Reading Wars or Reading Reconciliation?: a critical examination of robust research evidence, curriculum policy, and teachers' practices for teaching phonics and reading.

Dominic Wyse and Alice Bradbury
IOE, UCL's Faculty of Education and Society
University College London
Helen Hamlyn Centre for Pedagogy (0 to 11 Years) (HHCP)

Corresponding author: d.wyse@ucl.ac.uk

NB: This document is a pre-publication copy that might contain errors. Please refer to the version published in the peer-reviewed journal *Review of Education*. https://doi.org/10.1002/rev3.3314

Abstract

Teaching children to read is one of the most fundamental goals of early years and primary education worldwide, and as such has attracted a large amount of research from a range of academic disciplines. The aims of this paper are, a) to provide a new critical examination of research evidence relevant to effective teaching of phonics and reading in the context of national curricula internationally; b) to report new empirical findings relating to phonics teaching in England; and c) examine some implications for policy and practice. The paper reports new empirical findings from two sources: 1. a systematic qualitative meta-synthesis of 55 experimental trials that included longitudinal designs; 2. a survey of 2205 teachers. The paper concludes that phonics and reading teaching in primary schools in England has changed significantly for the first time in modern history, and that compared to other English dominant regions England represents an outlier. The most robust research evidence, from RCTs with longitudinal designs shows that the approach to phonics and reading teaching in England is not sufficiently underpinned by research evidence. It is recommended that national curriculum policy is changed and that the locus of political control over curriculum, pedagogy and assessment should be re-evaluated.

Context and Implications

• Rationale for this study
Teaching children to read is one of the most important elements of primary education because it is fundamental to children's educational development. For this reason it is vital that the teaching of reading, and curriculum policies on reading, are informed by robust research.

• Why the new findings matter
If children are not being taught to read in the most appropriate way, because curriculum policy and teaching practices are not informed sufficiently by robust research evidence, then children's education will not be as effective as it should be.

• Implications for practitioners, policymakers, researchers.

The outcomes of the survey of teachers in England, and the new analysis of systematic reviews and meta-analyses, and randomised controlled trials with longitudinal designs, reported in the paper show the need for changes to the teaching of reading and to national curriculum policy on the teaching of reading. The teaching of phonics and reading in curriculum policy and practice should more closely reflect the evidence that contextualised teaching of reading, or balanced instruction, is the most effective way to teach reading.

Keywords
Phonics; reading; policy; assessment.

Teaching children to read is one of the most fundamental goals of early years and primary education worldwide. Reading is vital for pupils' cognitive development if they are to progress successfully throughout education, as birth cohort studies have shown (Sullivan & Brown, 2013), and reading gives children access to nearly all areas of the school curriculum as children move from early years through to secondary education, and beyond. Because reading is so fundamental to children's learning it is also one of the key ways in which the quality of education is measured. For example, settings and schools are held to account by parents and by wider society for the progress in reading that children make as a result of the teaching in those settings. In some countries governments have policies for education monitoring processes that may include statutory national assessment and testing programmes, and inspections of education settings taking account of such monitoring. At the international level, the education and curriculum policies in different countries and regions are compared on the basis of the levels at which children and young people are able to read.

The importance that societies place on children learning to read is also seen in the large amount of research, from a range of disciplines such as education, psychology, neuroscience, history, philosophy, and interdisciplinary work, much of which is devoted to understanding how reading might be taught more effectively, and which curriculum policies might be advantageous. An aspect of reading that has attracted a wealth of research has focused on how children can be taught to understand the ways in which graphemes represent the phonemes of speech: in alphabetic languages this is often referred to as the 'alphabetic code'. If children do not learn to decode the ways that letters represent phonemes to make the meanings in words, and learn the ways that words are combined to make sentences and whole texts, they will not learn to read. In other words if they do not learn to fully decode written language they cannot comprehend the meanings composed by writers.

The societal importance placed on learning to read has also been evident in an unusually prolonged and at times fierce debate about how to teach reading. Jean Chall (1983) famously depicted the argument about the best ways to teach reading as the 'great debate', and the description 'the reading wars' quickly followed. It is not known who first used the phrase 'the reading wars' although it has been attributed to debates in the USA (Castles, Rastle and Nation, 2018). The phrase is still prevalent today as part of academic debates, for example in the research paper by Torgerson et al (2019), in the Castles (op cit.) paper itself, and in a paper by Solity (2020). The phrase also continues to appear in the mainstream media for example in a recent

article about reading policy in Mississippi in *The Economist* titled “The reading wars; Literacy” (Economist Intelligence Unit, 2021).

Although the debate started by Chall has been seen as a binary one between ‘top-down’ approaches to teaching reading versus ‘bottom-up’ approaches, the main variants of the debate have since become focused on three main orientations to the teaching of reading:

1. Synthetic phonics: a focus first and foremost on teaching children about *phonemes and letters*. As part of this approach at key moments in the teaching programme phonics teaching is separate from practising reading with whole texts. In the early stages of the approach in particular, whole text reading is required to be done with ‘decodable’ books which are reading scheme/basal books with vocabularies controlled to enable repetition of key words learned during the phonics programme.

2. Whole language: a focus first and foremost on *whole texts*, ‘real’ books (trade books created by authors as part of standard publication practices), that it is theorised children will enjoy more and will be motivated by. The whole language approach is driven by reading for meaning. Phonics teaching, and other aspects of reading, are taught in a relatively non-systematic way, and carried out through examples related to the real books being read.

3. Balanced instruction: a focus first and foremost on the *balance* between teaching based on use of whole texts and systematic teaching about the alphabetic code *and* also other linguistic features. With this approach the importance of comprehending the meaning of written language is carefully balanced with the acquisition of a range of skills and knowledge. Lessons make explicit links between phonics teaching and other linguistic aspects with whole texts, which are often a combination of real books and reading scheme books with controlled vocabularies.

Synthetic phonics is sometimes categorised within the term *systematic phonics* but systematic phonics incudes a wider range of variants of approaches to teaching about phonemes and letters, including for example analytic phonics approaches. The development of children’s phonological awareness (PA) is considered by some people to be part of phonics teaching but others regard PA as a separate approach. Later in the paper we link these three orientations with key theories and other publications relevant to learning to read and teaching reading. As will become evident in this paper the debate in England is concerned mainly with synthetic phonics as a variant of systematic phonics approaches.

Every category set of this kind can be criticised on the grounds that teaching approaches do not necessarily neatly fit into categories. For example, almost no approach to teaching reading described as synthetic phonics first and foremost completely excludes a focus on whole texts, at least as part of all the reading teaching activities in a classroom. Equally no approach described as whole language completely excludes attention to letters and phonemes. Never-the-less these three categories do help to distinguish what are important aspects relevant to the study of effective teaching of reading. The extent to which different emphases on whole texts and on letters and phonemes as part of teaching programmes are more or less

effective is important to advancing knowledge about effective teaching of reading. For example, if children are more motivated by approaches that focus on 'real books', those books created by authors to engage young readers as part of standard 'trade' publishing processes, then evidence that a whole language approach was the most effective way to teach reading would provide empirical support for addressing both children's motivation for books *and* learning about phonemes within the same teaching approach and in reading policies. Children, teachers and society more generally are likely to benefit if it can be determined which of these overall approaches to teaching is most effective.

The effectiveness of the teaching of reading is of paramount importance for education systems, and effectiveness should be determined through rigorous research. However, the extent to which relevant research evidence is reflected in curriculum policy and teachers' practice is affected by a range of influences. For example political ideology can be an influence on the development of national curricula and teaching methods, sometimes contrary to the research evidence (e.g. as documented by James, 2012). Another important influence on curriculum and teaching is the nature of national assessment systems which may influence teachers' approaches and schools' policies. For instance it has been shown that In England the 'high stakes' nature of statutory assessments at age five/six, six/seven and age ten/eleven have had an impact on teaching (Bradbury, 2018; Bradbury, Braun and Quick, 2021). To determine what is effective teaching of reading, and hence to be reflected in education policy, requires consideration of a range of evidence in order to identify the most robust sources. It also requires understanding of the historical context of debates; identification of relevant theories of reading and their appropriateness as models for teaching and policy; and a clear practical understanding of how robust research might be adapted at large scale in schools and classrooms.

This aims of this paper are, a) to provide a new critical examination of research evidence relevant to effective teaching of phonics and reading in the context of national curricula internationally; b) to report new empirical findings relating to phonics teaching in England; and c) examine some implications for policy and practice. The original contribution to knowledge made by the paper is through new findings in each of the following four areas: 1) analyses of approaches to teaching reading in national curriculum models in relation to international and national pupil assessment data; 2) an evidence-based account of effective teaching of phonics and reading based on a systematic critical synthesis of the most relevant and methodologically robust systematic reviews, meta analyses, and experimental trials; 3) an account of teachers' views about approaches to teaching reading and to assessment of reading, based on a recent survey of teachers in England; 4) insights into how the teaching of reading could be optimised including through curriculum and assessment policies.

The first section of the paper locates the three orientations of synthetic phonics, whole language, and balanced instruction in relevant theories of reading teaching. This section is followed by a brief outline of some key moments in the debates about reading particularly in the UK and the USA. The next section contextualises the paper's frames of reference in national curriculum and assessment in a selection of English language-dominant nations. This section includes an examination of

correlations between orientations to reading in national curriculum texts and outcomes of pupil tests in the Programme for International Student Assessment (PISA) and the Progress in International Reading Literacy Study (PIRLS).

The main part of the paper reports the findings from two new analyses: 1. a systematic qualitative meta-synthesis (SQMS) of 55 randomised controlled trials (RCTs) that had longitudinal research designs and which compared approaches to the teaching of phonics and reading; 2. a questionnaire survey of 2205 teachers in England. The paper concludes with evidence-based recommendations for teaching phonics and reading, including for national curriculum and assessment policies.

Theories underpinning the three orientations to the teaching of reading

The three orientations to teaching reading: synthetic phonics, whole language, and balanced instruction, can each be linked to different theories that to varying degrees underpin them. One of the most highly regarded theories, from the perspective of psychology, that has been linked with advocacy for the *synthetic phonics* orientation to teaching reading is the *simple view of reading* (SVR. Gough and Tunmer, 1986). Gough and Tunmer's seminal paper from 1986 theorised that: Reading = Decoding × Comprehension (R = D × C) which subsequently has been shown in countless studies to be relevant to understanding children's reading development (e.g. see Savage, et al. 2015). It is important to note that Gough and Tunmer did not regard this model as a model of teaching, it is a model of children's reading development, although in their paper Gough and Tunmer acknowledged that reading development and the teaching of reading are interconnected.

One of the issues at the heart of our exploration in this paper is how reading research and reading theories can be most appropriately *interpreted* in relation to classroom practice and curriculum policies. There are many robust academic theories about reading but for an academic theory to have relevance to the practice of teaching reading in schools, and to have relevance for policy, it requires interpretation and acceptance by practitioners and policy makers that the theory is relevant to classroom practice as a result of testing the theory through teaching. The SVR became influential in policy on reading in England as a result of the report led by Sir Jim Rose, known as the Rose Report (Rose, 2006). This influence of Gough and Tunmer's SVR was not direct, it rested on two connected interpretations of the SVR. The first interpretation was in the appendix to the Rose Report written by Morag Stuart and Rhona Stainthorp as part of their critique of two other reading models: a) Marie Clay's model of reading *cueing* that was part of Clay's *Reading Recovery* approach originating in New Zealand (Clay, 1979); and b) The *Searchlights Model* that had emerged as part of the policy developments that resulted in the National Literacy Strategy (NLS; Department for Education, 1998).

The second interpretation of the SVR was done by Jim Rose himself, also in the Rose Report, in making recommendations for reading policy derived in part from the Stuart and Stainthorp appendix. As a result of the two interpretations of the SVR that featured in the Rose Report. The following recommendation was the catalyst for a series of changes to curriculum and pedagogy in England:

> What best practice should be expected in the teaching of early reading and synthetic phonics [*sic*.]
> …
> • High quality, systematic phonic work as defined by the review should be taught discretely. The knowledge, skills and understanding that constitute high quality phonic work should be taught as the prime approach in learning to decode (to read) and encode (to write/spell) print. (Rose, 2006, p.70, underline added)

The *whole language* orientation to reading teaching was theorised in the USA in seminal work by Ken Goodman, Yetta Goodman, and Frank Smith. The work by the Goodmans is probably the most well-known theory of whole language. One of the components of this theory was that learning to read was almost the same as learning to talk, so was seen as a natural process. Ken Goodman's influential paper of 1967 published in *The Journal of the Reading Specialist*, now called *Literacy Research and Instruction*, is regarded as central to the debate. The motivation of his paper was to replace "pre-existing, naïve, common sense notions" to "offer a more viable scientific alternative" (Goodman, 1967, p. 126). Goodman's perspectives on reading were informed by his early career experience as a teacher: his theory of reading was based on his experience of researching how children read in school. The idea of analysing children's word reading errors when reading aloud, which he called *miscue analysis*, to gain insight into their mental processing provided teachers with a tool to think more deeply about children's reading (Clay *op cit.* also used the idea of miscues in her Reading Recovery approach). The central most controversial element of Goodman's theory is summed up in this quote from the original paper:

> More simply stated, reading is a psycholinguistic guessing game. It involves an interaction between thought and language. Efficient reading does not result from precise perception and identification of all elements, but from skill in selecting the fewest, most productive cues necessary to produce guesses which are right the first time. The ability to anticipate that which has not been seen, of course, is vital in reading, just as the ability to anticipate what has not yet been heard is vital in listening. (Goodman, 1967, p. 127, underline added)

Claims were made, in support or in opposition to Goodman's theory, on the basis of whether studies of eye-movements while reading proved that people attend to every letter when reading words (e.g. Perfetti, 1995) or whether they only attend to some letters in a word in order to read it. Irrespective of these arguments about 'precise perception' the lack of attention to any *systematic* phonics teaching in Goodman's theories can, with hindsight, be seen as a weakness. Since Goodman published his theory multiple research studies have confirmed the beneficial effect of teaching children in the early stages of learning to read about letters and the speech sounds that they represent, as the SQMS later in this paper shows.

In England the equivalent to whole language was *the real book approach*. The real book approach emphasises the importance of high-quality children's literature first and foremost as a way to engage children with reading. Reading skills are taught as needed based on one-to-one assessments of children reading aloud and assessments of their reading in small group activities. The theory of the real book approach was underpinned by the work of Margaret Meek, advanced in her book

How Texts Teach What Readers Learn (Meek, 1988), and in the more practical work of Liz Waterland whose book was written as a reaction against what she saw as the narrow focus on phonics at the time something which was demotivating the children that she taught (Waterland, 1985).

The *balanced instruction* orientation to reading was, in the USA, underpinned particularly by the theories of Michael Pressley in his book *Reading Instruction that Works: The Case for Balanced Teaching*. Pressley (2006) argued that neither an exclusive focus on whole language or on what he called reading skills was sufficient. Presley's balanced teaching combined attention to whole texts, for example through teaching to help children's reading comprehension, with teaching of reading skills. However systematic teaching about letters and phonemes was not a particularly strong feature in Pressley's seminal book.

In the UK the phrases balanced instruction or balanced teaching were not particularly prominent nor associated with one theorist. Several authors made contributions that we can describe as making a case for balanced instruction for example Roger Beard (Beard, 1993); David Wray and Jane Medwell (Wray and Medwell, 1994); Dominic Wyse (Wyse and Jones, 2001) and Kathy Hall (Hall, 2003). The probable reason for the greater range of sources for balanced instruction in England is that *de facto* the orientation draws on a wider range of research, sometimes from multiple disciplines. In part, balanced instruction was also originally based on a reaction against extreme advocacy for either whole language or synthetic phonics.

A key test for the validity of reading theories is the extent to which the teaching methods that link with the theory are effective in practice when used by teachers as part of early years and primary education. Ultimately the theories of reading most appropriate as part of the evidence base for curriculum policies are those that not only make an original and lasting contribution to understanding children's reading development but also to effective teaching of reading. However, determining effectiveness also requires robust evidence resulting from robust research designs that can be linked with appropriate reading theories. Qualitative enquiry has much to offer understanding of the detail and depth of reading teaching in early years settings and primary schools but robust experimental designs, including the careful selection of appropriate test measures, are needed to determine the effectiveness of teaching approaches at sufficient scale.

Key moments in the history of teaching and assessing reading

The history of the English language in England includes some of the first books worldwide aimed to improve children's learning of the English language, for example head teacher Richard Mulcaster's *THE FIRST PART OF THE ELEMENTARIE WHICH ENTREATETH CHEFELIE OF THE right writing of our English tung* first published in 1582. Debates about teaching reading, including attention to the letters of the alphabet and their associated sounds in oral language, are part of this long history.

> “Therefore let the scholler [pupil], being thus traded [schooled] from letters to syllables of one Consonant: from syllables of one Consonant, to syllables of many Consonants: from syllables of many Consonants, to words of many syllables; proceede to sentences.” (Michael, 1984, p. 57)

The quote above is from one of the ‘spelling books’ that were common in England in 1610, and represents some of the first published writing about approaches to teaching literacy. The approach implied by the quote has some elements of what we now call systematic phonics teaching: the smaller units of letters and single syllables were to be taught first followed by the larger units such as multiple syllable words and finally to sentences.

Related to developments in England, as a result of migration patterns from England to other countries, a long history can also be traced in the USA. Although Joseph Neef was originally from Alsace in France, he is recognised not only for the first education method book in the USA, published in 1808, but also the first book to detail an approach to teaching phonics, published in 1813. In common with the early spelling book from England, the idea of the separation of phonics from the experience of books was a key feature of Neef’s approach. Historical evidence of accounts from Neef’s pupils who experienced his approach suggests that until children were aged 10 books were not deemed suitable for them or their education (Brooks, 2021).

In the 20th century the debates about the teaching of reading were ignited once more in the USA. The seminal text in the debate was Jean Chall’s (1983) book *Learning to Read: The Great Debate*, which was first published in the 1960s. Chall framed the debate as an opposition between ‘bottom-up’ approaches to the teaching of reading, which emphasised first and foremost teaching of the alphabetic code, versus ‘top-down’ approaches that emphasised first and foremost a focus on meaning and comprehension of texts. The reading debates have included some extreme views for example in this extraordinary email exchange from the USA about the use of federal funds for balanced instruction or whole language versus ‘scientifically based’ approaches to teaching reading. The denigration of people who advocated whole language even alludes to violence:

> Beat the [expletive deleted] out of them in a way that will stand up to any level of legal and [whole-language] apologist scrutiny. Hit them over and over with definitive evidence that they are not SBRR [scientifically based reading research], never have been and never will be. They are trying to crash our party and we need to beat the [expletive deleted] out of them in front of all the other would-be party crashers who are standing on the front lawn waiting to see how we welcome these dirtbags. (Cummins, 2007, p.566)

More recently the following quote from *The Economist* reveals a similarly unequivocal position that phonics is the best way to teach and other methods are wrong:

> PHONICS, WHICH involves sounding-out words syllable by syllable, is the best way to teach children to read. But in many classrooms, ff-on-ics is a dirty sound. Kymyona Burk, who implemented Mississippi's statewide literacy

> programme, says that some teachers have had to sneak phonics teaching materials into the classroom, like some kind of samizdat. Teaching reading any other way is "malpractice", says Ms Burk. And yet for reasons that include politics, partisanship and personal experience, most American children are taught to read in a way that study after study has found to be wrong. (Economist Intelligence Unit, 2021, online)

A feature of the teaching of reading in England in the 20th century was a long period of relatively little change in classroom practice, in spite of policy developments, until policy changes were introduced that began with the publication of 'The Rose Report' in 2006. Alexander (1995) summarised a series of in-depth studies of primary education over a ten-year period and concluded that primary teaching had undergone a change in the culture of schools but maintained relative continuity in approaches to teaching in spite of the introduction of the England's National Curriculum in 1988. Further evidence of stability in reading teaching methods came, for example, from a report by Her Majesty's Inspectorate (HMI) that drew on evidence from visits to 120 primary schools and concluded that, "The teaching and learning of reading were observed in 470 classes and over 2,000 children read aloud to HMI … phonic skills were taught almost universally and usually to beneficial effect" (HMI, 1990 p. 2, underline added). A research study carried out by The National Foundation for Educational Research (NFER) in 1992, that based its findings on a random selection of 234 primary schools across England and Wales resulting in survey returns from 122 head teachers and visits to some of the schools to observe teaching, concluded that their results replicated the findings from HMI (Cato et al. 1992). With regard to the teaching of reading prior to 2006 the teaching of reading in most classrooms in England is best described as balanced instruction, in which some phonics teaching has always been part of the teaching of reading typically for children in the infant years (aged five to seven) although not necessarily 'systematic phonics' instruction.

England's first national curriculum of 1988 was subject to periodic review then in 1998 was augmented by the NLS. The introduction of the NLS increased the specification of phonics teaching to a level not previously seen in national curriculum documentation. However in 2006 the Rose Report recommended that there should be even more emphasis on phonics teaching. The NLS was modified and renamed the Primary National Strategy – Primary Framework for Literacy and Mathematics (PNS): the PNS required that "high-quality, systematic phonic work should be taught discretely" (Department for Education, 2008, p. 7, underline added). One of the developments following the Rose report was the development of the *Letters and Sounds* programme of phonics teaching (Department for Education and Skills, 2007).

A further intensification of phonics teaching, and another major curriculum policy change, occurred in 2012 with the introduction of an additional test to be taken by all Year 1 (age 5 to 6) children in England in state-funded primary schools: the Phonics Screening Check (PSC). The PSC consists of 40 words and pseudo-words (phonetically regular letter combinations represented as plausible words) which children are asked to read out loud as part of a one-to-one assessment with their teacher. Children who do not meet the expected standard in the test have to be retested in Year 2 (age 6 to 7).

A range of associated means to ensure compliance to synthetic phonics teaching were also implemented by government, for example the requirement for the national inspectorate Ofsted to use outcomes in the PSC as one of its means to judge schools' effectiveness as part of school inspections (Ofsted, 2019, point 335). The 2019 Inspection Framework also included the requirement that inspectors assess to what extent "In Reception, staff teach children to read systematically by using synthetic phonics and books that match the children's phonic knowledge" (Ofsted 2019, point 325). Other developments linked with the intensification of synthetic phonics teaching from 2006 onwards included, for the first time, the Department for Education (DfE) reviewing and approving published phonics teaching schemes, a process that allowed further control of teaching by the DfE. The process for approving programmes was updated in 2021 resulting in the *Letters and Sounds* programme (a free resource developed by the Department for Children, Families and Schools during the New Labour government term) being removed from the approved phonics schemes list: the DfE explanation was that the scheme "isn't a full Systematic Synthetic Phonics (SSP) programme" (Department for Education, 2021, online).

In 2013 the national curriculum that had been developed by the New Labour government, led by what was at the time the Qualifications and Curriculum Authority (QCA), was disapplied, then archived in January 2014 (Department for Education, 2014), as a result of the Conservative Liberal-Democrat coalition government coming to power in 2010. It was eventually to be replaced by the 2014 national curriculum developed by the Conservative government which further intensified the specifications for synthetic phonics teaching.

A significant feature of England's curriculum policies that has an impact on teaching is the nature of national assessment processes. The most relevant of these to our analyses in this paper is the Phonics Screening Check (PSC), that is taken by all Year 1 (age six to seven) children in England's schools that receive state funding. Other relevant national assessments include the range of statutory assessments for children in Reception (aged four to five), in Year 2 (aged six to seven) and in Year 6 (aged ten to eleven). Statutory assessments in England have been subject to many changes both to the range and nature of assessments as Table 1 shows.

Table 1: Statutory assessment scores mapped against main literacy policy changes in England.[1]

	PSC* Y1	PSC Y2	KS1 R TA	KS1 W TA	E Test	E TA	KS2 R Test	KS2 R TA	GP&S	KS 2 W TA	Main Literacy Policy Changes in England
2020											Covid disrupts all testing and examinations
2019	82	91	75	69	na	na	73	na	78	78	English reading statutory TA not required from 2018-19 school year onwards.
2018	82	92	75	70	na	na	75	80	78	78	Changes made within the 2017/18 writing TA frameworks mean that judgements in 2018 are not directly comparable to those made using the previous interim frameworks in 2016 and 2017.
2017	81	92	76	68	na	na	72	79	77	76	
2016	81	91	74	65	na	na	66	80	73	74	"Pupils were assessed against the new more challenging curriculum, which was introduced in 2014, for the first time this year." Figures for 2016 KS1 and KS2 not comparable with previous years.
2015	77	90	90	88	na	na	89	90	80	87	

[1] All assessment scores and quotes are taken from national statistics reported in Department for Education annual reports on assessment outcomes.

2014	74	88	90	86	na	88	89	89	76	85	2014 national curriculum introduced. "The majority of this national curriculum was introduced in September 2014, with English and maths coming into force for all year groups from September 2016."
2013	69	85	89	85	na	87	86	87	74	83	Phonics Screening Check in place for one year. GP&S new this year.
2012	58	na	87	83	85	85	87	86	na	See English score	Phonics Screening Check first used. "There were significant changes to the Key Stage 2 assessment arrangements in 2012 that affect this release. In 2012, schools were no longer required to administer a writing test and submit this for external marking ... Therefore, this year's figures for English cannot be compared to the figures for English that were published in earlier years, which were based solely on tests."
2011	na	na	85	81	82	81	84	na	na	75	
2010	na	na	85	81	80	81	83	na	na	71	Conservative-Liberal Democrat government proposed to replace New Labour national curriculum with a new national curriculum. Primary National Strategy - Literacy also discontinued. Level 4 or above was expected standard for statutory assessments.
2009					80	79					

2008					81	79					Primary National Strategy - Literacy in place for one year
2007					80	78					National Literacy Strategy Ends. Primary National Strategy - Literacy Begins ("high-quality, systematic phonic work should be taught discretely" DfE, p. 7).
2006					79						Jim Rose Final Report published. Advocates more systematic phonics: "High quality, systematic phonic work as defined by the review should be taught discretely. The knowledge, skills and understanding that constitute high quality phonic work should be taught as the prime approach in learning to decode (to read) and encode (to write/spell) print."
2005					79						
2004					78						
2003					75						
2002					75						
2001					75						
2000					75						
1999					71						National Literacy Strategy in place for one year.
1998					65						National Literacy Strategy Begins
1997					63						National Literacy Project is precursor to National Literacy Strategy.

1996					57						
1995					49						
1988											First national curriculum in England
***Key**: PSC Phonics Screening Check; KS Key Stage; R Reading; TA Teacher Assessment; E English; W Writing; GP&S Grammar, Punctuation and Spelling.											

An important possible use of national assessment data should be the ability to evaluate long-term trends in the impact of significant changes in curriculum and pedagogy on children's learning (related to their birth cohort and hence experience of different curriculum changes). Unfortunately England's national assessment data is not sufficiently reliable for analysis of long trends of national curriculum change because the nature of the tests has changed so frequently, (for example see the notes in the rows for 2016 and 2018 of Table 1). The changes in the nature of statutory assessments mean that too frequently it is not possible to compare like with like in order to evaluate major curriculum changes. For shorter trends, say of three years or so, following significant changes to curriculum and/or assessment requirements the assessment outcomes generally show gradual increases in test scores, probably as teachers become more adept at teaching to the test and new curricula, and finally plateaus in scores (also see Wyse and Torrance, 2009). The plateaus also reflect to some degree ceiling effects related to the normal distribution of results in a given assessment.

One useful aspect of these relatively short-term trends in data is to underline how important it is to note changes to assessment outcomes about one year after the introduction of new curricula and/or teaching method, when the change has become relatively established but not sufficiently established to allow for teaching to the test to have an effect.

An international comparative context for national curricula and assessment.

One of the underlying priorities of this paper is to form conclusions based as much as possible on longitudinal evidence. As far as the assessment of reading is concerned there have been more cycles of the Programme for International Student Assessment (PISA) tests than Progress in International Reading Literacy Study (PIRLS). Another difference between PISA and PIRLS is that PISA measures the performance of 15-year-old pupils. This provides a stronger longitudinal measure than other international comparative tests in two senses: a) there have been six cycles of PISA since it was first carried out in 2000, and b) PISA's testing of 15-year-old pupils is a measurement point just before a significant proportion of young people may finish their schooling so it enables reflection on correlations, for example, between early reading curricula and final school outcomes for significant numbers of pupils.

Preliminary work for this paper included review and discourse analysis of the national curriculum texts in regions where the English language is dominant, as a way to contextualise and compare our focus on England in relation to specifications for the teaching of reading in national curricula. In addition to the English language being dominant the selection of countries was also made on the basis of those that have consistently been present in the PISA and PIRLS assessments. As well as identifying those countries that have consistently scored highly in pupil assessments the aim for the discourse analysis was to categorise the orientation to reading represented in the wording of the national curriculum texts in order to explore possible correlations between orientation to reading in the national curricula and total scores in PISA and PIRLS. Although approaches to teaching reading in classrooms and individual schools will vary to some degree from the national and state

curriculum text requirements, the texts themselves are important because they represent policy goals which influence teaching and learning.

National level curriculum texts were selected for the discourse analysis, but in some regions curriculum is also developed at state level. State level curricula were also checked for consistency in relation to the orientation to phonics and reading teaching in the national level texts. As Canada has been the strongest performer of English language dominant nations in PISA and PIRLS, analyses were also done of the curriculum texts of three high performing Canadian states: Alberta, Ontario and Quebec. Australia has not reported state level outcomes in PISA or PIRLS. The USA has not consistently reported state level outcomes in PISA and PIRLS but we did include one additional analysis, of Massachusetts.

The discourse analysis focused on the following aspects of the curriculum texts: rationale for the approach to teaching reading; aims of reading teaching; description of programme of study for phonics and reading. The frequency of the stem ‘phon’ in each curriculum text was also counted. Analysis of these frequencies included checking that each instance was a new point semantically not simply a repetition of a sentence applied across multiple grades or year groups, nor only an entry in a glossary.

Table 2 reports the outcomes of the discourse analysis of curriculum texts at national level, and some at state level, of regions included in PISA and PIRLS for which English is a dominant language, and attributes one of our orientations to reading to each curriculum text.

[Blank space due to page layout]

Table 2: Discourse analysis of national and state curricula of English dominant nations.

Country/State	Curriculum Subject Title	Rationale for approach to reading teaching	Aims of reading teaching	Description of programme of study for phonics and reading	Frequency of stem "phon"	Orientation to Reading
England	English	Reading The programmes of study for reading at key stages 1 and 2 consist of two dimensions: § word reading § comprehension (both listening and reading). It is essential that teaching focuses on developing pupils' competence in both dimensions; different kinds of teaching are needed for each. p. 15	"§ read easily, fluently and with good understanding § develop the habit of reading widely and often, for both pleasure and information § acquire a wide vocabulary, an understanding of grammar and knowledge of linguistic conventions for reading, writing and spoken language" p. 14	Has statutory and non-statutory elements. Starts with very strong emphasis on phonics teaching outlined in many pages of detailed statutory requirements for reading and for spelling. Reading comprehension sections are listed after phonics section for each year group.	78	Synthetic phonics

Canada, Alberta	English Language Arts	"Reading and writing are powerful means of communicating and learning. They enable students to extend their knowledge and use of language, increase their understanding of themselves and others, and experience enjoyment and personal satisfaction. Reading provides students with a means of accessing the ideas, views and experiences of others. By using effective reading skills and strategies, students construct meaning and develop thoughtful and critical interpretations of a variety of texts." (p. 2)	"Five general student outcomes serve as the foundation for the program of studies. General outcomes are broad statements identifying the knowledge, skills and attitudes that students are expected to demonstrate with increasing competence and confidence from Kindergarten to Grade 12." (p. 3) "listen, speak, read, write, view and represent to explore thoughts ..." (p. 4) are applied to all five general outcomes.	General outcome 2 includes "2.1 Use Strategies and Cues" (p. 18). These begin with "prior knowledge" then "Use comprehension strategies". Later there is "use textual cues". Later still "Use phonics and structural analysis" (p. 26). These are general PoS to apply flexibly not a systematic programme.	10. 26 in total because repeated "(graphophonic (phonological)" section across grades 1 to 8. This short bulleted section is the only phonics content.	Whole language

Canada, Ontario	Language	Includes UNESCO quote. Literacy as a communal project with skills embedded across the curriculum. Language central to intellectual, social and emotional growth, and more than just basic skills. Set of principles include language learning as life-enhancing reflective process.	The Reading strand has four overall expectations, e.g. Students will:1. read and demonstrate an understanding of a variety of literary, graphic, and informational texts, using a range of strategies to construct meaning. (p. 11)	"Teaching approaches should be informed by the findings of current research into best practices in literacy instruction, as described in the Expert Panel reports on literacy instruction in Ontario (see the list of resources on the preceding page). Instruction should include a balance of direct, explicit instruction; teacher modelling; shared and guided instruction; and opportunities for students to rehearse, practise, and apply skills and strategies and make choices." (p. 23)	2. 32 in total because repeated “graphophonic (phonological and graphic) cues” section across grades 1 to 8. This short bulleted section is the only phonics content.	Whole language

Canada, Quebec	English Language Arts	"The noted Brazilian educator, Paulo Freire, described literacy as knowing how to “Read the world and the word.” This program is centred in the connection between the learner’s world and words, since language is both a means of communicating feelings, ideas, values, beliefs and knowledge, as well as a medium that makes active participation in democratic life and a pluralistic culture possible." (p. 72) Explicitly notes the connections with the previous curriculum: "children's literature; writing as a process; responding to and interpreting texts; ... the four linguistic cueing systems ... (p. 72)	To develop the students’ capacity for oral (speaking and listening) and written (reading and writing) communication so as to enable him/her to express his/her view of the world, to enter into relationships with young people and adults from near and far, and to acquire and transmit cultural knowledge. (p. 70)	Only two occurrences of the stem phon: "phonetic representation" for spelling. Emphasis on "the four cuing systems" (p. 77) The first two web page tabs of nine tabs are "engagement" and "motivation and choice".	2	Whole language

Ireland	Primary Language Curriculum	Language learning enables children to understand the world around them and to communicate effectively with others. Communication takes many forms, from the non-verbal and verbal to print-based and digital texts. Through interacting with adults, children are initiated into, and engage in communicative relationships through which they come to understand,interpret, construct meaning and critically appreciate the communication of others. (p. 16	"Attend to, take part in and enjoy listening to reading and talking about the meaning and interpretation of written words and illustrations with others, recognising themselves as readers.Discover and explore texts in various languages."	Has a developmental linear sequence for Communicating, Understanding, Exploring, then Using. The next web page tab is "Conventions of Print and Sentence Structure". The emphasis is on understanding these to aid the understanding of text. The following caution is noted: "Direct comprehension instruction, although a vital aspect of the reading process, can often be neglected, particularly in the infant classrooms where a great emphasis is placed on	11 excluding glossary	Whole language

				phonics." (p .18 of guidance material)		
New Zealand	English.	"Literacy in English gives students access to the understanding, knowledge, and skills they need to participate fully in the social, cultural, political, and economic life of New Zealand and the wider world. To be successful participants, they need to be effective oral, written, and visual communicators who are able to think critically and in depth." (p. 18)	English is divided into two categories: Listening, Reading, and Viewing; and Speaking, Writing and Presenting. "Processes and strategies Students will: • Acquire and begin to use sources of information, processes, and strategies to identify, form, and express ideas." (p. 2) Other	Just one page for 'Level One' PoS. This is the section of relevance to phonics: "Language features • Recognise and begin to understand how language features are used for effect within and across texts. INDICATORS: – begins to recognise that oral, written, and visual language features can be	Zero, but 4 in levels document	Whole language

			sections are Purposes and audiences; Ideas; Language features; Structure.	used for effect; – recognises a large bank of high-frequency and some topic specific words; – shows some knowledge of text conventions, such as: capital letters, full stops, and word order; volume and clarity; and simple symbols." (p. 2)		
Australia	English	"The study of English is central to the learning and development of all young Australians. It helps create confident communicators, imaginative thinkers and informed citizens. It is through the study of English that individuals learn to analyse, understand, communicate and build relationships with others and with the world around them.	"The English curriculum is built around the three interrelated strands of language, literature and literacy. Teaching and learning programs should balance and integrate all three strands." (Year 1 Level Description)	Divided into strands: Language; Literature; Literacy. Language strand includes "Language variation and change" and "Language for interaction". Phonics and word knowledge is the fifth category. Includes this: "Manipulate	37 in "English Sequence of content F-6" text.	Balanced instruction

		The study of English plays a key role in the development of reading and literacy skills which help young people develop the knowledge and skills needed for education, training and the workplace. It helps them become ethical, thoughtful, informed and active members of society." (First page of English Learning Area).		phonemes in spoken words by addition, deletion and substitution of initial, medial and final phonemes to generate new words" and other very detailed phoneme and grapheme aspects.		

USA	English Language Arts	English Language Arts Standards » Anchor Standards » College and Career Readiness Anchor Standards for Reading Key Ideas and Details: CCSS.ELA-LITERACY.CCRA.R.1 Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions drawn from the text. CCSS.ELA-LITERACY.CCRA.R.2 Determine central ideas or themes of a text and analyze their development; summarize the key supporting details and ideas. CCSS.ELA-LITERACY.CCRA.R.3 Analyze how and why	**An integrated model of literacy** Although the Standards are divided into Reading, Writing, Speaking and Listening, and Language strands for conceptual clarity, the processes of communication are closely connected, as reflected throughout this document. For example, Writing standard 9 requires that students be able to write about what they read. Likewise, Speaking and Listening standard 4 sets the expectation that students will share findings	English Language Arts Standards » Reading: Foundational Skills » Kindergarten Print Concepts Phonological Awareness Phonics and Word Recognition Fluency	7 in the "Reading Foundational Skills: Grade 1 section. A further 8 in Grades 2 to 5 Reading Foundational Skills sections.	Balanced instruction

		individuals, events, or ideas develop and interact over the course of a text.	from their research.			

USA, Massachusetts	English Language Arts & Literacy	"The Massachusetts Curriculum Framework for English Language Arts and Literacy builds on the Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects (“The Standards”). " (no page number)"An integrated model of literacyAlthough the Standards are divided into Reading, Writing, Speaking and Listening, and Language strands for conceptual clarity, the processes of communication are closely connected, as reflected throughout this document"	Guiding Principle 7An effective English language arts curriculum provides explicit skill instruction in reading and writing.In some cases, explicit skill instruction is most effective when it precedes student need. Systematic phonics lessons, in particular decoding skills, should be taught to students before they use them in their subsequent reading. Systematic instruction is especially important for those students who have not developed phonemic awareness — the ability to pay	Grade 1 students:Phonics and Word Recognition3. Know and apply grade-level phonics and word analysis skills in decoding words.a. Know the spelling-sound correspondences for common consonant digraphs. b. Decode regularly spelled one-syllable words. c. Know final -e and common vowel team conventions for representing long vowel sounds.Use knowledge that every syllable must have a vowel sound to determine the number of syllables in a printed word.	31 excluding notes and glossary. Some repetition of grade level statements.	Synthetic phonics

			attention to the component sounds of language	Decode two-syllable words following basic patterns by breaking the words into syllables. d. Read words with inflectional endings. e. Recognize and read grade-appropriate irregularly spelled words. (p. 10)		

As part of the comparison of curricula in the six regions, the mapping of curriculum orientations to reading was considered in relation to PISA data. This analysis included a comparison of relevant birth cohorts of pupils in England with PISA outcomes data to explore correlations with significant curriculum and/or assessment changes in England. Cohorts of children were mapped against relevant years when PISA testing was undertaken (Table 3). Table 3 shows the changes in total test scores in reading, since 2000, of our selection of PISA countries/states where English is a dominant language and also maps each year there was a PISA assessment against main changes in England's curriculum policy[2] on the teaching of reading based on relevant pupils' birth cohorts.

[Blank space due to page layout]

[2] Pupils in England were initially included as part of UK data in PISA. From PISA 2015 onwards England was reported separately in addition to the UK ranking.

Table 3: PISA total scores for English dominant regions and map of key policy changes in England.

Country/Region	PISA Rank	Score		Year	Key Policy Changes in England
PISA 2022	**na**	**na**		**2022**	Data from PISA from 2021 onwards relevant for the 2006 cohort who first experienced Conservative-Liberal Democrat removal of New Labour national curriculum.
PISA 2018					PISA 2018 data relevant to 2003 cohort who experienced Primary National Strategy-Literacy ("high-quality, systematic phonic work should be taught discretely" p. 7) in place for one year.
Canada Alberta	sub national	532			
Canada	6	520			
Ireland	8	518			
New Zealand	12	506			
United States	13	505			
England	**national**	**505**			
United Kingdom	**14**	**504**			
Australia	16	503			
PISA 2015					PISA 2015 results relevant to 2000 cohort who experienced National Literacy Strategy.
Canada British Columbia	sub national	536		2014	2014 national curriculum introduced.

Canada	3	527		2013	Phonics Screening Check in place for one year.
US Massachusetts	sub national	527			
Ireland	5	521			
New Zealand	10	509			
Australia	16	503			
England	**national**	**500**			
United Kingdom	**22**	**498**			
United States	24	497			
PISA 2012				2012	PISA 2012 relevant to 1997 cohort who experienced the National Literacy Strategy (plus one years of Primary National Strategy - Literacy). Phonics Screening Check first used.
Ireland	7=	523		2010	Conservative-Liberal Democrat government made decision to remove New Labour national curriculum from websites including Primary National Strategy - Literacy.
Canada	7=	523			
New Zealand	13=	512			
Australia	13=	512			
United Kingdom	**23**	**499**			

United States	24	498			
PISA 2009					PISA 2009 relevant to 1993 cohort who experienced National Literacy Strategy in place for one year.
Canada	6	524		2008	Primary National Strategy - Literacy in place for one year.
New Zealand	7	521		2007	Primary National Strategy-Literacy begins ("high-quality, systematic phonic work should be taught discretely") PISA results relevant for Phonics Screening Check, in place for one year, for 2007 cohort from 2023 onwards.
Australia	9	515			
United States	17	500			
Ireland	21	496			
United Kingdom	**25**	**494**			
PISA 2006				2006	PISA 2006 relevant to 1991 cohort who experienced the National Literacy Strategy from 1998 onwards. Jim Rose Report Final published (advocates more synthetic phonics).
Canada	4	527			
New Zealand	5	521			
Ireland	6	517			
Australia	7	513			
United Kingdom	**17**	**495**			

United States	Not included	Not included			
PISA 2003					PISA 2003 would have been relevant to 1988 cohort, who experienced the National Literacy Strategy, but UK not included in PISA 2003.
Canada	3	528			
Australia	4	525			
New Zealand	6	522			
Ireland	7	515			
United States	18	495			
United Kingdom	**Not included**	**Not included**			
PISA 2000					PISA 2000 relevant to 1985 cohort: England's first national curriculum only.
Canada	2	534		1999	National Literacy Strategy in place for one year.
New Zealand	3	529		1998	National Literacy Strategy begins
Australia	4	528		1997	National Literacy Project begins
Ireland	5	527		1988	England's first national curriculum.
United Kingdom	**8**	**523**			
United States	16	504			

We now consider the information from tables two and three together. The whole language orientation is correlated with all the highest ranked regions in our analysis of PISA data. Canada has had the highest score across all PISA cycles, however, relating the score to an orientation to teaching reading is complicated by the fact that individual states in Canada have powers to determine their own curricula. Our additional comparison of the details of the curriculum specifications for primary education in Alberta, Ontario and Quebec revealed a very similar orientation to teaching reading, one that we categorise as whole language. New Zealand's curriculum has remained unchanged since 2006, and is another we categorise as whole language. Ireland's national curriculum did not change from 1999 until 2020 (apart from some new specifications for the Primary Language Curriculum 2015 and 2019) when a review was started, the outcomes of which was due to be implemented from 2022 onwards. We also categorise Ireland's curriculum text as a whole language orientation to teaching reading.

Similar to Canada, in Australia and the USA the state level requirements have been augmented by federal level requirements: in Australia the 'Australian Curriculum', and in the USA the 'Common Core Standards'. The Australian curriculum and the USA Common Core standards represent balanced instruction models of reading teaching. However, an analysis of the 2010 curriculum document for Massachusetts, which was included only in the subnational analysis of the 2015 cycle of PISA, revealed a synthetic phonics orientation.

England's national curriculum has more emphasis on phonics teaching than any other of the six curricula in our comparison. This is evident in the general emphasis on phonics; in the separation made between the requirements of phonics teaching in relation to other aspects of teaching reading; and in the amount of detail in the programme of study for phonics. England's curriculum is the only one at national level in our comparison of English dominant regions that we categorise as synthetic phonics. The frequency of the stem "phon" in England's 2014 national curriculum is 78 which is twice as frequent as the next highest region which is Australia (see Table 2).

Exploring correlations between international comparative data and national curricula in England.

The largest change in overall score for England, across all PISA cycles to date, shows a positive correlation in favour of England's first national curriculum compared to the change to the more structured phonics teaching of the NLS (Table 3). Children born in England in 1985 took the 2000 PISA tests for which the overall score was 523. These children had experienced England's 1988 national curriculum from age six onwards. Children born in 1991 took the 2006 PISA tests (England was not included in PISA 2003) for which the overall score was 494, a significant drop in score. These children, born in 1991, were being taught at the time of the increased phonics teaching recommended in the National Literacy Project which was the prototype for the NLS which provided more detailed programmes of study for the literacy aspects of the national curriculum including its greater emphasis on phonics teaching. This cohort experienced the NLS, in its first year, from age seven onwards.

A much smaller change in score is reflected in the results from PISA 2018. The cohort of children born in 1997 experienced the NLS. The score in PISA 2012 was 499. The cohort of children born in 2000 also experienced the NLS. The score for England in PISA 2015 was 500. The cohort of children born in 2003 experienced the PNS for which "high-quality, systematic phonic work should be taught discretely" DfE, 200?, p. 7). The score in PISA 2018 of 505 reflects a small positive correlation in favour of the greater intensity of phonics teaching.

Overall, the correlations between the PISA scores and changes to reading curriculum policy data favour the less systematic phonics that was part of England's first national curriculum rather than the subsequent increased emphases on synthetic phonics. While there are differences in the scores for England over time that we have related to curriculum changes, when England is compared with trends in other countries the OECD averages since PISA 2006 do not achieve the difference to demonstrate statistical significance at the 5% level (Sizmur, et al. 2019).

The PIRLS data show a contradictory pattern compared with the the PISA data, and when comparing internal trends between the different PIRLS assessment years (see Table 4).

[Blank space due to page layout]

Table 4: PIRLS total scores for English dominant regions and map of key policy changes in England.

Country/Region	**PIRLS Rank**	**Score**		**Year**	**Key Policy Changes in England**	
PIRLS 2021	na	na		2022	PIRLS 2021 results will be relevant to 2011 cohort who experienced Phonics Screening check and Conservative government emphasis on synthetic phonics in 2014 national curriculum and associated mechanisms.	
PIRLS 2016					PIRLS 2016 results relevant to 2006 cohort who for one year had experienced removal of New Labour national curriculum including Conservative-Liberal Democrat government continued increase in phonics emphasis.	
Ireland	4	567				
Northern Ireland	7	565				
England	**10**	**559**				
United States	14	549				
Australia	21	544				
Canada	23	543				
New Zealand	33	523				
PIRLS 2011					PIRLS 2011 results relevant to 2001 cohort who had experienced National Literacy Strategy then the start of the Primary National Strategy-Literacy.	

Northern Ireland	5	558		2008	Primary National Strategy-Literacy in place for one year.	
United States	6	556		2007	Primary National Strategy-Literacy Begins ("high-quality, systematic phonic work should be taught discretely")	
Ireland	10	552				
England	**11**	**552**				
Canada	12	548				
New Zealand	23	531				
Australia	27	527				
PIRLS 2006				2006	PIRLS 2006 data relevant to the 1996 cohort who experienced England's National Literacy Strategy. Jim Rose Final Report published (advocates more synthetic phonics)	
Canada Alberta	3	560				
Canada British Columbia	5	558				
Canada Ontario	7	555				
Canada Nova Scotia	16	542				
United States	18	540				
England	**19**	**539**				
Canada Quebec	23	533				
New Zealand	24	532				

PIRLS 2001					PIRLS 2001 data relevant for the 1991 cohort who experienced England's national curriculum of 1988.	
England	**3**	**553**		1999	National Literacy Strategy in place for one year.	
Canada (O, Q)	6	544				
United States	9	542		1998	National Literacy Strategy Begins	
New Zealand	13	529		1988	England's first national curriculum.	
Canada Ontario	n/a	548				
Canada Quebec	n/a	537				

[Blank space due to page layout]

There have been two contradictory large changes in overall score for England across all PIRLS cycles to date. For the children tested in PIRLS 2001, the cohort born in 1991 who experienced England's first national curriculum, England's score was 553. For the children tested in 2006, the cohort born in 1996 who experienced England's NLS, England's score was 539. These outcomes revealed a positive correlation for the children who experienced England's first national curriculum that had very little specification of phonics teaching. For the children tested in PIRLS 2011, the cohort born in 2001 who experienced the teaching of the NLS, and the beginning of the PNS that included increased specification of phonics teaching, England's score was 552. When compared with the 2006 PIRLS score for England this outcome revealed a contradictory pattern of a positive correlation for more phonics teaching. A much smaller rise in England's score from 552 in PIRLS 2011 to 559 in PIRLS 2016 is correlated with the period when the emphasis on discrete phonics was once again being increased, at the time when the Conservative government's 2014 new national curriculum was in development but before the PSC had been established.

Because of the greater longitudinal aspects of the PISA data it is a more valid source for our purposes than PIRLS. The correlations are in favour of less systematic teaching of phonics. However, the more general problem with this analysis of correlations between orientations to teaching reading and PISA test outcomes is that there are many factors that could have contributed to these correlations not just changes to curriculum and assessment, including problems with sampling of pupils (Anders et al. 2019), and other issues identified in relation to trends in PISA, PIRLS and statutory tests in England (e.g. see Bowers, 2020). Different analyses of international comparative data have produced important findings (e.g. Jerrim & Moss, 2019) but not that relate closely enough to our research interest in key changes to national curriculum approaches to teaching reading. The PISA assessments and their reports provide an important international context for the reading debates, and a wealth of data for further analyses and, as we have shown, some correlations suggest an advantage for whole language orientation to the teaching of reading, but in the end they are not a sufficient way of determining which approaches to the teaching of phonics and reading are most effective in a curriculum. In view of the lack of reliability in the correlations between curriculum policies for reading and national and international tests the most appropriate evidence to judge the most effective teaching of phonics and reading comes from experimental trials particularly, RCTs with longitudinal designs, undertaken with pupils in the region of interest. It is evidence from these kinds of studies that the findings of our SQMS reports later in the paper.

Methodology

The research questions for the research reported in this paper were as follows:

1. How does the orientation of England's national curriculum, and assessment system, to teaching reading at primary level compare to other high-performing countries where the English language is a dominant language?
2. What are the most effective approaches to teaching phonics and reading, for typically developing readers, according to the most robust research evidence internationally?

3. What are the views of teachers about teaching and assessing reading in England?
4. To what extent does the appropriate evidence support a case for change to curriculum policy and practice in primary schools in England?

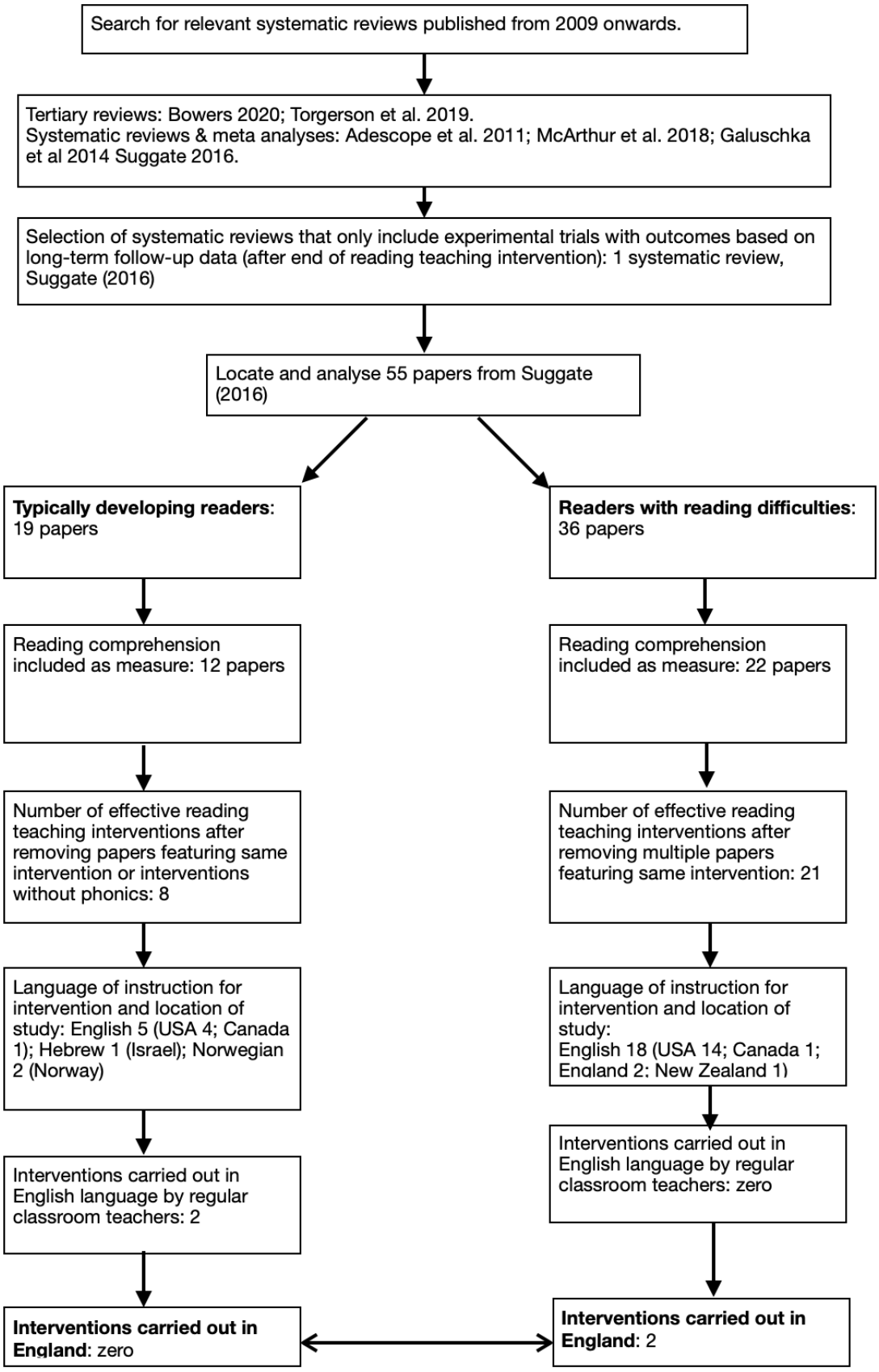
To address these questions the selected research design was a mixed methods concurrent design featuring, a) a systematic qualitative meta-synthesis (SQMS) and, b) a questionnaire survey of teachers. The design of the survey was influenced by the authors' experience in large-scale surveys of teachers (see (Bradbury 2018, Bradbury, et al. 2021; Wyse and Ferrari, 2014), and initial work that had located relevant systematic reviews and meta-analyses (see below).

Systematic Qualitative Meta-Synthesis

The SQMS process is shown as a whole in figure 1.

[Blank space due to page layout]

Figure 1: The process of the SQMS.



The first stage of the SQMS was to locate and review relevant tertiary reviews then systematic reviews (SRs) and meta analyses (MAs). The review of SRs and MAs for this paper covered the period from 2008 to 2020 because previous research and related research publications had addressed systematic reviews published before 2008 and had developed some of the methods for the qualitative meta-synthesis reported in this paper (Wyse and Goswami, 2008; Wyse, 2010). The particular focus, and methods, for different SRs meant that some were more relevant to the research questions reported in this paper than others. To be selected for inclusion in the SQMS the following criteria had to be met by a SR:

1. published after 2008;
2. a main focus on the teaching of phonics and reading;
3. included MA;
4. included evaluation of methodological quality of studies;
5. included analysis of publication bias;
6. identified by tertiary reviews of Bowers 2020 and/or Torgerson 2018 as relevant to evaluation of effective teaching of phonics and reading;

Ultimately Suggate 2016 was chosen as the main source of the studies analysed in the SQMS because it met all the criteria and crucially was also the only SR and MA to exclusively focus on high quality trials that had longitudinal measures: follow-up tests undertaken after the end of the successful reading teaching intervention (mean = 11.7 months after intervention end). Longitudinal designs are an important methodological element because the success of interventions' effects are often only measured at the end, or near the end, of an intervention yet the impact on pupils' reading may not extend beyond the point of the post-intervention measures. The most effective teaching interventions to support children's reading should show effects that have a long-term positive impact.

All 55 research papers reporting the experimental trials cited in the Suggate 2016 SR were located and read in full. The next stage of the SQMS was to establish criteria for the selection of individual papers for in-depth qualitative analysis. The criteria were those which enabled the relevant research questions for our study to be addressed. Ultimately the analysis was seeking to provide evidence-based recommendations for reading teaching practice and for national curriculum policy in England. The studies cited in Suggate 2016 were mapped against the following inclusion criteria as part of our final analysis:

1. longitudinal follow-up measures;
2. sample of children whose reading development was typical;
3. measure of reading comprehension;
4. the language of instruction for the teaching intervention was English;
5. sample of children from age five to age eight;
6. methods used to evaluate fidelity to the intervention;
7. intervention delivered by teachers as part of normal teaching.

In order to fix the final selection of studies for more detailed analysis of their teaching approaches the following key information was extracted from the papers as direct quotations and/or notes: sample size and selection criteria; sample characteristics – children at-risk/not at risk; age of children in the intervention; country and region

location of the research; inclusion of reading comprehension measure or not; overall nature of intervention; who delivered the intervention; allocation to experimental groups and nature of control condition; characteristics of the intervention including its overall approach to teaching phonics and reading; nature of any fidelity measures; main findings of the study. The key information from each study was entered into an Excel spreadsheet. Spreadsheet rows were then progressively filtered, based on each of the criteria in turn, to exclude studies in order to reach the final selection. For those studies that met most of the seven criteria above, further analysis looked in particular at the nature of the teaching in the intervention in relation to our three orientations to reading, and included other aspects relevant to teaching such as calculations of the total hours required for the programme, the frequency of the sessions in the programme, and the total hours required to deliver the programme. As can be seen in Figure 1, eight out of 55 studies met the criteria of longitudinal design; sample of typically developing readers; reading comprehension measure included. Only two studies met all the criteria. No studies that had been undertaken in England met all the criteria.

Appendix 1 provides details, including methodological quality ratings, for the studies in the final selection for the SQMS. Studies were categorised as high quality if they met at least four of the criteria for the SQMS selection as well as: random allocation to intervention and comparison groups; interventions delivered by classroom teachers; and processes to evaluate fidelity to the intervention. Random allocation was chosen as a criterion due to its effect on minimising bias. Intervention by classroom teachers was a criterion due to the focus of our research questions on effective teaching for typically developing readers. Measures of fidelity were important to provide data on the extent to which the intervention was delivered as intended and hence a plausible causal impact on test measures. Studies were categorised as medium quality if they met at least four of the criteria for the SQMS selection; included random allocation to intervention and comparison groups; and processes to evaluate fidelity to the intervention. Studies were described as low quality if they met at least four of the criteria for the SQMS selection, in addition to random allocation to intervention and comparison groups.

Questionnaire Survey

An online questionnaire survey was designed, using *Opinio* software, to establish, a) how teachers in Nursery, Reception and Year 1 classes in England were teaching reading; and b) how Year 2 teachers were responding to the use of the PSC in Year 2. The PSC is usually taken in the summer term by all Year 1 children and by the minority of Y2 children who have not met the expected standard in their previous Year 1 tests. However, in 2020 the Year 1 cohort were unable to take the PSC due to school closures during the Covid pandemic, and so the PSC was moved to the autumn of Year 2 for this cohort of children. Thus the unusual circumstances of the Covid pandemic allowed the researchers to collect additional data from teachers based in Year 2 on the impact of the PSC in what an unprecedented event[3] (see Bradbury, 2020).

[3] The PSC will also operate in Year 2 in the autumn of 2021 as statutory assessments were suspended again due to school closures in the spring term of 2021.

In total there were 22 survey questions. After two initial screening questions the survey had two 'branches': branch 1) for those teaching in Year 2 (14 questions), and branch 2) those teaching in early years or Year 1 or in other roles including leadership (six questions). The analyses for this paper use data from both branches of the survey. For the non-Year 2 teachers the analysis focused on three survey questions of the six (the remaining three were focused on the pandemic and provided context for the wider work on the PSC in Year 2); these three question were:

1. How would you describe your main approach to the teaching of phonics? (choice of three options)
2. To what extent does the Phonics Screening Check affect your practice in relation to phonics? (choice of 4 options)
3. If you could change national curriculum policy on teaching and assessing reading, including phonics, what kinds of changes would you recommend? (free text response)

For the Year 2 teacher respondents we report on their responses to the following questions:

1. Do you think Year 2 children should be doing the Phonics Screening Check this term?
2. What are the main ways which doing the Phonics Screening Check has affected your teaching this term?
3. To what extent do you agree or disagree with the following statements? Doing the Phonics Screening Check in autumn of Year 2 has:
    a. reduced the time spent on other literacy activities
    b. had a positive impact on my teaching this term.

The invitation to complete the survey was distributed from 3 November 2020 to 20 December 2020 via the network of affiliates of the authors' research centre, and the networks of the affiliates, and via social media. Affiliates of the research centre are people interested in research, practice and policy relevant to early years and primary/elementary pupils. 2205 people answered all or some of the survey questions. 1271 respondents identified themselves as Year 2 teachers; 934 respondents identified themselves as non-Year 2 teachers. Of the non-Year 2 teachers there were more Year 1 teachers (n=270), than Reception (n=170) or Nursery teachers (n=42), or those in mixed year groups (n=47). There were 105 respondents from other non-year specific roles, such as leadership positions. The total numbers of responses per question varied as not all respondents answered every question, and so totals for each question are indicated in the findings.

Findings

Evidence from the SRs and MAs

In addition to our existing knowledge of the field, and preliminary literature searching, two *tertiary* reviews were used to support the location of and selection of appropriate SRs and MAs: Bowers (2020) and Torgerson et al. (2019). Tertiary reviews systematically locate and analyse SMs and MAs. In addition to their use in

highlighting relevant SRs and MAs the findings of tertiary reviews are in themselves relevant to our research questions. On the basis of their review of SRs and MAs relevant to phonics teaching Torgerson et al (2019) concluded that although there was evidence that phonics teaching was beneficial for young readers the evidence did not support a "phonics only" teaching policy because "many studies have *added* phonics to whole language approaches, balanced instruction is indicated." (Torgerson et al., 2019, p. 27. Italics in original.) Bowers' (2020) tertiary review concluded that "the above research provides little or no evidence that systematic phonics is better than standard alternative methods used in schools … the findings undermine the claim that systematic phonics is more effective than alternative methods including unsystematic phonics (such as whole language)" (Bowers, 2020. p. 16). The explicit mention of 'whole language' as a relevant consideration, on the basis of multiple SRs, MAs, and RCTs, in both the Torgerson et al (2019) and Bowers et al (2020) reviews is particularly notable given the trend over time towards greater emphasis on, and attention to, synthetics phonics and the parallel denigration of whole language as an approach that as we showed above is sometimes seen.

The relevance of our categorisation of three approaches to reading that inform this paper can be seen in the concerns of both Torgerson and of Bowers for example in their use of the phrases "phonics only"; "systematic phonics"; "balanced instruction"; "whole language". The use of rigorous systematic review techniques, to locate high quality studies that use experimental methods appropriate to the evaluation of claims of teaching effectiveness, is an important contribution to many debates about teaching. However it is important to note the differing conclusions that these robust SRs and MAs can reach even when their analyses are based on some of the same kinds of original sources. For example Bowers questions some of the methods and conclusions in the Torgerson et al study, and in turn Bowers' conclusions have also been questioned in a response to Bowers' paper by Fletcher, Savage and Vaughn (2020) which Bowers responds to in a further paper (Bowers 2021). The differing conclusions, and the methodological limitations, underline the complexity of the debate but also the need for analyses which not only take account of the statistical outcomes of SRs and MAs but go beyond these to systematically examine the contextual details of studies relevant to teaching and reading policies in particular regions of the world, and locate these in a wider historical and political context for the debate, an approach to analysis that we took to inform this paper. Not-with-standing the methodological limitations of these SRs and MAs we regard it as important to report their headline findings in recognition that all research studies have limitations and, particularly if published in reputable peer-reviewed research journals, are worthy of consideration.

The SR and MA published earliest in our selected date range focused on effective teaching of reading and writing for pupils who had acquired English as a second language (Adescope et al. 2011). Two outcomes from the study were of particular importance: 1. peer interaction to negotiate a shared understanding of the meaning of texts produced larger effects for increased competence in reading, and writing, than systematic phonics instruction; 2. it was recommended that policy makers take account of contextual factors, such as particular contexts for schools, when making decisions about optimal pedagogies.

McArthur, et al. (2018) was an update of a review first carried out in 2012. This Cochrane Library Intervention Review focused on phonics training for “English-speaking poor readers” (p. iii. We use the term ‘at-risk readers’ in this paper). The review concluded that phonics training may have improved at-risk readers’ accuracy for reading phonetically regular real and pseudo words, and only slightly improved reading comprehension, but the evidence for both these findings was “low-quality”. Overall it was concluded that more studies are needed to improve the precision of outcomes including in relation to reading comprehension and reading fluency.

Galuschka *et al.* (2014) found that phonics instruction was the most effective method for the reading and spelling performance of “reading disabled children and adolescents” (p. 9). However, the description of phonics instruction in this SR included reading fluency, described as “repeated word or text reading practice”, so could be described as a balanced instruction orientation. Also of interest was the view that “The Anglo-American region far outweighs other countries in quantity and quality of the published work in this research domain.” (Galuschka, 2014, p.10)

The SR and MA that was finally selected for greater analysis in our SQMS was Suggate (2016). In addition to being cited by Torgerson 2018 and Bowers 2020 as relevant to a focus on the teaching of phonics and reading, and because it met our criteria for selection of SRs and MAs, it also was the only one to focus explicitly on the long-term effects of reading interventions. Suggate’s (2016) SR examined the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. The overall findings from Suggate (2016) were that phonemic awareness training was more effective than phonics interventions and that “the greatest effect sizes at follow-up appeared to result from interventions with a comprehension component.” (p. 87). The nature of the work by Suggate and its fit with our inclusion criteria meant that it was selected for in-depth analysis of all the individual papers that were part of Suggate’s SR in order to learn more about key characteristics relevant to our three overall approaches to reading and other aspects related to effective reading teaching.

Systematic qualitative meta-synthesis

In order to determine what is the most effective teaching of reading for the majority of children in England the studies to be selected for the SQMS were filtered on the basis of a set of criteria relevant to our research questions. The first criterion was the nature of the sample of children. The potential for generalisability of the research findings in relation to teaching methods required consideration of whether the sample in a research study was of children whose reading development was typical or the sample was of children with reading difficulties. 19 of the papers in the SQMS had samples that were typical readers and 36 papers had samples that were of children with reading difficulties. The papers reporting on the basis of samples of children whose reading development was typical were selected because the main focus of our research is effective teaching for these children and their teachers. The orientations of the effective interventions across all 19 studies that included typically developing readers were as follows: synthetic phonics – 1; other phonics – 6; balanced instruction – 8; whole language – 4 although three of these whole language interventions emphasised reading comprehension and/or strategy instruction, and one included some systematic phonics.

The ultimate goal of reading teaching is improvement in pupils' comprehension of texts, therefore studies needed to include a measure of children's reading comprehension to be selected for the SQMS. Of the 19 papers with samples of children whose reading was typical 12 papers including measures of reading comprehension. Of these 12 papers eight different reading teaching interventions were identified by excluding papers reporting the same intervention, with results from a prior measurement point, or those papers that had interventions that did not include phonics (e.g. interventions that were reading comprehension only). Details about the final selection of eight studies can be found in the Appendix, including: sample; total teaching time of dose; allocation to groups; main finding of the study; and an assessment of the overall quality of the study.

Another criterion relevant to generalisability was the language of instruction. As phonics teaching was a key component of all the studies, the language of instruction was important because the transparency of the orthography is theorised to be relevant to the teaching strategies adopted (Seymour, Aro & Erskine, 2003). Five of the eight remaining interventions were English language. Four of the reported research studies had been undertaken in the USA. The other English language study was undertaken in Canada. One study was undertaken with Hebrew in Israel and two with Norwegian in Norway.

The age of the sample of children in the final section of eight studies ranged from a mean age of five years and two months to children aged eight years. Five studies included participant children age five to age six. Two studies had children age six to seven, and one study included children age six to age eight.

In four of the eight studies the interventions were not delivered by teachers, for example delivered by paraeducators or volunteer tutors. The positive aspect of this is the evidence that people other than teachers can effectively contribute to children's reading development. The limitation is the more limited data about effective interventions that teachers deliver in the course of their normal teaching.

Another important consideration for effective teaching of reading is how much time should be devoted to any particular intervention. The first thing to note is that the first year of the successful interventions typically included daily lessons/activities up to four times per week typically lasting about 30 minutes duration (one intervention also had an adjusted programme for a second year). The minimum amount of hours total duration for an intervention was approximately 9.1 hours and the maximum was approximately 60 hours. The interventions carried out in Norway reported in Lyster (2002) were notable for the lowest number of hours for delivery of the effective interventions: approximately 9.1 hours delivered in about 35 minutes per week in one or two lessons.

One key point of disagreement about the teaching of phonics for reading has been whether it is more effective to focus on the phonics separately from reading with whole texts or to integrate the teaching of phonics with whole texts. Six of the seven interventions included lessons that included both the teaching of phonics and teaching with whole texts. All the selected studies included the use of texts specially created to enable reading often referred to as 'decodable texts', however these

appeared to vary considerably. The merits of decodable texts versus 'real' books and texts has not been researched using an RCT with longitudinal design. The pupils in all the studies experienced non-decodable texts in classroom activities and lessons outside of the phonics programmes.

Only two studies carried out in the English language were taught by classroom teachers, rather than paraeducators/other assistants. The study undertaken in Canada (Phillips, Norris and Mason, 1996) included 'arms' of the experimental trial where parents delivered the intervention, however only the school-only condition, i.e. not involving parents' delivering the intervention, showed positive gains for reading when students were retested in Grade 4, four years after the intervention was first introduced. One possible explanation for the school-only finding is that the arm of the trial that included parents *and* teachers was a more complicated instruction context. Parents would not have the skills of trained teachers, in particular the knowledge that develops as a results from teaching reading to multiple classes of children over many years. There is a significant separate field of studies focused on the most appropriate ways that parents can help their children with reading.

The successful intervention, carried out in Canada, included materials clearly built on a rationale of the importance of texts to contextualise the teaching of the alphabetic code, for example:

> The reading intervention materials (McCormick & Mason, 1990) consisted of a series of booklets with the following features and rationale: (a) They were thematic and contained familiar topics to increase the child's expectation that text should make sense … c) There was a strong fit between illustrations and text to develop the concept that both text and picture frame the meaning … (op cit. p. 180)

This intervention is best described as a whole language approach, however both intervention groups and control groups also had separate teaching from "the Language Development Reading Series (McInnes, 1988), the purpose of which is "to cultivate familiarity with print" (McInnes, 1988, p. ix)." (Op cit., p.180). Fidelity to the intervention and control arms was not reported. Overall, the materials were designed "to fit closely the early literacy needs of children entering school at risk of failure" (p. 180) so once again, even with this study, we have to be cautious about the extent to which the findings are generalisable to children with typical reading development. And although this study meets more of the inclusion criteria than any other studies it was still was not undertaken in England. The other English language intervention delivered by classroom teachers was carried out in the USA (Gunn, Smolkowski and Vadasy, 2011) but the final outcomes were not statistically significant.

The other studies delivered by classroom teachers were undertaken in Norway (Lyster, 2002 and Lie, 1991). The interventions in Lie (1991) showed positive effects for both the positional analysis group (that helped to develop children's skills in identifying initial, final and medial phonemes in words) and the sequential analysis group (that helped children to identify the phones in spoken words in the right sequence) were both effective compared to control. The interventions included the use of specially prepared stories that used the pictures and text to enable practice with the phonemes that the children had learned about. The orientations of these

interventions was balanced instruction. The interventions in Lyster (2002) were phonological awareness or morphological training, both of which were effective in comparison with the control group. The morphological group outperformed phonological awareness and control on the measure of word reading, and both experimental groups performed better than the control group on text reading. These interventions were phonics orientations but not synthetic phonics.

The other English language studies were those undertaken by Vadasy and colleagues in the USA although the interventions were implemented by paraeducators not by teachers. The intervention lessons in three Vadasy papers (Vadasy and Saunders, 2011; Vadasy and Saunders, 2012a; Vadasy and Saunders, 2012b) addressed both teaching in the alphabetic code and teaching with whole text, although it is not clear to what extent connections were made between these two aspects:

> Students assigned to treatment received individual systematic and explicit phonics tutoring instruction in English, which included letter-sound correspondences, phonemic decoding, spelling, and assisted oral reading practice in decodable texts. … In a typical tutoring session, paraeducators spent 20 min on phonics activities and 10 min scaffolding students' oral reading practice in decodable texts. (Vadasy and Sanders, 2012, p.990)

These interventions are best described as balanced instruction orientation.

The other effective intervention using a language other than English also connected.the phonics teaching with whole texts. The successful intervention in the Hebrew language (Kozminsky, & Kozminsky, 1995) was also a balanced instruction orientation which included an unstructured phonological awareness activity centre in the classroom; collections of stories poems and games; and was undertaken in addition to the normal general language enrichment program that was part of the nation kindergarten curriculum.

In summary, no studies met all the criteria of: experimental design with random allocation; longitudinal design; sample of children whose reading was typical; delivered by standard class teachers; reading comprehension measures included; and undertaken in England with the English language. This is an important limitation for those interested in developing national curriculum and assessment policy to support evidence-based reading teaching in England.

While acknowledging the overall limitations of the existing research, and that direct experimental comparisons for many aspects of teaching that need answers are not part of the research designs, our interpretation of the most robust research evidence from Tertiary reviews, from SRS and MAs, and from our SQMS of longitudinal studies of most relevance suggests that phonics teaching is likely to be effective if it is:

1. implemented with children aged five to six (In England in Year One);
2. carefully connected with the reading of whole texts, both decodable and real books, including a focus on reading for meaning, in all lessons;

3. undertaken during the course of not more than one whole school year featuring several lessons per week between 36 hours and 60 hours in total teaching time.

There is much that we do not know about optimal total amount of time for phonics teaching including the frequency of lessons, although the research indicates that daily sessions are appropriate. It is also not known when sufficient phonics teaching has been done so that the focus of teaching can move productively to an emphasis on other aspects of reading and literacy more generally. Given that there was evidence that effective phonics and reading teaching could be delivered in the lower end of the range at 9.1 hours this could be equated with 6 weeks of teaching of 30 minutes per day, a model that would be considerably shorter than the 2014 national curriculum requirements in England which specify phonics teaching from the Early Years Foundation Stage (for children from age four) and then in the national curriculum for children aged five to age seven. Only when children are in Year 3 (age seven to eight) do the non-statutory requirements note that "At this stage, teaching comprehension should be taking precedence over teaching word reading directly." (Department for Education, 2013, p.36) another clear example of how phonics teaching and comprehension are separated.

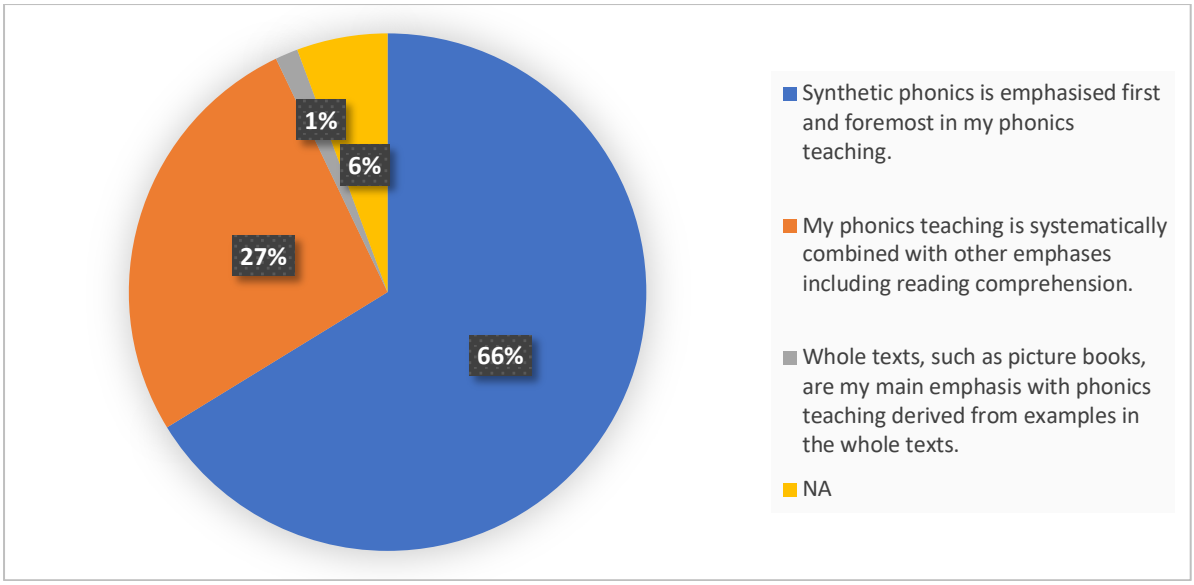
The studies in Canada and Norway clearly showed that effective teaching of phonics teaching and reading was delivered by class teachers who adjusted their normal practice to fit the aims and strategies of the interventions. The interventions in the other studies were additional to normal classroom practice, and indeed delivered by paraeducators and assistants, however these practices could also be adapted as part of normal classroom reading teaching practice by most teachers, possibly supported by classroom assistants.

Survey Findings

Approaches to teaching phonics

Our findings from the analyses of the survey data begins with the responses from non-Year 2 respondents. In relation to the survey question, "How would you describe your main approach to the teaching of phonics?", Figure 2 shows that "synthetic phonics first and foremost" was the way the majority of respondents reported that they taught phonics (420 out of 634 responses: 66%). Whole texts were only seen as the main emphasis and context for phonics teaching in a very small number of responses (9 out of 634 responses: 1%).

Figure 2: Responses to the question "How would you describe your main approach to the teaching of phonics?"



The majority of respondents who entered open comments in relation to this question were those who had said that their approach was synthetic phonics first and foremost (62 out of 101 open comments: 61%). For most of these 62 respondents, phonics teaching was implemented at a different time from teaching other aspects of teaching reading. For example one respondent said, "We teach SSP discretely and systematically. Literacy is taught separately though of course vocabulary may overlap." (underlines added). Another respondent said, "Synthetic phonics using letters and sounds really helps our children learn to decode words quickly. We use good quality, whole texts for comprehension in English lessons" (.e. not in the phonics lessons which are separate). Similarly another respondent wrote, "Reading comprehension is done through 1:1 rather than whole class phonics sessions and through story time." A very clear articulation of the separation was made like this:

> We consider the teaching of reading to have 3 distinct and equally important strands: phonics, comprehension and reading for pleasure. We teach phonics systematically and discretely but we also separately teach comprehension and develop an environment which engenders a love of reading. (underline added)

A few respondents said that they reinforced their phonics teaching in separate lessons on English, literacy and/or comprehension but this is not the same as close and systematic integration of whole texts as part of the phonics lesson, hence learning the alphabetic code is decontextualised from whole texts.

In the open comments from respondents who used synthetic phonics first and foremost, one commercially published scheme/basal was referred to more than any other: *Read Write Inc*. This is in line with previous research in England which has identified the dominance of this scheme (Bradbury 2018). There were however differences in opinion about how the *Read Write Inc* approach to teaching phonics was described, as Table 5 shows.

Table 5: A comparison between comments about *Read Write Inc* and Letters and Sounds.

	Mentioned in 'phonics first and foremost' responses	Mentioned in 'phonics systematically combined' responses
Comments citing *ReadWrite Inc.*	13	9
Comments citing *Letters and Sounds*	7	2

More respondents indicated that they saw *Read Write Inc* as an approach to teaching phonics first and foremost (13 responses) rather than one that systematically combined phonics teaching with books and other aspects such as reading comprehension (9 responses), although the responses relevant to this point are low in number and so have to be treated as less secure. In our view *Read Write Inc* is built on an approach that is more accurately described as phonics first and foremost, and hence a synthetic phonics orientation. For example the handbook gives this guidance:

> .. once children have learnt the Set 1 Speed Sounds and can blend words made up of these sounds, they can start on the Sound Blending Books, then the Red Ditty Books and the *Get Writing!* Red Ditty Books. When they move onto the Green Storybooks and the *Get Writing!* Green Books, they are taught the Set 2 Speed Sounds and continue to review Set 1 Speed Sounds and blending." (Miskin, 2020. p. 11).

Certainly in the first phase of the approach the separation between phonics teaching and whole texts is explicit. Children are taught some sounds and blends separately, and only then can they try these with books, and only decodable books. Also if children have not been assessed as knowing the 'Set 1 phonemes' they cannot move onto Set 2 which includes reading of whole texts. However, the ultimate test of teaching reading is whether pupils can engage with a range of real texts not only decodable texts.

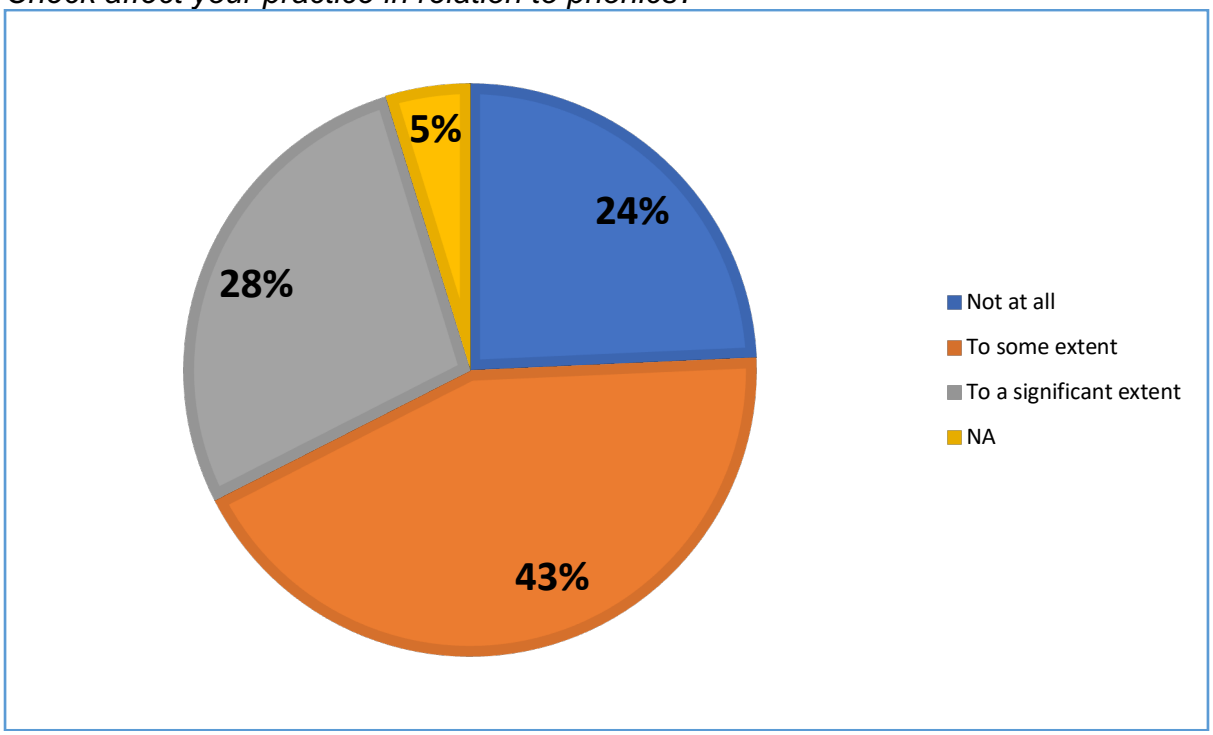
A less pronounced difference in opinion was evident about the next most cited resource, *Letters and Sounds*, which was developed in 2007 as a result of the Rose Review in the time of the NLS in England. Only two respondents categorised this approach as phonics systematically combined. This resource is also best categorised as a synthetic phonics orientation, as the resource's authors made clear: "Systematic high quality phonics work … is best taught in short, discrete daily sessions, with ample opportunities for children to use and apply their phonic knowledge and skills throughout the day." (underline added, Department for Education and Skills, 2007 p.7).

The influence of the phonics screening check on teaching reading

For the Year 2 teacher responses the survey data showed that the Phonics Screening Check had an impact on respondents' teaching. There were 634 responses to the question 'To what extent does the Phonics Screening Check affect

your practice in relation to phonics?', with a majority responding 'To some extent' (43%) and a further 28% responding 'To a significant extent', as shown in Figure 3.

Figure 3: Responses to the question 'To what extent does the Phonics Screening Check affect your practice in relation to phonics?'



In response the question, "If you could change national curriculum policy on teaching and assessing reading, including phonics, what kinds of changes would you recommend?" there were 72 written comments. The PSC was mentioned in 52 of these comments. 49 responses were negative, referring to removal or reform of the PSC. The remaining three comments on the PSC were unclear or ambiguous; there were no positive comments about the PSC.

The comments that referred to the removal of the PSC altogether included this: "I think teaching phonics is important, but don't really see the value of the screening [check]. We assess phonics anyway and see progress through reading and writing". Another respondent commented:

> We assess children in their phonics anyway and we teach them to use the sense of the text to self-correct and learn new vocabulary when reading but the screening check does not allow for that. If a child reads a word one way, then self corrects, they are penalised. It is not an accurate system and a very demoralising way of assessing phonics.

There were also a number of comments relating to the use of pseudo-words in the PSC such as "[There should be] no formal assessment [of phonics] or if there is omit pseudo words, as too many children these just confuse further"

The remaining comments in response to this question were mainly about reducing the emphasis on phonics in favour more wholistic approach to the teaching of reading. For example the need for ‘a more wholistic approach to the teaching of reading’ or ‘More emphasis on comprehension and enjoyment of books’. Another respondent expressed a desire for, ‘A reflection on the mass of skills involved in reading rather than solely focusing on phonics’. The only positive comment read: ‘I think synthetic daily phonics is a vital part of developing children's early reading skills and their phoneme, grapheme correspondence’.

In response to the survey question, ‘What are the main ways which doing the Phonics Screening Check has affected your teaching this term?’ there were 936 written responses from Year 2 respondents. These were categorised as shown in Table 6. Note that some comments were coded under more than one theme.

Table 6: Categories and frequencies of mentions in relation to the PSC affecting teaching.

Category	**Number of mentions**
No change in my teaching	90
Preparation for the PSC (including practice tests)	285
More teaching of pseudo-words	73
Extra phonics lessons as part of classroom teaching	237
Interventions to remove children from typical lessons for extra phonics	157
Changes to classroom grouping practices	18
Delaying teaching other Year 2 curriculum content	108
Reduction in time spent on other literacy activities and in other subjects	190
Phonics homework and advice for parents	11
Experiencing more pressure and additional workload	111

While there were a number of comments that suggested teaching had not been affected (90 mentions), the majority commented on some changes, including most frequently: increased time spent preparing for the test (285 mentions), and extra phonics lessons (237 mentions), and the use of intervention groups to target some children (157 mentions). Teachers’ comments explained how the typical Year 2 curriculum had been delayed (108 mentions) due to the need to focus on preparation for the PSC, including avoiding introduction of new phonemes to “avoid confusion” in the test.

Other comments suggested a dominant focus and emphasis on phonics, such as “English has been replaced by RWI phonics teaching”. The test had effects on the teaching of wider reading skills, as one respondent explained: “We have had to miss Guided Reading sessions, which build children up for the SATS reading paper, to focus on phonics. This then means the children are a term behind consolidating key comprehension skills for the KS1 SATS further up the school”. This distinction between phonics and wider reading skills was common in teachers’ comments, for example, “More phonics focused teaching. Less reading teaching”. These kinds of

comments revealed an emphasis on phonics as a discrete subject separate from other aspects of reading.

These effects have to be considered in the light of the number of comments about the increased pressure exerted on Year 2 teachers in the autumn of 2020 to ensure children who had experienced a disrupted Year 1 reached the expected standard in the PSC. Several respondents mentioned the pressure from their senior leadership teams; one commented 'We are being asked to "live and breathe phonics"'. Indeed, the word 'pressure' appeared 97 times in the responses, referring to pressure on both teachers and children.

Taken as a whole the main finding from the analyses of the survey data show that a phonics first and foremost approach dominates the teaching of reading in England. The survey data also provide clear evidence of the impact of the PSC. The PSC is a high stakes assessment, in that the data form part of the package of results used by government to hold schools to account. This accountability is also reflected in the monitoring of schools' performance by the national inspectorate, Ofsted, in its grading of primary schools as outstanding, good, requires Improvement or inadequate. The nature of the high stakes assessment means that schools are under pressure to ensure high proportions of children pass the PSC each year, something that impacts on how phonics and reading is taught.

Discussion and conclusions

The findings from the survey reported in this paper showed that synthetic phonics first and foremost is the dominant approach to teaching reading in England, and that assessment policy in the form of the PSC has contributed to the stronger emphasis on phonics as part of the teaching of reading. Taken as a whole the responses to the survey showed that policy changes have resulted in adaptations to pedagogy including devotion of a greater proportion of teaching time to phonics, separation of phonics from other literacy activities, and reliance on a limited number of phonics schemes. When considered in relation to policy changes in England since 1988, it is clear that the emphasis of the teaching of reading in primary schools in England has moved significantly for the first time in more than 100 years: the main emphasis in the teaching of reading is now synthetic phonics. Historically, this change is remarkable given the relative stability of reading pedagogy over many decades. The change did not happen quickly, instead it was a result of a series of very significant changes in national curriculum policy. The NLS of 1998 was a well-funded attempt to change pedagogy nationally. This was followed by the increased emphasis on discrete teaching of phonics recommended by the Rose report and the PNS from 2006 onwards. Further intensification of synthetic phonics teaching was seen in England's national curriculum of 2014, along with a range of other measures to ensure teacher compliance with the prescribed method of teaching reading, including the use of the PSC; the vetting of phonics teaching schemes; and the use of the inspectorate to focus on outcomes in statutory reading assessments as a prime focus in school inspections.

The key question that we address in this paper is whether robust research evidence supports this historically significant change in reading pedagogy. Our findings from analysis of tertiary reviews, systematic reviews and from the SQMS do not support a

synthetic phonics orientation to the teaching of reading: they suggest that a balanced instruction approach is most likely to be successful. They also suggest the need for a new more careful consideration of the strengths and weaknesses of whole language as an orientation to teaching reading. The reading wars have often resulted in some very dismissive attitudes to whole language, a position that is not underpinned by the research. While there remains no doubt that phonics teaching in general is one important component in the teaching of reading, the research certainly does not suggest the complete exclusion of whole language teaching.

Systematic tertiary reviews are a relatively new phenomenon in relation to research in a range of education topics because they rely not only on multiple experimental trials but also multiple systematic reviews. The findings from tertiary reviews are important because, similar to systematic reviews, they are accounting for multiple research studies where the quality of the studies is also explicitly reported. Too often views about effective policy on the teaching of reading can be influenced by single research studies (Wyse and Goswami, 2008). The approach in this paper, including the development of the SQMS, represents another way to ensure the appropriate range of research is systematically accounted for when seeking evidence-informed recommendations for effective teaching and related curriculum policies.

The goal of reading teaching is for pupils to be able to comprehend texts, ultimately in ways that include sophisticated understanding of texts and well-justified views about texts. The importance of comprehension means that the most relevant research about the teaching of reading needs to include measures of comprehension. The effectiveness of any reading teaching intervention also needs to be measured longitudinally. If for example an approach to teaching reading proves to be effective after four or more years of measurement we can have more confidence in its effectiveness than if the measures are only carried out just after the end of the intervention. The main finding of our SQMS is that there is no study that has been carried in England with typically developing readers that fits our rigorous criteria. Having noted once again that important caveat we can hypothesise what is likely to be effective on the basis of the studies carried out in other nations where the English language is dominant including the summary evidence from tertiary reviews and SRs and MAs.

The most effective interventions relevant to our research questions carefully connected the reading of whole texts with the teaching of phonics and other relevant aspects within all lessons. We describe this as *contextualised teaching of reading*. The undue separation of the teaching of the alphabetic code from the context of whole texts as part of teaching in primary/elementary schools is unlikely to be as effective as contextualised teaching of reading, and as such poses a significant risk to typically developing children's education and life chances because it is not optimal robust evidence-based teaching. If education policies also fail to sufficiently reflect the robust research evidence this risk is compounded. The positive effects sizes in the effective interventions that we report can be seen as a quantifiable potential negative effect on children's learning that was not achieved due to lack of implementation of optimal evidence informed reading teaching in England.

In addition to the importance of contextualised reading teaching as an evidence-based orientation to the teaching of reading we hypothesise the following

pedagogical features that are likely to be effective. Phonics teaching is most likely to be effective for children aged five to six. Phonics teaching with children younger than this is not likely to be effective. A focus on whole texts and reading for meaning, to contextualise the teaching of other skills and knowledge, should drive pedagogy. Classroom teachers using their professional judgement to ensure coherence of the approach to teaching phonics and reading with other relevant teaching in their classroom is most likely to be effective. Insistence on particular schemes/basals, scripted lessons, and other inflexible approaches is unlikely to be optimal. Well-trained classroom assistants, working in collaboration with their class teachers, could be a very important contribution to children's reading development. Although the most relevant studies in the SQMS showed approaches that were effective usually from between 9.1 hours and 60 hours of teaching time we hypothesis that effective teaching of the alphabetic code could be delivered in 30 hours or less of instruction time. If so this would mean that greater emphasis on aspects such as reading comprehension could begin much earlier in England's national curriculum programmes of study than in the current national curriculum of 2014.

There were various limitations in our research design. The SQMS restricted is analysis to experimental trials with longitudinal designs. These were located from a previously published systematic review. it is possible that there are other RCTs with longitudinal designs that we did not locate. Studies that do not have longitudinal designs still have important findings but we did not include these in the SQMS for the reasons outlined in the paper. The survey respondents were contacted via links with the authors' research centre. It is possible that there was a bias in the sample of respondents in relation to their views on the teaching of reading. We have no reason to believe that that there was systematic bias towards one particularly orientation to the teaching of reading, and the data reveal a range of views, but we cannot rule out that possibility.

In comparison with the national curricula in the other English language dominant countries that we reviewed, England's national curriculum of 2014 represents an outlier. Whereas other countries either have a whole language or balanced instruction orientation to the teaching of reading, England heavily emphasises synthetic phonics teaching. This emphasis is not only in the amount of specification of phonics in its programmes of study but also in the requirement for a particular variant of phonics teaching, synthetic phonics. Our analyses of the PISA data suggest that teaching reading in England has been less successful since the introduction of more emphasis on synthetic phonics, although the correlations reported here require further research. In relation to the national curricula of the regions that we reviewed there is little evidence to suggest that a synthetic phonics first-and-foremost orientation to national curricula is likely to be the most effective orientation.

The influence of successive governments in England on reading pedagogy, and the disjuncture between evidence from the most appropriate robust research and national curriculum policy, suggest a wider problem. In our view it is now apparent from the example of successive governments in England, of different political parties, that it is not appropriate for governments and their individual ministers of education to have the power to directly control curriculum and pedagogy. Where robust research evidence exists this should be the required basis for governments'

recommendations. An independent body, possibly like the Office for Education Research proposed by the Royal Society/British Academy report (British Academy & The Royal Society, 2018), is one appropriate way forward. The power of the Secretary of State for Education in England to construct England's national curriculum and assessment system could be revoked under new legislation to allow the independent body to make recommendations to Ministers to inform policies. There is a need for a much more collaborative policy ethos where policy makers, teachers and researchers work together, over longer time scales than in the past, to evolve national curriculum and associated pedagogy on the basis of the most robust evidence of effective teaching methods: in the context of the decades of the reading wars this is what we call *reading reconciliation*.

References

* = studies that are part of the SQMS.

Alexander, R. J. (Ed.) (1995). *Versions of Primary Education*. London: Routledge and The Open University.

Anders, J., Has, S., Jerrim, J., Shure, N., & Zieger, L. (2020). Is Canada really an education superpower? The impact of non-participation on results from PISA 2015. *Educational Assessment, Evaluation and Accountability, 33*, 229-249.

* Antoniou, F., & Souvignier, E. (2007). Strategy instruction in reading comprehension: An intervention study for students with learning disabilities. *Learning Disabilities: A Contemporary Journal, 5*, 41-57.

Beard, R. (1993). *Teaching Literacy: Balancing Perspectives*. London: Hodder Education.

* Blachman, B. A., Schatschneider, C., Fletcher, J. M., Francis, D. J., Clonan, S. M., Shaywitz, B. A., & Shaywitz, S. E. (2004). Effects of intensive reading remediation for second and third graders and a 1-year follow-up. *Journal of Educational Psychology, 96*, 444-461.

* Blachman, B. A., Tangel, D. M., Ball, E. W., Black, R., & McGraw, C. K. (1999). Developing phonological awareness and word recognition skills: A two year intervention with low-income, inner-city children. *Reading and Writing: An Interdisciplinary Journal,, 11*, 239-273.

Bowers, J. (2020). Reconsidering the Evidence That Systematic Phonics Is More Effective Than Alternative Methods of Reading Instruction. *Educational Psychology Review*. doi:https://doi.org/10.1007/s10648-019-09515-y

Bradbury, A. (2018). The impact of the Phonics Screening Check on grouping by ability: A 'necessary evil' amid the policy storm. *British Educational Research Journal*, *44*(4), 539-556.

Bradbury, A. (2019). "Taking the 'early yearsy' route: resistance and professionalism in the enactment of assessment policy in early childhood in England." Education 3-13 **47**(7): 819-830.

Bradbury, A. (2020). The use of the Phonics Screening Check in Year 2: The views of Year 2 teachers and headteachers HHCP Working Paper 1. London: UCL Institute of Education

Bradbury, A., A. Braun and L. Quick (2021). "Intervention culture, grouping and triage: high-stakes tests and practices of division in English primary schools." British Journal of Sociology of Education **42**(2): 147-163.

* Brady, S., Fowler, A., Stone, A., & Winbury, N. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia, 44*, 26-59.
British Academy, & Royal Society. (2018). *Harnessing Educational Research*. Retrieved from https://royalsociety.org/topics-policy/projects/royal-society-british-academy-educational-research/:
Brooks, G. (2021). Joseph Neef (1770-1854): a forgotten pioneer of applying phonetics and regularised phonic materials to the initial teaching of literacy in English. *Language and History, 64*, 1-26.
* Burns, M. K., Senesac, B. J., & Silberglitt, B. (2008). Longitudinal effect of a volunteer tutoring program on reading skills of students identified as at-risk for reading failure: A two-year follow- up study. . *Literacy Research and Instruction, 47*, 27-37.
* Büyüktaşkapu, S. (2012). Effect of family supported pre-reading training program given to children in preschool education period on reading success in primary school. . *Educational Sciences: Theory & Practice, 12*, 309-316.
Castles, A. R., K & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest, 19*, 5-51.
Cato, V., Fernandes, C., Gorman, T., Kispal, A., & White, J. (1992). *The Teaching of Initial Literacy: How do teachers do it?* Slough: National Foundation for Educational Research.
* Center, Y., Wheldall, K., Freeman, L., Outhred, L., & McNaught. (1995). (1995). An evaluation of reading recovery. *Reading Research Quarterly, 30*, 240-263.
Chall, J. (1983). *Learning to Read: The Great Debate* (Updated ed.). New York: McGraw-Hill.
* 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*, 744-781.
Clark, M. (2013). "The Phonics Check for all Year 1 children in England: its background, results and possible effects." Education Journal(160).
* Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*, 1106-1116.
Clay, M. (1979). *The Early Detection of Reading Difficulties* (3rd ed.). Auckland: New Zealand: Heinemann Education.
Darnell, C. A., J. E. Solity and H. Wall (2017). "Decoding the phonics screening check." British Educational Research Journal **43**(3): 505-527.
Department for Education. (2014). Primary National Curriculum until 2014. Retrieved from https://webarchive.nationalarchives.gov.uk/20140107083211/https://www.education.gov.uk/schools/teachingandlearning/curriculum/primary
Department for Education (2019) Phonics screening check and key stage 1 assessments: England 2019 Available at: https://www.gov.uk/government/statistics/phonics-screening-check-and-key-stage-1-assessments-england-2019 (Retrieved 11 August 2021)
Department for Education (DfE). (2021). The removal of Letters and Sounds 2007 from the Department's list of validated phonics programmes – teachers' questions answered. Retrieved from https://dfemedia.blog.gov.uk/2021/05/17/the-removal-of-letters-and-sounds-2007-from-the-departments-list-of-validated-phonics-programmes-teachers-questions-answered/

Department for Education and Employment (DfEE). (1998). *The National Literacy Strategy Framework for Teaching*. Sudbury: DfEE Publications.

Department for Education and Skills. Primary National Strategy. (2007). Letters and Sounds: Principles and Practice of High Quality Phonics. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/190599/Letters_and_Sounds_-_DFES-00281-2007.pdf

Department for Education and Skills (DfES) (2006). *Primary National Strategy: Primary Framework for literacy and mathematics.* DfES.

* Dion, E., Brodeur, M., Gosselin, C., Campeau, M., & Fuchs, D. (2010). Implementing research-based instruction to prevent reading problems among low-income students: Is earlier better? *Learning Disabilities Research & Practice, 25*, 87-96.

* Ecalle, J., Magnan, A., & Calmus, C. (2009). Lasting effects on literacy skills with a computer-assisted learning using syllabic units in low progress readers. *Computers and Education, 52*, 554-561.

Economist Intelligence Unit. (2021, 12 June). The reading wars; Literacy. *The Economist*.

* Elbro, C., & Peterson, D. K. (2004). Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia. *Journal of Educational Psychology, 96*, 660-670.

* Fälth, L., Gustafson, S., Tjus, T., Heimann, M., & Svensson, I. (2013). Computer-assisted interventions targeting reading skills of children with reading disabilities—A longitudinal study. *Dyslexia, 19*, 37-63.

* Fawcett, A. J., Nicolson, R. I., Moss, H., & Nicolson, M. K. (2001). Effective reading intervention in junior school. *Educational Psychology, 21*, 299-312.

Fletcher, J., Savage, R., & Vaughn, S. (2020). A Commentary on Bowers (2020) and the Role of Phonics Instruction in Reading. *Educational Psychology Review*. doi:https://doi.org/10.1007/s10648-020-09580-8

Galuschka, K., Ise, E, Krick, K, Schulte-Korne, G. (2014). Effectiveness of Treatment Approaches for Children and Adolescents with Reading Disabilities: A Meta-Analysis of Randomized Controlled Trials. *PLoS ONE, 9*, 1-12. doi:e89900. doi:10.1371/journal.pone.0089900

* Gittelman, R., & Feingold, I. (1983). Children with reading disorders. Efficacy of reading remediation. *Journal of Child Psychology and Psychiatry and Allied Disciplines,, 24*, 167-191.

Goodman, K. (1967). Reading: A psycholinguistic guessing game. *Literacy Research and Instruction, 6*, 126-135.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education, 7*(1), 6-10. DOI: 10.1177/074193258600700104

* Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction: Results for Hispanic and non-Hispanic students. *Journal of Special Education, 39*, 66-85.

* Gunn, B., Smolkowski, K., & Vadasy, P. (2011). Evaluating the effectiveness of Read Well Kindergarten. *Journal of Research on Educational Effectiveness, 4*, 53-86.

Hall, K. (2003). *Listening to Stephen read: Multiple perspectives on literacy*. Buckingham: Open University Press.

* Hatcher, P. J., Hulme, C., & Snowling, M. J. . (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychiatry, 45*, 338-358.

* Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65*, 41-57.
Her Majesty's Inspectorate (HMI). (1990). *The Teaching and Learning of Reading in Primary Schools.* London: Department of Education and Science (DES).
* Hook, P. E., Macaruso, P., & Jones, S. (2001). Efficacy of fast forward training on facilitating acquisition of reading skills by children with reading difficulties—A longitudinal study. *Annals of Dyslexia, 51*, 75-96.
* Houtveen, T., & van de Grift, W. (2012). Improving reading achievements of struggling learners. . *School Effectiveness and School Improvement, 23*, 71-93.
James, M. (2012). Background to Michael Gove's response to the Report of the Expert Panel for the National Curriculum Review in England. *BERA News.* Retrieved from https://www.bera.ac.uk/bera-in-the-news/background-to-michael-goves-response-to-the-report-of-the-expert-panel-for-the-national-curriculum-review-in-england
Jerrim, J., & Moss, G. (2019). The link between fiction and teenagers' reading skills: International evidence from the OECD PISA study. *British Educational Research Journal, 45*, 181-200.
* Kjeldsen, A.-C., Niemi, P., & Olofsson, A. (2003). Training phonological awareness in kindergarten level children: Consistency is more important than quantity. *Learning and Instruction, 13*, 349-365.
* Kozminsky, L., & Kozminsky, E. (1995). The effects of early phonological awareness training on reading success. *Learning and Instruction, 5*, 187-201.
* Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdom: GG rime and GG phoneme. *Reading Research Quarterly, 48*, 61-76.
* Lie, A. (1991). Effects of a training program for stimulating skills in word analysis in first-grade children. *Reading Research Quarterly, 26*, 234-250.
* Loeb, D. F., Gillam, R. B., Hoffman, L., Brandel, J., & Marquis, J. (2009). The effects of Fast ForWord language on the phonemic awareness and reading skills of school-age children with language impairments and poor reading skills. . *American Journal of Speech-Language Pathology, 18*, 376-387.
* Lyster, S. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. *Reading and Writing: An Interdisciplinary Journal, 15*, 261-294.
* Mantzicopoulos, P., Morrison, D., Stone, E., & Setrakian, W. (1992). Use of the SEARCH/TEACH tutoring approach with middle-class students at risk for reading failure. *Elementary School Journal, 97*, 573-586.
* McArthur, G., Sheehan, Y, Badcock, N A, Francis, D A, Wang, H C, Kohnen, S, Banales, E, Anandakumar, T, Marinus, E, Castles, A. (2018). *Phonics training for English-speaking poor readers. Cochrane Database of Systematic Reviews 2018, Issue 11. Art. No.: CD009115.* . London: John Wiley and Sons.
Meek, M. (1988). *How Texts Teach What Readers Learn*. Stroud: Thimble Press.
Michael, I. (1984). Early evidence for whole word methods. In G. Brooks & A. K. Pugh (Eds.), *Studies in the History of Reading*. Reading: Centre for the Teaching of Reading, University of Reading.
* Morris, R. D., Lovett, M. W., Wolf, M., Sevcik, R. A., Steinbach, K. A., Frijters, J. C., & Shapiro, M. B. (2012). Multiple component remediation for developmental reading

disabilities: IQ, socioeconomic status, and race as factors in remedial outcome. *Journal of Learning Disabilities, 45*, 99-127.

* Nancollis, A., Lawrie, B., & Dodd, B. (2005). Phonological awareness intervention and the acquisition of literacy skills in children from deprived backgrounds. . *Speech & Hearing Services in Schools, 36*, 325-335.

* O'Connor, R. E., Notari-Syverson, A., & Vadasy, P. . (1998). First-grade effects of teacher-led phonological activities in kindergarten for children with mild disabilities: A follow-up study. *Learning Disabilities Research & Practice, 13*, 43-52.

Ofsted (2019) School Inspection Handbook, Available at: https://www.gov.uk/government/publications/school-inspection-handbook-eif/school-inspection-handbook (Retrieved 11 August 2021)

* Olusola, O., Lavin, T., Thompson, T., & Ungerleider, C. Pedagogical strategies for teaching literacy to ESL immigrant students: A meta-analysis. *British Journal of Educational Psychology, 81*, 629-653.

Perfetti, C. A. (1995). Cognitive research can inform reading education. . *Journal of Research in Reading, 18*, 106-115.

* Phillips, L. M., Norris, S. P., & Mason, J. M. Longitudinal effects of early literacy concepts on reading achievement: A kindergarten intervention and five-year follow-up. *Journal of Literacy Research, 28*, 173-195.

Pressley, M. (2006). *Reading Instruction that Works: The Case for Balanced Teaching* (3rd ed.). New York: Guilford.

* Regtvoort, A. G. F. M., & van der Leij, A. Early intervention with children of dyslexic parents: Effects of computer based reading instruction at home on literacy acquisition. . *Learning and Individual Differences*, 35-53.

* Reitsma, P., & Wesseling, R. (1998). Effects of computer assisted training of blending skills in kindergarteners. *Scientific Studies of Reading, 2*, 301-320.

Rose, J. (2006). *Independent Review of the Teaching of Early Reading: Final Report*. Nottingham: DfES Publications.

* Rothe, E., Grunling, C., Ligges, M., Fackelmann, J., & Blanz, B. (2004). Erste Auswirkungen eines Trainings der phonolo gischen Bewusstheit bei zwei unterschiedlichen Altersgruppen im Kindergarten [First effects of training phonological awareness in two age cohorts of kindergarten children]. *Zeitschrift für Kinder- und Jugendpsychiatre und Psychotherapie, 32*, 167-176.

* Ryder, J. F., Tunmer, W. E., & Greaney, K. T. (2008). Explicit instruction in phonemic awareness and phonemically based decoding skills as an intervention strategy for struggling readers in whole language classrooms. *Reading and Writing: An Interdisciplinary Journal, 21*, 349-369.

Savage, R., Burgos, G., Wood, E., & Piquette, N. (2015). The Simple View of Reading as a framework for national literacy initiatives: a hierarchical model of pupil-level and classroom-level factors. *British Educational Research Journal, 41*(5), 820-844.

* Schachter, J., & Jo, B. (2005). Learning when school is not in session: A reading summer day-camp intervention to improve the achievement of exiting first-grade students who are economically disadvantaged. *Journal of Research in Reading, 28*, 158-169.

* Segers, E., & Verhoeven, L. (2005). Long-term effects of computer training of phonological awareness in kindergarten. *Journal of Computer Assisted Learning, 21*, 17-27.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143-147.

* Spörer, N., Brunstein, J. C., & Kieschke, U. . (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction, 19*, 272-286.
Solity, J. (2020). Instructional psychology and teaching reading: Ending the reading wars. *The Educational and Developmental Psychologist, 37*, 123-132. doi:10.1017/edp.2020.18
Suggate, S. (2016). A Meta-Analysis of the Long-Term Effects of Phonemic Awareness, Phonics, Fluency, and Reading Comprehension Interventions. *Journal of Learning Disabilities, 49*(1), 77-96.
Sullivan, A., & Brown, M. (2013). *Social inequalities in cognitive scores at age 16: The role of reading*. Retrieved from UCL Institute of Education, Centre for Longitudinal Studies: https://cls.ucl.ac.uk/wp-content/uploads/2017/04/CLS-WP-2013-10-.pdf
Torgerson, C., Brooks, G., Gascoine, L., & Higgins, S. (2019). Phonics: reading policy and the evidence of effectiveness from a systematic 'tertiary' review. *Research Papers in Education, 34*, 208-238.
* Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia, 60*, 40-56.
* Vadasy, P. F., Jenkins, J. R., & Pool, K. (2000). Effects of tutoring in phonological and early reading skills on students at risk for reading disabilities. . *Journal of Learning Disabilities, 33*, 579-590.
* Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006). Code oriented instruction for kindergarten students at risk for reading difficulties: A randomized field trial with paraeducator implementers. *Journal of Educational Psychology, 98*, 508-528.
* 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*, 786-803.
* Vadasy, P. F., & Sanders, E. A. (2011). Efficacy of supplemental phonics-based instruction for low-skilled first graders: How language minority status and pre-test characteristics moderate treatment response. *Scientific Studies of Reading, 15*, 471-497.
* Vadasy, P. F., & Sanders, E. A. . (2012a). 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*, 821-843.
* Vadasy, P. F., & Sanders, E. A. (2012b). Two-year follow-up of a kindergarten phonics intervention for English learners and native English speakers: Contextualizing treatment impacts by classroom literacy instruction. *Journal of Educational Psychology, 104*, 987-1005.
* van der Kooy-Hofland, V. A. C., van der Kooy, J., Bus, A. G., van IJzendoorn, M. H., & Bonsel, G. J. . (2012). Differential susceptibility to early literacy intervention in children with mild perinatal adversities: Short- and long-term effects of a randomized control trial. *Journal of Educational Psychology, 104*, 337-349.
* van Keer, H. (2004). Fostering reading comprehension in fifth grade by explicit instruction in reading strategies and peer tutoring. *British Journal of Educational Psychology, 74*, 37-70.

* Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Hagan, E. C., . . . & 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*, 449-487.
* Vaughn, S., Linian-Thompson, S., Mathes, P. G., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., . . . & Francis, D. J. (2006). Effectiveness of Spanish intervention for first-grade English language learners at risk for reading difficulties. *Journal of Learning Disabilities, 39*, 56-73.
* Vaughn, S., Mathes, P., Linian-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S. D., . . . & Francis, D. J. (2006). Effectiveness of an English intervention for first grade English language learners at risk for reading problems. *Elementary School Journal, 107*, 153-180.
* Vaughn, S., Cirino, P. T., Tolar, T., Fletcher, J. M., Cardenas-Hagan, E., Carlson, C. D., & Francis, D. J. . (2008). Long-term follow-up of Spanish and English interventions for first-grade English language learners at risk for reading problems. *Journal of Research on Educational Effectiveness, 1*, 179-214.
* Warrick, N., Rubin, H., & Rowe-Walsh, S. . (1993). Phoneme awareness in language-delayed children: Comparative studies and intervention. *Annals of Dyslexia, 43*, 153-173.
Waterland, L. (1985). *Read With Me. An Apprenticeship Approach to Reading*. Stroud: Thimble Press.
* Williams, J. P. (1980). Teaching decoding with an emphasis on phoneme analysis and phoneme blending. *Journal of Educational Psychology, 72*, 1-15.
* Wolff, U. (2011). Effects of a randomised reading intervention study: An application of structural equation modelling. *Dyslexia, 17*, 295-311.
Wray, D., & Medwell, J. (1994). *Teaching Primary English: The State of the Art*. London: Routledge.
Wyse, D. (2010). Contextualised Phonics Teaching. In K. Hall, U. Goswami, C. Harrison, S. Ellis, & J. Soler (Eds.), *Interdisciplinary Perspectives on Learning to Read: Culture, Cognition and Pedagogy.* London: Routledge.
Wyse, D., & Ferrari, A. (2014). Creativity and Education: Comparing the national curricula of the states of the European Union with the United Kingdom. *British Educational Research Journal, 41*(1), 30-47.
Wyse, D., & Jones, R. (2001). *Teaching, English, Language and Literacy*. London: RoutledgeFalmer.
Wyse, D., & Goswami, U. (2008). Synthetic phonics and the teaching of reading. *British Educational Research Journal,* 34(6), 691-710.
Wyse, D., & Torrance, H. (2009). The development and consequences of national curriculum assessment for primary education in England. *Educational Research, 51*(2), 213-228.

Appendix 1: Details in final selection of studies including quality assessments[4]

References	Sample	Age of children approx	Context for study	Reading Comp. measure?	Orientation: SP; BL or WL	Delivered by
Gunn, B., Smolkowski, K., & Vadasy, P. (2011). Evaluating the effectiveness of Read Well Kindergarten. Journal of Research on Educational Effectiveness, 4, 53–86.	54 teachers participated across the 3 school years, 26 in the RWK condition and 28 in the control condition, in the 24 schools. There were 37 intervention classes, 23 full day and 14 half 59 day. Of the 41 comparison classes, 23 were full-day classes and 18 were half-day classes. Within those classrooms, we assessed 1,519 kindergarten students in the fall and 1,427 in the spring.	5 to 6	USA Oregon and New Mexico	yes	BI	Classroom teachers
Kozminsky, L., & Kozminsky, E. (1995). The effects of early phonological awareness training on reading success. Learning and Instruction, 5, 187–201.	Seventy students from two adjacent public kindergarten classes participated in thestudy.	5 years 2 months mean	Israel. Hebrew language.	yes	BI	Trained student teachers

[4] Additional columns of the table are located halfway down the table.

Lie, A. (1991). Effects of a training program for stimulating skills in word analysis in first-grade children. Reading Research Quarterly, 26, 234–250.	The subjects were 208 students enrolled in first-grade classes in Halden (Norway) at the time of the study. Out of a total of 15 classes in 9 schools, 10 classes were selected from 7 schools, using as much as possible the criterion of deliberate sampling for heterogeneity.	6 to 8	Norway. Norwegian language.	yes	BI	classroom teachers
Lyster, S. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. Reading and Writing: An Interdisciplinary Journal, 15, 261–294.	A total of 273 monolingual Norwegian children attending 25 different preschool groups in two different communities outside Oslo participated in the study. The analyses presented here are partly run on the basis of results from the 237 non-reading children who were still available at the end of grade 1 and partly on the basis of the results from the 225 non-reading children about whom there was information about the mothers' education.	5 to 6 Their age was then 5 years 10 months to 6 years 9 months.	Norway. Norwegian language.	yes	Phonics but not SP	classroom teachers

Phillips, L. M., Norris, S. P., & Mason, J. M. (1996). Longitudinal effects of early literacy concepts on reading achievement: A kindergarten intervention and five-year follow-up. Journal of Literacy Research, 28, 173–195.	At the beginning of kindergarten, there were 318 children in the sample on whom Metropolitan Reading Readiness Test Level 1 reading achievement data were collected. By the end of fourth grade, this had reduced to 214 on the CTBS Level 10 or 67% of the original number (instruments are described in the next section).	5 years 2 months mean	Canada.	yes	WL	classroom teachers or parents
Vadasy, P. F., & Sanders, E. A. (2011). Efficacy of supplemental phonics-based instruction for low-skilled first graders: How language minority status and pretest characteristics moderate treatment response. Scientific Studies of Reading, 15, 471–497.	After attrition, the final sample included 93 treatment students (48 LM students) and 94 controls (50 LM students) from 29 classrooms across 11 schools.	6 to 7	USA	yes	BI	paraeducators

Vadasy, P. F., & Sanders, E. A. (2012). 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, 821–843. doi:10.1007/s11145-012-9393-4	In October of 2007–2008, all students (n = 903) in first-grade classrooms at 13 U.S. urban public elementary schools (45 first-grade classrooms) in the Pacific Northwest known for relatively large proportions of English learner (EL) student enrolment were invited to participate in the research study. EL 95. Non-EL 85	6 to 7	USA	yes	BI	paraeducators.
Vadasy, P. F., & Sanders, E. A. (2012). Two-year follow-up of a kindergarten phonics intervention for English learners and native English speakers: Contextualizing treatment impacts by classroom literacy instruction. Journal of Educational Psychology, 104, 987–1005. doi:10.1037/a0028163	Finally, we were able to capture follow-up classroom literacy block observation data (procedures described below) for 106 (62 LM and 44 non-LM) of the 137 students who had at least two test waves completed (including seven of the eight retained students).	5 to 6	USA	yes	BI	paraeducators

References	Total Hours of teaching time for reading programme	Allocation to groups	Fidelity measures	Main findings	Unweighted Average Effect Size at Follow-Up	Quality Rating (High; Medium; Low)
Gunn, B., Smolkowski, K., & Vadasy, P. (2011). Evaluating the effectiveness of Read Well Kindergarten. Journal of Research on Educational Effectiveness, 4, 53–86.	57.5	Random assignment schools	Yes	Analyses of final outcomes revealed a statistically significant difference favoring intervention students on the curriculum-based measures of sight words and decodable words at the end of kindergarten, with no significant differences at either fall or spring follow-up assessments in first grade. The main effects for other measures, however, were not statistically significant. We were therefore unable to confirm benefit of RWK for students on standardized measures of sight word reading, decoding, phonemic awareness, and vocabulary as well as ORF.	-0.13	High

Kozminsky, L., & Kozminsky, E. (1995). The effects of early phonological awareness training on reading success. Learning and Instruction, 5, 187–201.	45	Random assignment kindergartens	not clear	These gains in phonological awareness were accompanied by differences in reading comprehension scores. The scores of the experimental group exceeded those of the control population by about one-half standard deviation at the end of first and third grades (SD = 0.56 and 0.46, respectively). This difference demonstrates the sustained effect of pre-school phonological awareness training on success in learning to read in the primary grades. This effect is confirmed by the significant positive correlations found between post-training phonological awareness measures (LAC and PAT) and first-grade reading comprehension for the experimental group.	0.70	Low

Lie, A. (1991). Effects of a training program for stimulating skills in word analysis in first-grade children. Reading Research Quarterly, 26, 234–250.	Approximately 38	Random assignment classes	not clear	The tendency however is clear: The training program in sequential analysis as well as the training in positional analysis had a facilitating effect on both reading and spelling acquisition.	0.44	Medium
Lyster, S. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development development. Reading and Writing: An Interdisciplinary Journal, 15, 261–294.	Approximately 9.1	Random assignment children	not clear	The comparison procedure showed that the Morphological Group outperformed both the other groups on Word Reading and that both experimental groups had significantly better performances on Text reading than the Control Group. Group differences were close to being significant for Sentence Reading, F(2,232) = 2.90, P = 0.57. No significant differences were found between the groups on Phonological Coding, F(2,332) = 1.63. The stress on phonics in most classrooms and the	0.27	Medium

				relatively regular orthographic structure of the Norwegian language might, to some extent, explain this result.		
Phillips, L. M., Norris, S. P., & Mason, J. M. (1996). Longitudinal effects of early literacy concepts on reading achievement: A kindergarten intervention and five-year follow-up. Journal of Literacy Research, 28, 173–195.	Flexible programme based on weekly use of booklets.	Random assignment schools	not clear	The results showed positive effects on reading achievement because of the treatments that lasted until the end of second grade for all treatment groups, and more modest positive effects that lasted until the end of fourth grade for theschool-only treatment group. In addition, it was argued that the increase in reading achievement was attributable to students' increased knowledge of	0.31	Medium

				early literacy concepts at the end of kindergarten, which in turn was due to the treatment.		
Vadasy, P. F., & Sanders, E. A. (2011). Efficacy of supplemental phonics-based instruction for low-skilled first graders: How language minority status and pretest characteristics moderate treatment response. Scientific Studies of Reading, 15, 471–497.	60	Random assignment children	Yes	Despite the positive findings overall for treatment, results showed that LM students tended to exhibit lower treatment effects than non-LM students, although significantly lower only on passage reading fluency. Simple effect sizes (Cohen's d) reported in Table 2 for LM students are approximately one third of the size of effects for non-LM students.	0.39	Medium

Vadasy, P. F., & Sanders, E. A. (2012). Two-year follow-up of a code-oriented intervention for lower-skilled first-graders: The influence of language status and word reading skills on thirdgrade literacy outcomes. Reading and Writing, 26, 821–843. doi:10.1007/s11145-012-9393-4	36	Random assignment children	Yes	In our original study, we found large, significant treatment benefits of a code oriented intervention for first-grade English learner (EL) and native English speaking (non-EL) students on a host of literacy outcomes, including word reading, spelling, and reading comprehension. The current study, which follows the original sample participants into grades two and three, shows that treatment benefits were maintained across all three outcomes for non-ELs, and for two of the three measures (all but spelling) for EL children.	0.19	Medium

Vadasy, P. F., & Sanders, E. A. (2012). Two-year follow-up of a kindergarten phonics intervention for English learners and native English speakers: Contextualizing treatment impacts by classroom literacy instruction. Journal of Educational Psychology, 104, 987–1005. doi:10.1037/a0028163	36	Random assignment children	Yes	This current study's findings show that kindergarten supplemental phonics intervention continued to have advantages for LM and non-LM students 2 years postintervention. For LM children, the advantages, after controlling for covariates, were for word level outcomes (i.e., word reading and spelling). For non-LM children, advantages were significant for word level, fluency, and comprehension outcomes.	0.23	Medium