attendance, engagement with school, well being etc., may benefit from targeted intervention for parents of children with EAL.

Many of the studies in this review focused on primary school children, with far fewer aimed at enhancing the English language and/or literacy skills of secondary school students, and only one which was targeted at older secondary school pupils. However, later secondary school years are particularly important in that most children take high stakes national examinations at this stage in their educational development. It would be worth expanding the breadth of research on children with EAL to encompass all of the formal education years, with studies aimed at facilitating the transition from primary to secondary as well as examining best practice and support for older secondary school students with EAL.

There were also rather mixed findings in terms of the strength of evidence of the in-depth review studies. There were no examples of pure RCT studies or those where the effectiveness of the intervention was evaluated by an independent review team. However, many studies did include many characteristics of RCT designs and incorporated a number of careful methodological practices to ensure fidelity to treatment and to establish appropriate matching (e.g., random assignment to groups). Nonetheless, there were some studies that had rather low ratings across a range of criteria suggesting an undesirable level of variability in the overall methodological rigour of these studies.

As with all research there are a number of limitations to this review and two main ones are worth highlighting here. As with all research the results are methodological artefact. In other words, the results obtained are a direct consequence of the methodology adopted. This review is no different in that the use of different search terms, different databases and/or different exclusion/inclusion criteria would have likely yielded different studies in the in-depth review. Secondly, there is a level of subjectivity in how reviewers weigh evidence of research studies in systematic reviews such as this one. Some reviewers may have weighed some studies' characteristics differently than others. We attempted to mitigate against this subjectivity as much as possible in having two independent (blind) reviewers, and in applying a two-tiered process of review. Despite these limitations, however, this review has objectively identified a range of studies that are relevant for the review questions and which identify a number of interesting studies which offer credible evidence concerning the development of EAL pupils' English language and/or literacy skills.

The results of this review then, suggest that there are some promising areas to explore further with respect to how to best support children with EAL in their English language and/or literacy outcomes. These are in the areas of developing English vocabulary and word-level skills, phonics and alphabetic knowledge for struggling readers (e.g., reading recovery) and developing book reading activities with appropriate explicit instruction of vocabulary together with comprehension strategies. Specifically, children with EAL who struggle with word reading can benefit most from those interventions that focus on alphabetic knowledge, phonics, phonological awareness and other such word-level skills. Children with EAL who have good decoding skills but who nonetheless fall short of non-EAL peers on measures of reading comprehension benefit most from explicit (academic) vocabulary instruction, and strategy training to enhance their abilities to analyse words through their morphological structure. However, as only one study was carried out in the UK it is not appropriate to assume that these effective interventions will be immediately transferable to a UK context. Therefore, there is considerable scope for research that implements interventions aimed at these linguistic features in children with EAL in the UK. Furthermore,

research which examines how best to support the CPD of teachers and how to support parents of children with EAL are also likely to be fruitful areas for future investigation. CPD was a component of many of the interventions discussed in this review, and while the parent-focussed interventions where themselves less effective in terms of improving English language/literacy outcomes in children with EAL, it is possible that such parental support interventions might be able to improve other outcome variables (such as attendance and engagement with school, etc) that in turn might have positive influences on EAL children's English language and literacy outcomes.   Further carefully conducted intervention studies in these areas are likely to yield useful results which in turn can equip teachers and schools with credible evidence upon which to develop effective support for children with EAL.

About the authors:

Professor Victoria Murphy is Professor of Applied Linguistics at the University of Oxford

Adam Unthiah is a doctoral student at the University of Oxford

References

Almaguer, I. (2005). Effects of Dyad Reading Instruction on the Reading Achievement of Hispanic Third-Grade English Language Learners. *Bilingual Research Journal*, *29(3)*, 509–526

Andrews, R. (2009). *Review of research in English as an Additional Language (EAL).* London: NRDC Institute of Education.

Bialystok, E., Luk, G., Peets, K.F. & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13(4)*, 525-531.

Burgoyne, K., Kelly, J.M., Whiteley, H.E. & Spooner, A. (2009). The comprehension skills of children learning English as an additional language. *British Journal of Educational Psychology, 79*, 735-747.

Burgoyne, K., Whiteley, H.E. & Hutchinson, J.M. (2011). The development of comprehension and reading-related skills in children learning English as an additional language and their monolingual English-speaking peers. *British Journal of Educational Psychology, 81,* 344-354.

Calhoon, M. B., Otaiba, S. A., Cihak, D., King, A., & Avalos, A. (2007). Effects of a Peer-Mediated Program on Reading Skill Acquisition for Two-Way Bilingual First-Grade Classrooms. *Learning Disability Quarterly*, *30(3)*, 169. doi:10.2307/30035562

Cameron, L. (2002). Measuring vocabulary size in English as an Additional Language. *Language Teaching Research, 6(2),* 145-173.

Carlo, M. S., August, D., Mclaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., Lively, T. J., & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, *39(2)*, 188–215. doi:10.1598/RRQ.39.2.3

Cena, J., Baker, D. L., Kame'enui, E. J., Baker, S. K., Park, Y., & Smolkowski, K. (2013). The impact of a systematic and explicit vocabulary intervention in Spanish with Spanish-speaking English learners in first grade. *Reading and Writing, 26(8)*, 1289–1316. doi:10.1007/s11145-012-9419-y

Chall, J. (1983). *Learning to Read: The Great Debate*. Wiley: New York.

Cirino, P. T., Vaughn, S., Linan-Thompson, S., Cardenas-Hagan, E., Fletcher, J. M., & Francis, D. J. (2009). One-Year Follow-Up Outcomes of Spanish and English Interventions for English Language Learners at Risk for Reading Problems. *American Educational Research Journal*, *46(3)*, 744–781. doi:10.3102/0002831208330214

Clancy, C. (2009). *Reading Recovery and young English language learners: A small scale study considering the effects of early intervention on learner outcomes.* Unpublished MSc thesis. University of Oxford, UK.

Clancy, C. (2010). *Investigating the suitability of the Reading Recovery tests for use with EAL children.* Unpublished MSc thesis. University of Oxford, UK.

Crevecoeur, Y. C., Coyne, M. D., & McCoach, D. B. (2014). English Language Learners and English-Only Learners' Response to Direct Vocabulary Instruction. *Reading & Writing Quarterly*, *30(1)*, 51–78. doi:10.1080/10573569.2013.758943

Cruz de Quirós, A. M., Lara-Alecio, R., Tong, F., & Irby, B. J. (2012). The effect of a structured story reading intervention, story retelling and higher order thinking for English language and literacy acquisition. *Journal of Research in Reading*, *35(1)*, 87–113. doi:10.1111/j.1467-9817.2010.01472.x

Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire.* Clevedon, UK: Multilingual Matters.

Cummins, J. (2012). The intersection of cognitive and sociocultural factors in the development of reading comprehension among immigrant students. *Reading and Writing, 25,* 1973-1990.

Denton, C. A., Anthony, J. L., Parker, R., & Hasbrouck, J. E. (2004). Effects of Two Tutoring Programs on the English Reading Development of Spanish-English Bilingual Students. *Elementary School Journal*, *104(4)*, 289.

Dressler, C., Carlo, M. S., Snow, C. E., August, D., & White, C. E. (2011). Spanish-speaking students' use of cognate knowledge to infer the meaning of English words. *Bilingualism: Language and Cognition*, *14(2)*, 243–255. doi:10.1017/S1366728910000519

Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading Rescue: An Effective Tutoring Intervention Model for Language-Minority Students Who Are Struggling Readers in First Grade. *American Educational Research Journal*, *44(2)*, 414–448. doi:10.3102/0002831207302175

Filippini, A. L., Gerber, M. M., & Leafstedt, J. M. (2012). A Vocabulary-Added Reading Intervention for English Learners At-Risk of Reading Difficulties. *International Journal of Special Education*, *27(3)*, 14–26.

García, O. (2009). *Bilingual education in the 21st century: A global perspective.* Oxford: Blackwell.

García, O. & Kleifgen, J. (2010). *Educating emergent bilinguals: Policies, programs and practices for English language learners.* New York: Teachers College Press.

Genesee, F., Lindholm-Leary, K., Saunders, W.M. & Christian, D. (Eds.) (2006). *Educating English Language Learners: A synthesis of research evidence.* Cambridge: Cambridge University Press.

Giambo, D. A., & Mckinney, J. D. (2004). The Effects of a Phonological Awareness Intervention on the Oral English Proficiency of Spanish-Speaking Kindergarten Children. *TESOL Quarterly*, *38(1)*, 95–117. doi:10.2307/3588260

Goodrich, J. M., Lonigan, C. J., & Farver, J. M. (2013). Do Early Literacy Skills in Children's First Language Promote Development of Skills in Their Second Language? An Experimental Evaluation of Transfer. *Journal of Educational Psychology*, *105(2)*, 414–426. doi:10.1037/a0031780

Gorard, S & See, B.H. (2013). *Do parental involvement interventions increase attainment? A review of the evidence.* Nuffield Foundation Briefing Paper.

Gough, D. (2007). Weight of Evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22(2),* 213-228.

Gough, D., Oliver, S. & Thomas, J. (2012). *An Introduction to Systematic Reviews.* London: Sage.

Gough, P.B. & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7(1),* 6-10.

Graves, A., Duesbery, L., Pyle, N., Brandon, R., & McIntosh, A. (2011). Two Studies of Tier II Literacy Development: Throwing Sixth Graders a Lifeline. *The Elementary School Journal*, *111(4)*, 641–661.

Greenfader, C. M., & Brouillette, L. (2013). Boosting Language Skills of English Learners Through Dramatization and Movement. *The Reading Teacher*, *67(3)*, 171–180. doi:10.1002/TRTR.1192

Harper, S., Platt, A., & Pelletier, J. (2011). Unique Effects of a Family Literacy Program on the Early Reading Development of English Language Learners. *Early Education and Development*, *22(6)*, 989–1008.

Hopewell, S. (2011). Leveraging bilingualism to accelerate English reading comprehension. *International Journal of Bilingual Education and Bilingualism*, *14(5)*, 603–620. doi:10.1080/13670050.2011.564274

Hutchinson, J.M., Whiteley, H.E., Smith, C.D. & Connors, L. (2003). The developmental progression of comprehension-related skills in children learning EAL. *Journal of Research in Reading, 26(1)*, 19-32.

Kamps, D., Abbott, M., Greenwood, C., Arreaga-Mayer, C., Wills, H., Longstaff, J., Culppepper, M., & Walton, C. (2007). Use of Evidence-Based, Small-Group Reading Instruction for English Language Learners in Elementary Grades: Secondary-Tier Intervention. *Learning Disability Quarterly*, *30(3)*, 153–168.

Kieffer, M. J., & Lesaux, N. K. (2012). Effects of Academic Language Instruction on Relational and Syntactic Aspects of Morphological Awareness for Sixth Graders from Linguistically Diverse Backgrounds. *Elementary School Journal*, *112(3)*, 519–545.

Kim, J. S., & Guryan, J. (2010). The Efficacy of a Voluntary Summer Book Reading Intervention for Low-Income Latino Children from Language Minority Families. *Journal of Educational Psychology*, *102(1)*, 20–31.

Kim, J. S., Olson, C. B., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D., Collins, P., & Land, R. E. (2011). A Randomized Experiment of a Cognitive Strategies Approach to Text-Based

Analytical Writing for Mainstreamed Latino English Language Learners in Grades 6 to 12. *Journal of Research on Educational Effectiveness*, *4(3)*, 231–263.

Kotler, A., Wegerif, R., & Levoi, M. (2001). Oracy and the Educational Achievement of Pupils with English as an Additional Language: The Impact of Bringing "Talking Partners" into Bradford Schools. *International Journal of Bilingual Education and Bilingualism*, *4(6)*, 403–419. doi:10.1080/13670050108667740

Lara-Alecio, R., Tong, F., Irby, B. J., Guerrero, C., Huerta, M., & Fan, Y. (2012). The effect of an instructional intervention on middle school English learners' science and English reading achievement. *Journal of Research in Science Teaching*, *49(8)*, 987–1011. doi:10.1002/tea.21031

Lawrence, J. F., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. E. (2012). Language proficiency, home-language status, and English vocabulary development: A longitudinal follow-up of the Word Generation program. *Bilingualism: Language and Cognition*, *15(3)*, 437–451. doi:10.1017/S1366728911000393

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The Effectiveness and Ease of Implementation of an Academic Vocabulary Intervention for Linguistically Diverse Students in Urban Middle Schools. *Reading Research Quarterly*, *45(2)*, 196–228.

Lesaux, N.K., Koda, K., Siegel, L. & Shanahan, T. (2006). Development of literacy. In D. August & T. Shanahan (Eds). *Developing literacy in second language learners: Report of the national literacy panel on language minority children and youth*. Mahwah, NJ: Lawrence Erlbaum.

Lugo-Neris, M. J., Jackson, C. W., & Goldstein, H. (2010). Facilitating Vocabulary Acquisition of Young English Language Learners. *Language, Speech, and Hearing Services in Schools*, *41(3)*, 314–327. doi:10.1044/0161-1461(2009/07-0082)

Macaruso, P., & Rodman, A. (2011). Benefits of Computer-Assisted Instruction to Support Reading Acquisition in English Language Learners. *Bilingual Research Journal*, *34(3)*, 301–315. doi:10.1080/15235882.2011.622829

Mancilla-Martinez, J. (2010). Word Meanings Matter: Cultivating English Vocabulary Knowledge in Fifth-Grade Spanish-Speaking Language Minority Learners. *TESOL Quarterly*, *44(4)*, 669–699. doi:10.5054/tq.2010.213782

Marx, A.E. & Stanat, P. (2012). Reading comprehension of immigrant students in Germany: Research evidence on determinants and target points for intervention. *Reading and Writing, 25*, 1929-1945.

Matsumura, L. C., Garnier, H. E., Correnti, R., Junker, B., & Bickel, D. D. (2010). Investigating the Effectiveness of a Comprehensive Literacy Coaching Program in Schools with High Teacher Mobility. *Elementary School Journal, 111(1)*, 35–62.

McKendry, M. & Murphy, V.A. (2011). A comparative study of listening comprehension measures in English as an additional language and native English-speaking primary school children. *Evaluation and Research in Education, 24*, 17 – 40.

Murphy, V.A. (2014). *Learning a second language in the early school years: Trends and contexts.* Oxford: Oxford University Press.

NALDIC (2014). National Association for Language Development in the Curriculum. http://www.naldic.org.uk/

Nation, K. & Snowling, M. (1998). Semantic processing skills and the development of word recognition: Evidence from children with reading comprehension difficulties. *Journal of Memory & Language, 39*, 85-101.

Nation, K., & Snowling, M. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*, 342-356.

Olson, C. B., & Land, R. (2007). A Cognitive Strategies Approach to Reading and Writing Instruction for English Language Learners in Secondary School. *Research in the Teaching of English, 41(3)*, 269–303.

Organisation for Economic Co-Operation and Development (2003). *Where immigrant students succeed - A comparative review of performance and engagement in PISA 2003.* Report published by the OECD.

Paradis, J., Genesee, F., & Crago, M. (2011). *Dual language development and disorders: A handbook on bilingualism and second language learning* (2nd Edition). Baltimore, MD: Brookes.

Peugh, J.L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology, 48(1),* 85-112.

Proctor, C. P., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C., & Neugebauer, S. (2011). Improving Comprehension Online: Effects of Deep Vocabulary Instruction with Bilingual and Monolingual Fifth Graders. *Reading and Writing: An Interdisciplinary Journal*, *24(5)*, 517–544.

Schmitt, N., Jiang, X. & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95(1),* 26-43.

Schoenbrodt, L., Kerins, M., & Gesell, J. (2010). Using Narrative Language Intervention as a Tool to Increase Communicative Competence in Spanish-Speaking Children. *Language, Culture and Curriculum*, *16(1)*, 48–59. doi:10.1080/07908310308666656

Short, D. J., Fidelman, C. G., & Louguit, M. (2012). Developing Academic Language in English Language Learners Through Sheltered Instruction. *TESOL Quarterly*, *46(2)*, 334–361. doi:10.1002/tesq.20

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating Knowledge of Academic Language among Urban Middle School Students. *Journal of Research on Educational Effectiveness*, *2*, 325–344.

Solari, E. J., & Gerber, M. M. (2008). Early Comprehension Instruction for Spanish-Speaking English Language Learners: Teaching Text-Level Reading Skills While Maintaining Effects on Word-Level Skills. *Learning Disabilities Research & Practice*, *23(4)*, 155–168. doi:10.1111/j.1540-5826.2008.00273.x

Spycher, P. (2009). Learning Academic Language through Science in Two Linguistically Diverse Kindergarten Classes. *Elementary School Journal*, *109(4)*, 359–379.

Stanat, P. & Christensen, G.S. (2006). *Where immigrant students succeed: A comparative review of performances and engagement in PISA 2003*. Paris: OECD.

Tong, F., Irby, B. J., Lara-Alecio, R., Yoon, M., & Mathes, P. G. (2010). Hispanic English Learners' Responses to Longitudinal English Instructional Intervention and the Effect of Gender: A Multilevel Analysis. *Elementary School Journal*, *110(4)*, 542–566.

Tong, F., Lara-Alecio, R., Irby, B., Mathes, P., & Kwok, O. (2008). Accelerating Early Academic Oral English Development in Transitional Bilingual and Structured English Immersion Programs. *American Educational Research Journal*, *45(4)*, 1011-1044. doi:10.3102/0002831208320790

Townsend, D., & Collins, P. (2009). Academic vocabulary and middle school English learners: an intervention study. *Reading and Writing*, *22(9)*, 993–1019. doi:10.1007/s11145-008-9141-y

Troia, G. A. (2004). Migrant Students with Limited English Proficiency: Can Fast ForWord Language? Make a Difference in Their Language Skills and Academic Achievement? *Remedial and Special Education*, *25(6)*, 353–366.

Vadasy, P. F., & Sanders, E. A. (2010). Efficacy of Supplemental Phonics-Based Instruction for Low-Skilled Kindergarteners in the Context of Language Minority Status and Classroom Phonics Instruction. *Journal of Educational Psychology*, *102(4)*, 786–803.

Vadasy, P. F., & Sanders, E. A. (2013). Two-year follow-up of a code-oriented intervention for lower-skilled first-graders: the influence of language status and word reading skills on third-grade literacy outcomes. *Reading and Writing*, *26(6)*, 821–843. doi:10.1007/s11145-012-9393-4

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Hagan, E. C., Pollard-Durodola, S., Fletcher, D. J., & Francis, D. J. (2006). Effectiveness of a Spanish Intervention and an English Intervention for English-Language Learners at Risk for Reading Problems. *American Educational Research Journal*, *43(3)*, 449–487. doi:10.3102/00028312043003449

Vaughn, S., Linan-Thompson, S., Mathes, P. G., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., Cardenas-Hagan, E., & Francis, D. J. (2006). Effectiveness of Spanish intervention for first-grade English language learners at risk for reading difficulties. *Journal of Learning Disabilities*, *39(1)*, 56–73.

Appendix A – Studies from the Systematic Keyword Map (n = 44)

Authors	Title	Date	Publication	Country	Setting	Participants	L1	Sample
Almaguer	Effects of Dyad Reading Instruction on the Reading Achievement of Hispanic Third-Grade English Language Learners	2005	Bilingual Research Journal: The Journal of the National Association for Bilingual Education, 29:3, 509-526	USA	Mainstream grade 3 classes	EAL only	Spanish	Control group=40 Experimental group=40
Calhoon, Al Otaiba, Cihak, King, and Avalos	Effects of a Peer-Mediated Program on Reading Skill Acquisition for Two-Way Bilingual First-Grade Classrooms	2007	Learning Disability Quarterly, Vol.30, No.3, p.169-184	USA	Bilingual grade 1 classes. Treatment delivered as supplemental reading class	EAL and NS combined in treatment and contrast groups. EAL results presented separately	Spanish	PALS group=43 Contrast group=33
Carlo, August, Mclaughlin, Snow, Dressler, Lippman, Lively, and White	Closing the Gap: Addressing the Vocabulary Needs of English-Language Learners in Bilingual and Mainstream Classrooms	2004	Reading Research Quarterly, Vol.39, No.2, p.188-215	USA	Mainstream and bilingual grade 5 classes	EAL and NS combined in treatment and comparison groups. EAL results presented separately	Spanish	Comparison groups=85 Treatment groups=169
Cena, Baker, Kame'enui, Baker, Park, & Smolkowski	The impact of a systematic and explicit vocabulary intervention in Spanish with Spanish-speaking English learners in first grade	2013	Reading and Writing, 24, 1289-1316	USA	Early transition model bilingual program	EAL only	Spanish	G-SETR intervention=26 VE-SETR intervention=24
Cirino, Vaughn, Linan-Thompson, Cardenas-Hagan,	One-Year Follow-Up Outcomes of Spanish and English Interventions for English Language	2009	American Research Journal, Vol.46, No.3, p.744-781	USA	Bilingual early transition grade 2 classes	EAL only	Spanish	English study comparison group Cohort I=11 English study intervention group

Fletcher, and Francis	Learners at Risk for Reading Problems							Cohort I=18 English study comparison group Cohort II=44 English study intervention group Cohort II=38
Crevecoeur, Coyne, and McCoach	English Language Learners and English-Only Learners' Response to Direct Vocabulary Instruction	2013	Reading and Writing Quarterly: Overcoming Learning Difficulties, Vol.30, p.51-78	USA	Mainstream kindergartens	EAL and NS combined in treatment and control groups. EAL results presented separately	Spanish	Control groups=42 Treatment groups=79
Cruz de Quirós, Lara-Alecio, Tong	The Effect of a Structured Story Reading Intervention, Story Retelling and Higher Order Thinking for English Language and Literacy Acquisition	2012	Journal of Research in Reading, Vol.35, Issue 1, p.87-113	USA	Transitional bilingual classes grades 1 to 2	EAL only	Spanish	Control group=38 Treatment group=34
Denton, Anthony, Parker, and Hasbrouck	Effects of two tutoring programs on the English reading development of Spanish-English bilingual students	2004	The Elementary School Journal, Vol.104, No.4, p.289-305	USA	Grades 2-5 bilingual classrooms	EAL only	Spanish	Real Well comparison group=14 Read Well treatment group=19 Read Naturally comparison group=28 Read Naturally treatment group=32
Dressler, Carlo, Snow, August, & White	Spanish-speaking students' use of cognate knowledge to infer the meaning of English words	2011	Bilingualism: Language and Cognition, 14(2), 243-255	USA	One on one with researcher	8 EAL and 4 NS	Spanish	12 across one control group and one treatment group
Ehri, Dreyer, Flugman, and	Reading Rescue: An Effective Tutoring	2007	American Educational	USA	Mainstream grade 1	EAL only	Spanish	Control group 1=62 Control group 2=60

Gross	Intervention Model for Language-Minority Students Who Are Struggling Readers in First Grade		Research Journal, Vol.44, No.2, p.414-448		classes			Intervention group=64
Filippini, Gerber, & Leafstedt	A Vocabulary-Added Reading Intervention for English Learners At-Risk of Reading Difficulties	2012	International Journal Of Special Education, 27(3), 14-26	USA	Supplementary classes to mainstream arts classes	EAL and 10 NS designated as limited English proficient.	Spanish	71 across 3 treatment groups
Giambo and McKinney	The Effects of a Phonological Awareness Intervention on the Oral English Proficiency of Spanish-Speaking Kindergarten Children	2004	TESOL Quarterly, Vol.38, Issue 1, p.95-117	USA	Mainstream kindergarten classes	EAL only	Spanish	Comparison group=40 Treatment group=40
Goodrich, Lonigan, & Farver	Do Early Literacy Skills in Children's First Language Promote Development of Skills in Their Second Language? An Experimental Evaluation of Transfer	2013	Journal of Educational Psychology 105(2), 414-425	USA	English only and transitioning from Spanish to English	EAL only	Spanish	94 across 2 treatment conditions and one control group
Graves, Duesbery, Pyle, Brandon, and McIntosh	Two Studies Of Tier II Literacy Development: Throwing Sixth Graders a Lifeline	2011	The Elementary School Journal, 111(4), p.641-661	USA	Mainstream grade 6 classes	EAL, NS, and participants with learning disabilities. Results presented separately for all groups.	n/a	ORF control group=28 ORF treatment group=31 Vocabulary control group=26 Vocabulary treatment group=31 MAZE control group=24 MAZE treatment group=28

Greenfader & Brouillette	Boosting language skills of English learners through dramatization and movement	2013	The Reading Teacher 67(3), 171-180	USA	Mainstream kindergarten and first grade classes	EAL only	88%-91% Hispanic	Kindergarten=130 treatment, 1407 control. First grade=131 treatment, 1544 control
Harper, Platt, and Pelletier	Unique Effects of a Family Literacy Program on the Early Reading Development of English Language Learners	2011	Early Education and Development, 22(4), 989-1008	Canada	Family literacy program	EAL and NS in both treatment and control groups. Separate results for EAL	South Asian language 49%, East Asian language=26%, European language=12%, Middle Eastern language=10%, Other=2%	132 across two experimental groups and two control groups
Hopewell	Leveraging bilingualism to accelerate English reading comprehension	2011	International Journal of Bilingual Education and Bilingualism, 14(5), 603-620	USA	Biliteracy classrooms	EAL only	Spanish	41 across two treatment groups
Kamps, Abbott, Greenwood, Arreaga-Mayer, Wills, Longstaff, Culpepper, and Walton	Use of Evidence-Based, Small-Group Reading Instruction for English Language Learners in Elementary Grades: Secondary-Tier Intervention	2007	Learning Disability Quarterly, Vol.30, p.153-168	USA	Mainstream grade 1 and grade 2 classes	EAL and NS combined in control and treatment groups. EAL results presented separately	Spanish=99% Somalian, Sudanese, Vietnamese	Comparison group=113 (of which 60 EAL) Experimental group=117 (of which84 EAL)
Kieffer, and Lesaux	Effects of Academic Language Instruction on Relational and Syntactic Aspects of Morphological Awareness for Sixth Graders from Linguistically Diverse	2012	The Elementary School Journal, Vol.112, Issue 3, p.519-545	USA	Mainstream grade 6 classes	EAL and NS. EAL and NS combined in treatment and comparison groups. EAL results presented separately	Spanish=59.9%, Vietnamese=1.9%, Lao=1.7%, Hmong=1.3%, Somali=1.5%, Pilipino/Tagalog=2.3%, Other=6.7%	Control group=183 Treatment group=299

	Backgrounds							
Kim, and Guryan	The Efficacy of a Voluntary Summer Book Reading Intervention for Low-Income Latino Children from Language Minority Families	2010	Journal of Educational Psychology, Vol.102, No.1, p.20-31	USA	Voluntary summer book reading program	EAL only	Spanish	Control group=110 Treatment group=102
Kim, Olson, Scarcella, Kramer, Pearson, van Dyk, Collins, and Land	A Randomized Experiment of a Cognitive Strategies Approach to Text-Based Analytical Writing for Mainstreamed Latino English Language Learners in Grades 6 to 12	2011	Journal of Research on Educational Effectiveness, 4:231-263	USA	Mainstream classes grades 6 to 12	EAL	Primarily Spanish	779 in total Control group=51 classrooms Treatment group=52 classrooms
Kotler, Wegerif, and LeVoi	Oracy and the Educational Achievement of Pupils with English as an Additional Language: The Impact of Bringing "Talking Partners" into Bradford Schools.	2001	International Journal of Bilingual Education and Bilingualism, Vol.4, No.6, p.403-419	UK	Mainstream classes for 5 to 8 year old pupils (years 1 to 3)	EAL only	Punjabi, Sylheti	Control group=63 Treatment group=64
Lara-Alecio, Tong, Irby, Guerrero, Huerta, & Fan	The Effect of an Instructional Intervention on Middle School English Learners' Science and English Reading Achievement	2012	Journal of Research in Science Teaching, 49(8), 987-1011	USA	Mainstream science and English classes	EAL and NS. Results combined in both treatment and control groups	Spanish	166 across 2 treatment groups and 80 across 2 control groups
Lawrence, Capotosto, Branum-Martin, White, and	Language proficiency, home-language status, and English vocabulary	2012	Bilingualism: Language and Cognition,	USA	Mainstream grades 7 and 8 classes	EAL, NS, and LEP in treatment and comparison groups. Results for EAL	n/a	Comparison groups=204-525 EAL Treatment groups=680-1179

Author	Title	Year	Source	Country	Setting	Participants	Language	Sample size
Snow	development: A longitudinal follow-up of the Word Generation program		15, pp 437-451			presented separately		EAL
Lesaux, Kieffer, Falley, and Kelley	The Effectiveness and Ease of Implementation of an Academic Vocabulary Intervention for Linguistically Diverse Students in Urban Middle Schools	2010	Reading Research Quarterly, Vol.45, No.2, p.196-228	USA	Mainstream grade 6 classes	EAL and NS combined in treatment and control groups. EAL results presented separately	Spanish=59.9%, Vietnamese=1.9%, Lao=1.7%, Hmong=1.3%, Somali=1.5%, Pilipino/Tagalog= 2.3%, Other=6.7%	Control group=180 Treatment group=296
Lugo-Neris, Jackson, and Goldstein	Facilitating Vocabulary Acquisition of Young English Language Learners	2010	Language, Speech, and Hearing Services in Schools, Vol.41, p.314-327	USA	Summer education program	EAL only	Spanish	22 across two treatment groups
Macaruso and Rodman	Benefits of Computer-Assisted Instruction to Support Reading Acquisition in English Language Learners	2011	Bilingual Research Journal: The Journal of the National Association for Bilingual Education, 34: 301-315	USA	Bilingual classes	EAL only	Spanish	37 in control group and 29 in treatment group
Mancilla-Martinez	Word Meanings Matter: Cultivating English Vocabulary Knowledge in Fifth-Grade Spanish-Speaking Language Minority Learners	2010	Teachers of English to Speakers of Other Languages, Inc. (TESOL), Vol.44, No.4, p.669-699	USA	Mainstream classes grade 5	EAL combined with 1 English/other participant in treatment group and 3 English/other participants in contrast group	Primarily Spanish	Control group=25 Treatment group=24
Matsumura, Garnier,	Investigating the Effectiveness of a	2010	The Elementary	USA	Mainstream classes	EAL, NS, and teachers. EAL and	n/a	TAKS group=1,229 DRP group=896

Correnti, Junker, and Bickel	Comprehensive Literacy Coaching Program in Schools with High Teacher Mobility		School Journal, Vol.111, No.1, p.35-62		grades 4 and 5	NS combined in both treatment groups. Some EAL results presented separately		
Olson, and Land	A Cognitive Strategies Approach to Reading and Writing Instruction for English Language Learners In Secondary School	2007	Research in the Teaching of English, Vol.41, No.3, p.269-303	USA	Grades 6 to 12 transition English language development classes, standard arts classes. 7% of participants in English-only programs	EAL, NS, and teachers. EAL and NS combined in treatment and control groups. Some EAL results presented separately	Spanish	94 teachers and approximately 2,000 pupils
Proctor, Dalton, Uccelli, Biancarosa, Mo, Snow, and Neugebauer	Improving comprehension online: effects of deep vocabulary instruction with bilingual and monolingual fifth graders	2011	Reading and Writing, 24:517-544	USA	Mainstream grade 5 classes. Intervention performed in computer laboratory	EAL and NS combined in control and treatment groups. EAL results presented separately	Spanish	Control group=111 (of which 59 EAL) ICON group=129 (of which 59 EAL)
Schoenbrodt, Kerins, and Gesell	Using Narrative Language Intervention as a Tool to Increase Communicative Competence in Spanish-Speaking Children	2010	Language, Culture, and Curriculum, Vol.16, No.1, p.48-59	USA	Participants attended mainstream public elementary school. Treatment was delivered in an after-school tutoring program for children aged 6 to 11	EAL only	Spanish	Treatment group one=6 Treatment group two=6
Short, Fidelman, & Louguit	Developing Academic Language in English Language Learners	2012	TESOL Quarterly 46(2), 334-	USA	Content area and ESL classes	EAL only	Spanish 24%-41%, Polish 6%, Arabic 5%-7%,	Comparison group=176, Treatment group=386

	Through Sheltered Instruction		361				Gujurati, Turkish, Pilipino, Albanian	
Snow, Lawrence, and White	Generating Knowledge of Academic Language Among Urban Middle School Students	2009	Journal of Research on Educational Effectiveness, 2:325-344	USA	Mainstream grade 6 to grade 8 classes	EAL and NS. EAL and NS combined in treatment and comparison groups. EAL results presented separately	n/a	Comparison group=133 Treatment group=256
Solari, and Gerber	Early Comprehension Instruction for Spanish-Speaking English Language Learners: Teaching Text-Level Reading Skills While Maintaining Effects on Word-Level Skills	2008	Learning Disabilities Research & Practice, Vol23. Issue 4, p.155-168	USA	Mainstream kindergarten classes	EAL only	Spanish	82 across three treatment groups and categorization by at risk or not at risk of reading difficulties
Spycher	Learning Academic Language through Science in Two Linguistically Diverse Kindergarten Classes	2009	The Elementary School Journal, Vol.109, No.4, p.359-379	USA	Mainstream kindergarten class	EAL and NS combined in treatment and control groups. EAL results presented separately	Spanish	Control group=20 Intervention group=19
Tong, Irby, Lara-Alecio, Yoon, and Mathes	Hispanic English Learners' Responses to Longitudinal English Instructional Intervention and the Effect of Gender: A Multilevel Analysis	2010	The Elementary School Journal, Vol.110, No.4, p.542-566	USA	Mainstream kindergarten to grade 2 classes	EAL only	Spanish	Control group=112 Treatment group=84
Tong, Lara-Alecio, Irby, Mathes, and Kwok	Accelerating Early Academic Oral English Development in Transitional Bilingual and Structured English Immersion Programs	2008	American Educational Research Journal, Vol.45, No.4, p.1011-1044	USA	Transitional bilingual education and structured English immersion classes	EAL only	Spanish	534 over 4 treatment groups

Townsend, and Collins	Academic vocabulary and middle school English learners: an intervention study	2009	Reading and Writing, 22:993-1019	USA	After school class	EAL only	Spanish=68%, Vietnamese=16%, Japanese=8%, Arabic=5%, Gujarati=3%	Treatment A=20 Treatment B=17
Troia	Migrant Students with Limited English Proficiency: Can Fast ForWord Language(trademark) Make a Difference in Their Language Skills and Academic Achievement?	2004	Remedial and Special Education, Vol.25, No.6, p.353-366	USA	Mainstream public elementary classes grades 1 to 6	EAL only	Spanish	Control group=92 Treatment group=99
Vadasy and Sanders	Efficacy of Supplemental Phonics-Based Instruction for Low-Skilled Kindergarteners in the Context of Language Minority Status and Classroom Phonics Instruction	2010	Journal of Educational Psychology, Vol.102, No.4, p.786-803	USA	Mainstream kindergarten classes	EAL and NS combined in both treatment and control groups. EAL results presented separately	Spanish=49%, Vietnamese=15% Somali=11%, Chinese=6%, Tagalog=3%	Control group=81 (of which 46 EAL) Treatment group=67 (of which 38 EAL)
Vadasy, and Sanders	Two-year follow-up of a code-oriented intervention for lower-skilled first-graders: the influence of language status and word reading skills on third-grade literacy outcomes	2013	Reading and Writing, Vol26, p.821-843	USA	Mainstream first grade classrooms	EAL and NS separated in treatment and control groups. EAL results presented separately	Spanish=64%, Vietnamese=9%, Chinese=6%, Somali=5%, Amharic=2%, Arabic=2%, French=2%, Russian=2%, Samoan=2%, Tagalog=2%, Cambodian=1%, Oromo=1%, Punjabi=1%, Tigrigna=1%	EAL group=95 (of which 48 in treatment condition) Non EAL group=85 (of which 43 in treatment condition)

Vaughn, Cirino, Linan-Thompson, Mathes, Carlson, Hagan, Pollard-Durodola, Fletcher, and Francis	Effectiveness of a Spanish Intervention and an English Intervention for English-Language Learners at Risk for Reading Problems	2006	American Educational Research Journal, Vol.43, No.3, p.449-487	USA	Bilingual Grade 1 classes	EAL only	Spanish	171 across four groups
Vaughn, Linan-Thompson, Mathes, Cirino, Carlson, Pollard-Durodola, Cardenas-Hagan, Francis	Effectiveness of Spanish Intervention for First-Grade English Language Learners at Risk for Reading Difficulties	2006	Journal of Learning Disabilities, Vol.39, No.1, p.56-73	USA	Bilingual Grade 1 classes	EAL only	Spanish	Comparison group=33 Treatment group=31