Received: 26 October 2021 | Accepted: 9 March 2022

DOI: 10.1002/rev3.3349

STATE OF THE ART REVIEW Review of Education BERA

Instructional Psychology and Teaching Reading: An analysis of the evidence underpinning government policy and practice

Jonathan Solity[1,2]

[1]Optima Psychology, Leamington Spa, UK
[2]University College London, London, UK

Correspondence
Jonathan Solity, Optima Psychology, Leamington Spa, UK.
Email: jonathan.solity@optimapsychology.com

Funding Information
This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Abstract

There have been few areas in England over the last 50 years where government has drawn more heavily on research to inform policy and practice than in the area of teaching reading. The focus of this article is an analysis of the research and evidence on early reading, in particular the role of phonics, on which government policy in England and the practice it promotes, are based. The article has three parts. The first examines the major policy initiatives in teaching reading from the Plowden Report in 1967 to the present day and provides an overview of their implications for practice. It highlights the tensions between opposing sides in the 'reading wars' between those prioritising phonics first, fast and only, versus those promoting a whole language approach to teaching reading. The second examines the evidence on which the government has drawn and the way in which teaching reading has become increasingly politicised in England to promote specific practices, some of which, but not all, are underpinned by research. Part 3 introduces instructional psychology that provides an alternative perspective on teaching phonics, whole language and the reading wars. The approaches to teaching

reading generated by instructional psychology will be summarised and research presented to demonstrate how, despite the rhetoric surrounding the positive impact of current government policy on teaching reading, it is potentially a major cause of reading difficulties. Part 3 ends by highlighting the implications of instructional psychology for future policy and practice.

KEYWORDS
decodable texts, instructional psychology, real books, synthetic phonics

Context and implications

Rational for this study

England is unique in the English speaking world in having a strategy for teaching early reading that has been developed over time across the whole country. This article provides the first detailed analysis of the research that has informed policy and practice over the last 50 years that has led to the current focus on teaching reading through systematic synthetic phonics and decodable texts.

Why the new findings matter

Government recommended practice is vulnerable to being amended or replaced if not well grounded in research. The article identifies current policies and practice that are, and are not, supported by evidence in order that future practice builds on what has been learned in the past, rather than, like a pendulum, swings back to revisit the debates that have dominated discussion, and been defined as, the reading wars. The article has not only identified what to teach but through the discussion on instructional psychology provided a possible explanation as to why certain methodologies are effective.

Implications for practitioners, policy makers, researchers

Long-standing debates about the most effective methods for teaching early reading have remained largely unresolved over the last 50 years. The analysis of research and practice discussed in this article, together with the description of instructional psychology, highlight one possible way of resolving the tension between advocates of teaching reading through phonics or whole language. This involves teaching a small number of the most frequently occurring and generalisable phonic skills in children‘s written English through real books rather than reading schemes. Such a position also raises a range of novel instructional questions that are best resolved through research rather than rhetoric.

PART 1: POLICY AND PRACTICE IN TEACHING READING—THE RHETORIC

There are many factors that have influenced how reading has been taught in England in recent years, and these vary according to political perspectives as well as views about the nature of teaching and learning. So any attempt to identify the precise factors that contribute to changes in government policy and recommended practice over a period of 50 years is a precarious balance between rhetoric and reality. Nevertheless the direction of travel over this period is twofold: increased governmental control over the curriculum and a move towards teaching early reading through what is termed systematic synthetic phonic (SSP). The focus of this article is on early reading in England and the extent to which the evidence supports the government's policies and recommended practice in teaching reading.

Part 1 has four phases documenting the significant changes that have contributed to the way reading is currently being taught. Phase 1 can be characterised as a conflict between two opposing sides in what have been called 'The Reading Wars' (Castles et al., 2018; Solity & Vousden, 2009; Wyse & Bradbury, 2022). These highlighted the tensions between phonic based and whole language approaches that were reflected in the Bullock (DES, 1975) and Plowden (CACE, 1967) reports, and their impact on teaching phonics and on reading standards. Turner (1990) initiated a politically charged debate when claiming there was strong evidence that reading standards had declined significantly since Plowden due to the absence of rigorous phonics teaching and the use of real books (books written by a variety of authors where no attempt is made to restrict the content in terms of vocabulary, word frequency or phonic regularity, or to sequence books in order of perceived difficulty), which are seen to represent a whole language approach to teaching reading, and are perceived to be the antithesis of a phonics-based approach.

Phase 2 saw an increased role for teaching phonics through the introduction of the National Literacy Strategy (NLS; DfEE, 1998). The resulting methodology, and rationale on which it was based, were increasingly questioned when the NLS failed to have the desired impact on standards. This led to a review of teaching reading by the Education and Skills Committee (2005) with the Rose Review (2006) recommending that SSP should be the primary methodology through which reading is taught. This view was endorsed as government policy through the publication of *Letters & Sounds* (DfES, 2007), which was made available to all primary schools in England.

Phase 3 represents the increasing influence of government on reading policy and classroom practice from 2010 to 2021. Nick Gibb (Minister of State for School Standards 2010–2012 and 2015–2021) was a member of the Education and Skills Committee and, following the election of the Conservative Liberal-Democrat coalition government in 2010, became the Minister of State for School Standards. Gibb has been a significant figure in requiring schools to teach reading through SSP, a policy that has been enshrined in the Phonics Screening Check (DfE, 2012, the National Curriculum (DfE, 2013a), and Ofsted Inspection Framework (2019).

Phase 4 saw the publication of 16 core criteria through which commercial SSP programmes will be validated (DfE, 2021a). This represented a change in direction through effectively requiring schools to teach reading through commercial SSP programmes that include a phonically decodable reading scheme, rather than using a programme, such as *Letters & Sounds*, that was free and not accompanied by decodable texts. The government's commitment to this change in direction was demonstrated in two further publications (DfE, 2021b,c).

Phase 1

The phonically irregular nature of written English has resulted in contrasting views about how best to teach reading, which are not only about methodology but educational philosophies, between teacher-directed versus child-centred, or traditional versus progressive education. It is beyond the scope of this article to review the origins of these philosophies but the publication of the Plowden Report (CACE, 1967) gave fresh impetus to child-centred education in the UK. Plowden drew on Piagetian theory, emphasised child-centred education and stressed that, 'at the heart of the educational process lies the child' (p. 7). Teachers were encouraged to immerse children in high quality books and it was suggested they would learn to read when they were ready. The notion of direct teaching was summarily rejected and seen as the antithesis of a child-centred approach. The philosophy espoused in Plowden became the new orthodoxy and was reflected in the writings of a number of leading educationalists at the time. These included Goodman (1976, 1986) and Smith (1973, 1978) in the USA who saw that learning to read was a natural process, like learning to talk and walk, and could not be taught explicitly.

Plowden also drew attention to practice that is perceived to have dominated the way reading was taught until the introduction of the NLS (DfEE, 1998):

> The most successful infant teachers have refused to follow the wind of fashion and to commit themselves to any one method. They choose methods and books to fit the age, interest and ability of individual pupils. Children are helped to read by memorising the look of words and phrases, often with the help of pictures, by guessing from a context which is likely to bring them success, and by phonics, beginning with initial sounds. They are encouraged to try all the methods available to them and not to depend on only one method.
>
> (Vol. 1, p. 212)

> Instead of relying on one reading scheme, many teachers use a range of schemes with different characteristics, selecting carefully for each child: some schemes emphasise sight reading, others phonics; some consist of short books, with a very slow build up of vocabulary, and suit children who need quick success; other schemes help children who are able to advance rapidly and to discard primers. Reading schemes should never determine the practices adopted for all children.
>
> (Vol. 1, p. 212)

The emphasis was on building on children's interests, experiences and knowledge and it was felt that they could not be pushed into formal reading before they were ready, giving rise to the concept of reading readiness. Plowden effectively advocated a balanced approach to teaching reading, reinforcing the view that learning to read was an individual process and that no one method would succeed with all children. Such a view offers a sharp contrast to the DfE's (2021a,b,c) most recent pronouncements on teaching reading, a contrast that is at the heart of the reading wars.

The Bullock Report (DES, 1975) was a major report that examined the way that English was taught in primary and secondary schools and described the debate about teaching reading in the following terms:

> Some would put so much emphasis on the 'mechanics' of reading that certain children would be handicapped rather than helped … [whereas] others advocate

> so keenly the virtues of mature reading from the beginning that they are in danger of leaving it too much to trust that the skills will be acquired on the way.
>
> (p. 77–78)

It should be noted that there was little evidence available to inform debates which were largely driven by rhetoric rather than research.

In 1976, Jim Callaghan (the British Prime Minister between 1976 and 1979) gave his landmark Ruskin College Speech, signalling the point at which the field of education became increasingly politicised. At the time of the speech, the Secretary of State for Education had just three powers over schools, but by 2016 they had more than 2500 (Miller, 2016). Much of the political debate in the following years addressed the freedom of the teaching profession to determine the curriculum and teaching methodology. It was increasingly argued that there was too much variation in how children were being taught. This culminated in the introduction of the National Curriculum (DES, 1988) and annual, nationwide, assessments of pupils' attainments to monitor standards. The National Curriculum provided a framework for teaching reading and introduced the levels pupils were expected to reach as they progressed through primary to secondary school.

The debate about standards began to shift from the general impact of opposing educational philosophies, more specifically to reading and the teaching of phonics with the publication of 'Sponsored Reading Failure' (Turner, 1990) by a local authority (LA) educational psychologist who claimed there was strong evidence from several London LAs that reading standards had declined significantly over a number of years. Although the LAs used different tests, over different time scales with varying numbers of students, the data presented was seen to represent a decline in reading standards. The cause of the decline was attributed to the continuing reading wars, the absence of rigorous phonics teaching and the use of real books in schools. Turner suggested that a further decline in reading standards could be predicted if the real books approach was widely adopted.

Her Majesty's Inspectorate (HMI; DES, 1990) responded almost immediately with a report drawing on inspections in 120 schools. They found: (i) the majority of teachers used a blend of methods to teach reading; (ii) a clear link between higher standards and systematic phonic teaching; and (iii) that published schemes of graded reading books were used in more than 95% of the classes. Although standards in reading in about 20% of schools were judged poor and required attention, the overall pattern of results suggested that little had changed since the primary survey of 1978 (DES, 1978). [Corrections made on 04 June 2022, after first online publication: Changes were made in the previous sentence for clarity.] This appeared to refute claims that the real books approach was either used widely or having a damaging impact on standards, or that classroom practice had been unduly constrained by the tensions created by the reading wars.

However, Ofsted's (1996) report on reading standards in 45 inner London primary schools in three LAs told a different story. The inspections focused on how reading was being taught in Years 2 (7 year olds) and 6 (11 year olds) and involved classroom observations, questionnaires completed by teachers, head teachers and representatives from the LAs, and a common reading test administered by the National Foundation for Educational Research (NFER) to over 800 pupils in each of the two year groups being inspected. Reading standards were found to be unacceptably low, with approximately only one-in-five pupils in Year 2 achieving a reading age (RA) at or above their chronological age (CA). In Year 6, 40% of pupils achieved a RA at or above their CA, but more than 40% of pupils had a RA that was 2 years or more behind their CA. HMI concluded that a major problem was a commitment to methods of teaching reading that were 'self-evidently' not working and that far too few schools had a clear idea about the place and purpose of phonic work and how it should be taught. In contrast, when reading was well taught, phonics contributed to the accuracy and fluency of children's reading.

The presence or absence of teaching phonics was increasingly the issue in debates about standards of literacy and reading failure. The main source of evidence for teachers to draw on came from HMI inspections of primary classrooms whereas the evidence from experimental research tended to focus on identifying the cause of pupils' difficulties in learning to read.

Phase 2

Given the concerns about reading standards, David Blunkett, the Shadow Secretary of State for Education and Employment established the Literacy Task Force (1997) that was asked to develop a strategy for substantially raising standards of literacy in primary schools in England. It is perhaps worth noting that the aim of any resulting strategy was that, 'by the end of a second term of a Labour government, all children leaving primary school will have reached a reading age of at least eleven' (p. 7). Over 20 years after the Bullock Report, the Task Force hinted that the reading wars might be nearing an end when it noted:

> there have been few more vigorous educational controversies in the last decade than the one over how reading should be taught. Opposing sides in a vigorous national debate took to the barricades with banners proclaiming their loyalty to 'phonics' or 'real books.' But while this debate has raged, research and the understanding of best practice have moved on.
>
> (p. 16)

The Task Force acknowledged current best practice and the potential impact of the recently developed National Literacy Project (NLP; DfEE, 1996) which ran from January 1997 to March 1998 and was a whole class approach to teaching reading. It introduced the daily literacy hour that divided the curriculum into four sections, which started with text level work followed by sentence and word level work with a final section devoted to a whole class review. The activities and time spent on each level were specified in considerable detail. It is interesting to note that the Task Force advised that an independent evaluation should be undertaken involving a comparison of the schools participating in the trial of the NLP with a control group, with the specific aim of showing how the model could be refined and built upon to form the basis of a national approach to the teaching of literacy in primary schools. In the intervening 25 years the recommendation from the Task Force that governments research and evaluate the impact of the interventions to teaching reading that they recommend has never been undertaken in the UK. Research has compared an intervention with untreated controls ('business as usual') but not an alternative, equivalent intervention. Instead the NLP was evaluated through HMI visits to approximately 20% of the participating schools and standardised reading tests carried out by the NFER. The NLP was the pilot version of what became the NLS (DfEE, 1998) which was introduced to all primary schools in England in September 1998.

For reasons that will be discussed in Part 2, concerns continued to be expressed about teaching phonics after the introduction of the NLS, despite the development of several programmes designed to provide greater detail on teaching of phonics (*Progression in Phonics*, DfEE [1999b]; *Additional Literacy Support*, DfEE [1999c]; *Early Literacy Support*, DfES [2001]; *Playing with Sounds*, DfES [2004]), none of which had been researched experimentally. The Education Select Committee (2005) investigated methods for teaching reading and reported that at age 11 around 20% of children still did not achieve the success in reading (and writing) expected for their age, a proportion perceived to be unacceptably high. The report continued, 'there is a wide variation in the results achieved by schools with apparently similar intakes.

This differential achievement suggests that problems do exist, either in the implementation of the government's strategies or inherently in the methodology it promotes' (p. 3).

The Committee also noted the continued tension between two contrasting approaches: those who argued that phonics programmes should have greater prominence in the early teaching of reading whereas others preferred to focus on the development of vocabulary and the enrichment of linguistic experience. The views of those supporting phonics first and fast were summarised by the committee in their final report:

> An important feature of the phonics approach is that children are only taught to read through texts fully within their current phonological ability. So, although children might encounter words they do not understand, they are not given texts they cannot decode and are therefore not expected to infer words from context or syntax. Many supporters of phonics deplore what they see as 'guessing' strategies and view them as actively damaging.
>
> (para. 26)

Contrary to the optimistic tone of the Literacy Task Force (1997), that the reading wars were drawing to a close, those giving evidence to the committee indicated the debates were as bitter, rivalrous and intense as ever.

Overall the report highlighted the tensions between two different approaches to teaching phonics (analytic and synthetic, see Part 2), suggesting that there were significant problems with the searchlight metaphor of reading that underpinned the NLS theoretically (Stannard & Huxford, 2007) and indicating that for some proponents of teaching phonics: (i) pupils should only read books that are consistent with their phonic knowledge; and (ii) perhaps most controversially, that it was more important that pupils could decode the words they read phonically rather than understand what they read. This was seen to be preferable to the alternative which was perceived to be pupils reading words through looking at picture cues, using context or guessing.

This was the first public acknowledgement, that for advocates of phonics first and fast, decoding was more important than understanding. Those giving oral evidence explained that if pupils could not decode what was written they would not be able to understand what they had read. This view was written into the revised National Curriculum (DfE, 2013a), which stated that alongside their knowledge of grapheme-phoneme-correspondences (GPCs), pupils need to develop the skill of blending the sounds into words which 'will be supported by practice in reading books consistent with their developing phonic knowledge and skill and their knowledge of common exception words' (p. 9).

Following the publication of the Select Committee report, Ruth Kelly, the Secretary of State for Education, invited Sir Jim Rose (formerly Her Majesty's Chief Inspector of Primary Education and Director of Inspection for the Office for Standards in Education in England) to investigate best practice in teaching early reading and synthetic phonics. A major finding of the review was that, 'the practice seen by the review shows that the systematic approach, which is generally understood as "synthetic" phonics, offers the vast majority of young children the best and most direct route to becoming skilled readers and writers' (Rose, 2006, p. 4). The review recognised the value of decodable books so that pupils could apply their phonic knowledge and experience early success in reading to boost their confidence and motivation. However, the review also acknowledged that the text in some favourite children's books could fulfil the same function as decodable readers. Significantly, the review continued that 'real books' could be used in parallel or possibly in place of decodable texts so that children are not denied 'access to favourite books and stories at any stage and particularly at the point when they need to read avidly to hone their skills, as the focus shifts from learning to read to reading to learn' (p. 27).

There were two notable outcomes of the review. The first was that the searchlights metaphor of reading that underpinned the NLS, and was based on a balanced literacy perspective, was seen to be theoretically and practically questionable and was replaced with the Simple View of Reading (Gough & Tunmer, 1986). The second was the publication of *Letters & Sounds* (DfES, 2007) that was offered to schools as an example of good practice in teaching synthetic phonics.

Phase 3

After the election of the Conservative Liberal-Democrat Coalition Government in May 2010 successive governments have become more involved than at any time in the past in providing clear details on how word recognition is to be taught through SSP programmes. Ofsted (2010) reported in *Reading By Six* on the practice of 12 outstanding schools in teaching reading, based on Ofsted inspections, rather than experimental research into 'what works'. The report could be seen to be a manifesto for teaching synthetic phonics. Christine Gilbert, Her Majesty's Chief Inspector, in her introduction wrote that rigorous, intensive and systematic phonics teaching underpins reading, spelling and writing. Systematic phonics programmes were defined as following a specific order for teaching GPCs through clearly prescribed methodologies that frequently involved scripted lessons. Gilbert claimed that the challenge for all schools was to emulate the identified practices in the 12 schools that are eminently transferable and which should be applied consistently and reliably everywhere.

The report resulted in the government adopting what is called SSP as the fundamental approach to teaching reading. The Department for Education (DfE) commissioned a catalogue of 'approved' SSP programmes that had been evaluated by assessors to ensure that the recommended programmes met strict criteria (ESPO, 2011). The DfE (2010) listed 11 criteria that authors and publishers had to meet to become a recommended programme, the first of which was:

> Presents high quality systematic, synthetic phonic work as the prime approach to decoding print, i.e. a phonics 'first and fast' approach.

Publishers of SSP programmes were required to confirm not only that they were teaching synthetic phonics but were teaching a version of synthetic phonics that was accepted and approved by the government assessors. Approved programmes were required to use phonically regular reading schemes that were entirely decodable so that pupils experienced success and learned to rely on phonemic strategies. The government's insistence on synthetic phonics, and determination to ensure that the approach was adopted by all schools, was further reflected in their offer of matched funding, up to a maximum of £3000, when schools bought programmes from the catalogue. So schools were effectively getting £6000 of phonic resources for £3000. Seven mainstream SSP programmes were listed in the catalogue of approved programmes (ESPO, 2011) but whose impact had not been compared to an alternative intervention along the lines recommended by the Literacy Task Force (1997). One additional mainstream programme was endorsed a year later (ESPO, 2012). In total the DfE provided matched funding of £23.7 million (DfE, 2015) although Clark (2018) reported, using the Freedom of Information Act, that the actual figure spent on commercial materials and training for teachers was double the DfE figure (£46 million).

One of the key differences between the recommended approaches to teaching reading before 2010 and after the election of the coalition government, is that prior to 2010 the DfES took responsibility for promoting the way reading was taught and for writing and making

programmes available to schools at no cost, whereas since 2010 the curriculum and the way reading is taught is now predominantly in the hands of the authors and publishers of SSP programmes. What they have in common is that the impact of the programmes advocated have never been investigated through randomised controlled trials (RCTs) and compared to either a control group or an alternative intervention along the lines recommended by the Education and Skills Committee (2005), the Literacy Task Force (1997) and Torgerson et al. (2006).

The role of SSP was further reinforced through the introduction of an annual, Year 1 (pupils aged six) Phonics Screening Check (PSC; DfE, 2011) to assess pupils' mastery of GPCs and blending skills. The PSC was introduced in 2012 and has always had a pass mark of 32 (maximum 40). Anyone falling below this figure was required to repeat the PSC at the end of Year 2 (aged seven). The increasing emphasis on SSP as the most effective way to teach phonics since the report of the Education Select Committee, led to it being endorsed in the revised National Curriculum (DfE, 2013a), the associated Programmes of Study (DfE, 2013b) and Ofsted Education Inspection Framework (Ofsted, 2019) which all reiterated that pupils should be taught to read systematically through synthetic phonics and books that match the children's phonic knowledge.

Phase 4

Phase 4 represents continuity with Phase 3 but intensified the government's commitment and determination to ensure that all schools taught early reading through a commercial SSP programme. Three DfE publications (2021a,b,c) sent clear messages to authors and publishers of SSP programmes, teachers and literacy leaders about the government's view of what constitutes an SSP programme. They reflected the philosophy of those whose evidence to the Education Skills Committee (2005) can be characterised as phonics first, fast and only, a philosophy that reinforces the rationale and content of the National Curriculum and PSC.

The DfE (2021a) published 16 core criteria that publishers had to meet if they wished their SSP programme to be validated. The core criteria make clear the current DfE requirements of a SSP programme and state that only 'complete' programmes will be validated. A complete programme is defined as: 'one that provides all that is essential to teach SSP to children in Reception and Key Stage 1'. 'All that is essential' is the key phrase and refers to whether an SSP programme includes a reading scheme containing decodable texts that reflect the phonic skills being taught. The 16 core criteria embrace five major themes:

- systematic phonics is the primary approach through which early reading is taught
- phonics is taught early in Reception through a sequence that progresses from simple to more complex knowledge covering all the major GPCs in English
- children are taught to segment words into their constituent phonemes for spelling
- children are taught a small number of common exception words through highlighting the graphemes that are not consistent with the phonic knowledge that they are taught
- phonically decodable texts are provided to enable pupils to apply their developing knowledge of GPCs to experience success so that pupils only read books that are consistent with their phonic knowledge.

There are some 'do nots' as well that refer to teaching common exception words (words that are phonically irregular) with reference to their shape or attempting to decode unknown words through guessing, looking at picture cues or the use of context. The DfE announced in December 2021 that 19 programmes had been validated and are now available commercially, potentially at a considerable cost to schools.

The criteria reflect the core belief of those advocating phonics first and fast, that were expressed to the Education Skills Committee, and in the DfE's (2010) list of 11 core criteria that publishers of SSP programmes had to meet, that decoding initially takes priority over understanding:

> A programme should promote the use of phonics as the route to reading unknown words, before any subsequent comprehension strategies are applied.
>
> (DfE, 2021a)

The 16 core criteria and Reading Framework demand that unknown words are read through knowledge of phonics first and only, before pupils are taught any other aspects of the reading curriculum.

The underlying assumption is that pupils cannot understand what they cannot decode and that pupils knowing the meaning of individual words in the texts they read will not facilitate their capacity to decode. This is the rationale for SSP programmes such as Read Write Inc. (*RWInc.*; Miskin, 2020) insisting that pupils read every book in the scheme three times, first of all for accuracy, then fluency and finally for understanding. It is assumed that pupils being able to decode with 100% accuracy will be motivating, which may well be the case, but if so, will potentially reinforce for beginning readers the idea that understanding is not the initial priority and that reading is not related to comprehension. It also fails to acknowledge that understanding the meaning of what is read can facilitate pupils' phonic decoding skills through mispronunciation correction and Set for Variability (Dyson et al., 2017; Tunmer & Chapman, 2012).

The government reinforced the 16 core criteria in *The Reading Framework* (DfE, 2021b) that laid out its current policy on teaching early reading. It provided a detailed rationale for teaching SSP, highlighted the importance of pupils only reading texts that are consistent with their phonic knowledge and offered guidance on how to select an appropriate commercial synthetic phonics programme. It reinforces the DfE's view that SSP should be taught by commercial programmes and decodable texts through stating:

> Schools should invest in books that have been carefully structured in cumulative steps for children learning to read, so that they can decode every word as their knowledge of the alphabetic code increases.
>
> (p. 46)

The Reading Framework also implicitly implies that when stories are read to children they are not shown the text directly since they may encounter GPCs or common exception words that they have not yet learned, which they may find confusing. This reflects the contrast between the rhetoric surrounding how reading should be taught as expressed in the National Curriculum programmes of study and the reality as expressed in the core criteria.

The rhetoric is reflected in the Year 1 statutory guidance on comprehension (DfE, 2013b) which states that pupils should be taught to 'develop pleasure in reading, motivation to read, vocabulary and understanding by':

> Listening to and discussing a wide range of poems, stories and non-fiction at a level beyond that at which they can read independently.
>
> (p. 21)

The focus on developing reading for pleasure is on listening to texts beyond what they can read independently rather than emerging from what children read. When it comes to reading the emphasis is very different and centres on decoding through pupils only applying their knowledge of GPCs:

> Read aloud accurately books that are consistent with their developing phonic knowledge and that do not require them to use other strategies to work out words;
>
> Re-read these books to build up their fluency and confidence in word reading. (p. 20)

This again implies the purpose in reading is to demonstrate knowledge of taught GPCs through accurate decoding and then developing confidence through increased fluency. The guidance does not explicitly mention that pupils need to read with understanding and excludes children checking whether what they have decoded makes sense. This only happens when texts are read fluently after being re-read several times.

The DfE (2021b) also took the controversial step of announcing that it no longer regarded *Letter & Sounds* as a synthetic phonics programme even though it was used in 6 of the 12 schools identified in *Reading by Six* (Ofsted, 2010) as representing outstanding practice in teaching phonics. Gilbert then claimed that the challenge for all schools was to emulate the identified practices in the 12 schools. Nevertheless the DfE stated:

> The 2007 Letters & Sounds handbook, published under the previous government, has never been a full SSP programme. For a number of years, effective teaching using Letters & Sounds has relied on schools themselves building a programme around the handbook. We recognise, however, that for many schools, especially those who want or need to improve their practice, Letters & Sounds is not fit for purpose and does not provide the support, guidance, resources or training needed.

This represents a significant change in direction that does not appear to be based on an evaluation of the programme's impact on pupils' attainments, only an expression of the DfE's views about what constitutes a complete phonics programme. It is also a curious statement in acknowledging that *Letters & Sounds* has been taught successfully for a number of years due to teachers' efforts in developing a programme around the handbook but that these schools must nevertheless now consider altering their practice. Equally, it is assumed that a school's difficulties in implementing *Letters & Sounds* successfully was down to the programme rather than any other factors and will be overcome by using one of its validated programmes.

There is not a mandatory requirement that schools use one of its approved programmes but it is likely that schools will feel compelled to do so. The DfE (2022) advice to schools stated:

> There is no statutory requirement for schools to choose one of the SSP programmes on the validated list. Ofsted does not have a preferred programme or approach. However, validation status indicates that a programme has been self-assessed by its publisher and judged by a small panel with relevant expertise and that both consider it to meet all of the Department of Education (DfE) criteria for an effective systematic synthetic phonics programme.

The DfE's decisions, like so many that have influenced policy and practice since Plowden, are based on a view of how reading should be taught but without presenting any specific evidence, experimental or non-experimental, to justify its decisions, particularly its decisions to recommend SSP is taught alongside decodable texts, and withdraw *Letters & Sounds* as representing an acceptable methodology for teaching SSP. Given Gibb's (2017a) statement that practice should be evidence based, teachers could reasonably have expected the DfE to provide convincing evidence on the impact of the programmes that it regards as complete. Equally it is reasonable to expect that the DfE provide evidence in support of its decision to remove *Letters & Sounds* as a validated programme.

Castles et al. (2018) wrote about ending the reading wars and argued that it has persisted for two reasons. The first is that although there is a wealth of evidence demonstrating that phonics teaching works, there is not a comparable account to explain why phonics works. The second reason is that debates about teaching reading, and the evidence presented in a public forum, have rarely progressed beyond the role of phonics in learning to read. Their view seems to characterise the current position in England. Gibb, in his introduction to *The Reading Framework* (DfE, 2021b), claimed we have moved on from the reading wars as 'all schools now use systematic phonics to teach reading' (p. 4). The key question then is whether there is compelling evidence to support Gibb's claim that after over 50 years the reading wars has ended due to the developments in policy and practice in recent years, and in particular, the government's current drive to persuade schools to teach SSP through its validated programmes.

PART 2: REVIEW OF THE EVIDENCE ON WHICH POLICY AND PRACTICE ARE BASED—THE REALITY

Part 1 illustrated the extent to which teaching reading in England is highly politicised and that the curriculum has increasingly been influenced by successive governments over the last 35 years and their interpretation of the evidence on teaching reading. Although the agendas of government and the teaching profession should be highly compatible, the fact that governments are working within a narrow time frame to demonstrate that their policies are impacting positively on pupils' attainments means that their goals are invariably short term. Despite repeated suggestions that experimental research is undertaken to determine the impact of their recommendations, this has not yet taken place. Similarly governments rarely permit themselves, schools and teachers the time to systematically review the scientific evidence on reading and introduce changes at a pace that can maximise their impact.

Since the introduction of the NLS, almost 25 years ago, there has been a gradual shift to ensuring that early reading is taught primarily through systematic phonics, and over the last 10 years through decodable, phonically regular reading schemes. The debates about the respective roles of teaching a sight vocabulary and phonics prior to the NLS were equally intense with a clear desire to inform practice with evidence. Part 2 reviews the evidence on teaching reading since the late 1970s, particularly since the introduction of the NLS, and explores the extent to which it has informed government policy and practice, both in the past and present.

The focus on the extent to which policy and practice are informed by evidence is justified due to Nick Gibb's (2017a,b) frequent claim that educational practice should be based on the best available evidence. In a speech to the Education World Forum in January 2017 (Gibb, 2017a) he stressed how governments have a responsibility to inform policy with evidence, that the evidence must constrain the education experts and that the most expensive and longest investigation into raising standards in reading found that Direct Instruction programmes were the most effective:

Gibb spoke immediately after the results of the Progress in International Reading Literacy Study 2016 (PIRLS) study were released in December in 2017 (Mullis et al., 2017). He confirmed his commitment to evidence based practice and its impact in England:

> Decades of evidence from around the world—including the influential longitudinal study from Clackmannanshire in Scotland—pointed to systematic phonics as the most effective way to teach children to read.
>
> (Gibb, 2017b)

Gibb claimed the arguments of those opposing the use of phonics had always relied more heavily on emotion than evidence and he concluded his speech by saying:

> Slowly but surely, the education sector and the teaching profession are embracing evidence and raising academic standards for all.

However, evidence about the role of phonics in teaching reading began to emerge in England before the publication of the Plowden Report.

Evidence based practice after Plowden: The research on teaching phonics and balanced literacy

There are many parallels in the debates about teaching reading prior to the introduction of the NLS with those that have taken place more recently. They offer a cautionary note when examining the extent to which current government policy is evidence based and suggest that reaching a consensus on how best to teach reading is still some way from being realised. The question is whether we are about to see the pendulum swing away from teaching phonics back towards past practice, as recommended by Wyse and Bradbury (2022) following their critique of government policies, or an opportunity to build on what is currently known to move forward through addressing any limitations in current practice.

Morris (1958) discussed the relative effectiveness of different methods for teaching reading and noted that there was an emerging consensus in favour of a mixed method that taught regular words phonically (either through analytic or synthetic phonics) and irregular words through look-say. Morris (1959) (quoted in Murray, 1969) wrote:

> Most educationalists suggest that systematic phonics instruction should be delayed until children have acquired a good vocabulary and developed an interest in reading as a means of obtaining information and enjoyment.
>
> (p. 18)

Morris (1984) described her analysis of several databases of written English and identified 396 GPCs in formulating a programme to teach phonics knowns as Phonics 44. She argued that English orthography is highly patterned and that only a relatively small proportion of words diverge completely from conventional patterns. The vast majority of words can be recognised and spelled by applying the alphabetic principle of phoneme-grapheme correspondence and a knowledge of the statistical probability of sound-symbol relationships in English. Morris (1994) addressed the opposition to the scientific rationale for teaching phonics in an article called 'phonicsphobia' which examined its origins and the resistance to data driven arguments in favour of including a phonics component to teach reading.

Murray claimed that the content of the Ladybird Keyword reading scheme was based on scientific principles through teaching 250 whole words, largely in order of frequency, following an analysis of written English. The most frequent were taught in the early books before introducing letter sounds, diagraphs and consonant blends, gradually and systematically in the rest of the scheme which were taught through the synthetic method. Today the position is reversed in England. The following sections review the research underpinning government policies and practice and explain how this situation was reached.

The National Literacy Strategy: From onset-rimes and analytic phonics to GPCs and synthetic phonics

The NLS marked the first time that the government in England attempted to provide teachers with a detailed framework for teaching reading. It introduced greater rigour into how reading

was to be taught and was influenced by research that investigated the causes of pupils' literacy difficulties. Bradley and Bryant (1978, 1983) drew attention to the fact that pupils whose RAs were behind their CAs had difficulty in identifying the phonemes in a series of words that differed in their initial, middle or final phoneme. The research was ground breaking in highlighting the importance of students being sensitive to the phonological similarities and differences in groups of words. Goswami (1986, 1988, 1990a,b, 1991, 1993) and Goswami and Bryant (1990) extended this research and examined the development of phonological awareness in young children. They found that onset and rime awareness precedes phonemic awareness which in turn influenced the content of the word level work in the NLS.

Onset-rimes represent a large unit approach to decoding. The onset in a word is the initial consonant or consonants and the rime is the vowel followed by the remaining consonants. So the onset in <s-t-o-p> is /st/ and the rime is /Op/ and the onset in <d-r-ea-m-s> is /dr/ and the rime is /i:mz/. The onset and rime are pronounced separately and then blended together to yield the correct pronunciation. Teaching onsets-rimes is a feature of analytic phonics and was initially the method through which phonics was taught in the NLP. This can be contrasted with synthetic phonics which involves decoding through stating the phoneme represented by each grapheme, one at a time, and then blending the phonemes to pronounce the word. So for example, in 'stop' the four phonemes represented by the graphemes <s-t-o-p> are /s/t/O/p and the five phonemes in <d-r-ea-m-s> are /d/r/i:/m/z/. The phonemes in these words are pronounced separately before they are put together to lead to the correct pronunciation. Teaching phonics through GPCs is also known as teaching small units.

The word level work for the Reception year in the NLP required pupils to: (i) recognise critical features of words through their shape, length and common spelling patterns; (ii) discriminate 'onsets' from 'rimes' in spoken language (e.g., tip, sip, skip, flip, chip); (iii) recognise and explore patterns of rhyme (e.g., top, hop, mop, chop, shop) and alliteration (top, ten, try, twist) and (iv) extend these patterns by analogy, generating new and invented words in speech and in spelling. The Ofsted (1998) evaluation of the NLP commented:

> The word level work caused teachers the greatest difficulty, largely because many of them did not have a sufficient knowledge and understanding of what the phonic component should be. They often gave insufficient time to word level work and, consequently, it was frequently superficial and lacked systematic coverage. The teaching of word level work in the Project schools has remained unsatisfactory in an unacceptably high proportion of lessons. Given its crucial importance, this is a very serious and significant finding.
>
> (p. 9)

Despite the Ofsted evaluation, criticisms of the word level work and the continued use of teaching onset-rimes and analytic phonics (Duncan et al., 1997, 2000; Hulme et al., 1998; Muter et al., 1994; Stuart, 1999), the major difference between the NLP and NLS was that in the NLP the word level work was presented in the right-hand column of three, with sentence and text level work in the two columns to the left, but in the NLS it appeared in the left-hand column with text and sentence level work to the right. The word level work within the NLS was therefore the same as the NLP but with the addition of pupils learning to identify and write the initial and final phonemes in consonant-vowel-consonant words. Whereas the only explicit direct teaching in the NLP was high frequency words, there was direct teaching in the NLS to develop pupils' knowledge of GPCs. Students were to be taught 45 high frequency words directly by sight in Reception (aged 4 to 5) but other aspects of familiarising pupils with high frequency words reflected the whole language methodology used in the NLP (i.e. with reference to their shape and length).

The resulting controversy about which phonological skills to teach first, and the methods through which they were taught, could potentially have been averted had the decision to introduce the NLS been delayed to take account of the completed evaluation of the NLP. The NLP was introduced to schools in January 1997 and ran until March 1998. The Labour government was elected in May 1997 and decided in February 1998, 1 month before the NLP was due to finish, and 10 months before the evaluation of the NLP was published (Sainsbury et al., 1998), that all schools would implement the NLS from September 1998. Two months after the introduction of the NLS the government commissioned a report into the research supporting the NLS that was published in February 1999 (DfEE, 1999a), 5 months after the commencement of the NLS. The review contained a two-page summary of the research into teaching phonics and phonological awareness but failed to make any reference to the ongoing debates and arguments about the teaching of phonics.

It would appear that a decision had been made by the incoming Labour government to persist with the NLS irrespective of the evaluations of the NLP and emerging criticisms of teaching phonics through onset-rimes. Research had established, through correlational and experimental studies, that word recognition was the fundamental skill in learning to read (Adams, 1990; Share & Stanovich, 1995), and that children's progress in word recognition is closely related to their phonological skills (Ball & Blachman, 1991; Goswami & Bryant, 1990; Lundberg et al., 1988). However, the decision to focus on onset-rimes rather than GPCs may have been unduly influenced by the assumption that identifying the order in which skills develop naturally would lead directly to the most effective instructional sequence. What may be easiest to learn in the absence of explicit teaching may not be useful in the long term because it is of low utility and has low generalisability to unseen items. Alternatively, skills and knowledge that might be seen to be difficult to learn may be highly generalisable and so prove to be highly valuable. Although there is impressive evidence that onset-rime awareness precedes phoneme awareness, research into word recognition also casts doubt on the merits of teaching beginning readers onsets-rimes rather than GPCs.

Brown and Deavers (1999) found that whether pupils read unknown words through onset-rimes or GPCs depended on their reading competence. Less skilled readers, though able to make use of rime-based spelling-to-sound correspondences (reading ‘by analogy’), predominantly used simple grapheme–phoneme-level correspondences in reading isolated unfamiliar items. However, skilled readers were more likely to adopt an analogy strategy.

Deavers et al. (2000) examined the role of early reading instruction on the non-word reading strategies employed by beginning readers. One group of readers was taught through synthetic phonics, a second group through analytic phonics and a third control group received their conventional teaching that used neither synthetic or analytic phonics. They found that early reading instruction had a significant impact on the reading strategies pupils used to read non-words as well as phonically irregular words and that the synthetic phonics group were significantly more accurate when reading phonically irregular words than the analytic phonics and control groups. The synthetic phonics group were also advantaged at reading the nonwords although the differences did not reach significance. The fact that the synthetic phonics group appeared to be more accurate in their application of their chosen strategy is consistent with the view that small units may be particularly helpful for beginning readers (Carnine et al., 1997; Engelmann & Carnine, 1982).

The pupils in the Deavers et al. research were part of a larger study (Solity et al., 2000) that compared two approaches to teaching beginning readers, one known as the Early Reading Research (ERR) and the second the NLP. The ERR group were taught synthetic phonics but alongside teaching high frequency phonically regular (initially by sight until the relevant GPCs had been mastered) and irregular words and the use of real books. The NLP group was taught word level work in the way described earlier and had been included in the national trial of this programme.

The ERR group significantly outperformed the NLP group on seven out of eleven measures (standardised reading tests, phoneme synthesis and segmentation, phonically regular words) the exceptions being sight vocabulary and letters sounds (where ceiling effects were observed), spelling and rhyme awareness which was assessed through a forced choice rhyming task (e.g., which word rhymes with sun: wet, lap, run?). The NLP children were explicitly taught to read through a combination of rhyme, onset-rime and analogy and although their mean score on the rhyming task was actually higher than the ERR children, this difference was not statistically significant. It appears, therefore, that the ERR group's acquisition of rhyming skills occurred incidentally and was not dependent upon explicit teaching. In contrast the NLP group's training and superior performance on the rhyming task did not transfer to phoneme synthesis and segmentation, and reading phonically regular words.

Finally, observations from computational models of word reading also suggest that good generalisation to new items is only apparent when spelling-to-sound translation occurs at the level of individual graphemes and phonemes (e.g., Brown, 1997; Brown & Chater, 2004; Plaut et al., 1996). Overall, the evidence was clear that irrespective of the order in which skills were acquired developmentally, this should not be a barrier to pupils initially being taught at the level of GPCs.

The Ofsted evaluations of the NLS (Ofsted, 1999, 2000, 2001, 2002) and those from the DfEE (2000, 2001) and DFES (2003) reinforced the notion that being taught onset-rimes through analytic phonics was proving problematic. Although Ofsted found that the quality of delivery of the NLS improved from year to year, they repeatedly commented on the poor quality of phonics teaching indicating that it was not being taught well in approximately 25% of classrooms. Representatives of the NLS asserted that the problem in teaching phonics was one of delivery rather than because of what and how teachers were asked to teach and although Ofsted commented in their 2001 and 2002 annual reports that the quality of phonics teaching had improved, there was not a commensurate improvement in the English (reading and writing) Statutory Assessment Tests (SATs) results, the assessments that pupils take at the end of their primary education when aged 11.

The expected standard for this age group was Level 4 but between 1999 and 2002 the percentage of pupils failing to reach this criterion was 30% in 1999 and 25% in the 3 years between 2000 and 2002. The SATs results in 2002 contributed to the resignation of Estelle Morris, the Secretary of State for Education, as the Labour Government had indicated that anything <80% of pupils reading at Level 4 would indicate that standards were not improving as quickly as desired.

Ofsted held an invited seminar in 2003 to review the word level work within the NLS, in particular the teaching of phonics. There was a consensus among those invited to make presentations that systematic phonics represented the most effective way of teaching phonics. Some speakers (Solity, 2003; Stuart, 2003) explained why synthetic phonics was more effective than analytic phonics although at the time of the seminar, there was no published research to demonstrate the superiority of any one method of teaching phonics over another. However, one of the speakers, Rhona Johnston, had completed a study with a colleague Joyce Watson in Clackmannanshire, stating that in two experiments synthetic phonics was more effective than analytic phonics (Johnston & Watson, 2004).

Is synthetic more effective than analytic phonics?—The Clackmannanshire research

Although research conducted over a number of years was critical of the role of teaching reading through onset-rimes and analytic phonics, the significance of the Johnston and

Watson research was that it was set up to directly compare the impact of synthetic and analytic phonics during everyday teaching. Johnston and Watson reported two experiments and acknowledged that in Experiment 1, pupils in two analytic phonics groups were taught fewer letters than the synthetic phonics group. Johnston and Watson stated that 'Experiment 2 was carried out to control for these differences.' However, this statement is misleading since Experiment 2 was undertaken before Experiment 1 (Ellis, 2009). There were also major differences between the experimental groups undermining any claims that synthetic phonics was more effective than analytic phonics.

Experiment 2 was a 10 week, 19 session intervention provided in addition to pupils' usual reading programme. There were 92 participants in the study, divided into three groups: a control group of 29 participants who received no additional training to the usual classroom teaching; an analytic phonics group of 33 children and a synthetic phonics group of 30 children. The training offered to the analytic phonics group focused on: (i) pointing to illustrations of objects and stating their names; (ii) pointing to words starting with a target letter; (iii) pointing to pictures of named objects; (iv) pronouncing the initial phonemes when presented with pictures of objects; (v) repeating sounds stated by teachers when pointing to letters; and (vi) pointing to the printed word for objects. However, they were not taught to identify the phonemes representing graphemes in the final or medial positions in CVC words or to recall and state the phonemes represented by graphemes in the initial position or identify the rimes in taught words or undertaking any blending activities. This was in stark contrast to the synthetic phonics group who recalled phonemes represented by graphemes which were then blended to pronounce words.

Although both groups undertook activities with the same 114 words, the fact that the analytic phonics group were never required to recall the phonemes for graphemes or blend onsets with rimes, which are fundamental components in teaching analytic phonics, undermines any suggestion that the experiment demonstrated the benefits of synthetic phonics over analytic phonics. It does, however, demonstrate that Johnston and Watson's synthetic phonics intervention was more effective than the alternative, even though it could not reasonably be defined as teaching analytic phonics. Equally it was of value to have a treated rather than business as usual control group (i.e. untreated control group).

Johnston and Watson's research, and the claim that synthetic phonics is more effective than analytic phonics, has been the focus of much criticism (Torgerson et al., 2018; Wyse & Goswami, 2008; Wyse & Styles, 2007). Torgerson et al. (2006) conducted a meta-analysis and focused on experimental research that used randomised controlled trials to investigate the impact of phonics instruction on reading and spelling. They concluded: 'systematic phonics instruction within a broad literacy curriculum was found to have a statistically significant positive effect on reading accuracy' (p. 8) but not reading comprehension and that 'no statistically significant difference in effectiveness was found between synthetic phonics instruction and analytic phonics instruction' (p. 8).

Although no statistically significant difference was found between synthetic and analytic phonics, Torgerson et al. acknowledge that the evidence on this matter was weak and only included three RCTs, one of which was the Clackmannanshire study. The two other studies quoted by Torgerson et al. were Skailand (1971) and Torgesen et al. (2001). The Skailand study was reported at a conference, has not been published and is not accessible online. In the Torgesen et al. study neither intervention was comparable to the synthetic and analytic phonics interventions in Johnston and Watson's research or those included in the NLP or NLS. Furthermore Skailand had a sample size of 42 and taught two 15 minute sessions a week for 10 weeks (total instructional time was 5 hours). Torgesen et al. had a sample of 60 pupils, with a mean age of 9 years 9 months, and an intervention period of between 8 and 9 weeks where pupils were taught on a one-to-one basis. Even if both the Skailand and Torgesen et al. studies had compared acceptable versions of synthetic and analytic phonics

programmes, it would be difficult to draw conclusion on their relevance for teaching reading throughout Key Stage 1 (KS1) in England given the NLS covered three school years and taught students on a whole class basis.

A more valid conclusion from Torgerson et al.'s meta-analysis is that there have not yet been any studies that effectively compared synthetic with analytic phonics rather than there was no statistical difference between them. Stuart and Stainthorp (2016) reviewed the type of phonics teaching that is best taught to students but found that very few studies had been designed in such a way as to effectively compare interventions that taught synthetic and analytic phonics. However, two that have (Christensen & Bowey, 2005; Hatcher et al., 2004) were not included in the Torgerson et al. (2006) meta-analysis and concluded that teaching small units is more effective than larger units.

The results of the Clackmannanshire research, and the impact of synthetic phonics, were quoted frequently by those giving evidence to the Education and Skills Committee (2005). Nevertheless, the Committee in their summary did not claim that synthetic phonics was more effective than analytic phonics and instead repeated the advice of the Literacy Task Force to undertake experimental research to identify the most effective ways of teaching phonics to beginning readers:

> In view of the evidence from the Clackmannanshire study, as well as evidence from other schools where synthetic phonics programmes have been introduced, we recommend that the Government should undertake an immediate review of the NLS. This should determine whether the current prescriptions and recommendations are the best available methodology for the teaching of reading in primary schools. We therefore strongly urge the DfES to commission a large-scale comparative study, comparing the NLS with 'phonics fast and first' approaches.
>
> (p. 25)

The Rose Review: Teaching synthetic phonics and the potential role of real books

Following the report of the Education and Skills Committee, Sir Jim Rose conducted a review into how reading was being taught in the Early Years. The resulting report (Rose, 2006) drew on the findings from the Clackmannanshire research and noted that a number of contributors to the review expressed reservations about its methodology and conclusions. However, members of the review team visited Clackmannanshire and were impressed with how synthetic phonics was being taught. The review concluded:

> Despite uncertainties in research findings, the practice seen by the review shows that the systematic approach, which is generally understood as 'synthetic' phonics, offers the vast majority of young children the best and most direct route to becoming skilled readers and writers.
>
> (p. 4)

The review also noted the value of decodable books and giving pupils texts where they could apply their phonic knowledge and experience early success in reading to boost their confidence. However, the review recognised that the simple text in some recognised favourite children's books could fulfil much the same function as that of decodable books. The review suggested that 'real books' could be used in parallel or possibly in place of decodable texts which should 'not deny children access to favourite books and stories at any stage and particularly at the point when they need to read avidly to hone their skills, as the focus shifts from

learning to read to reading to learn' (p. 27). A major outcome of the review was the publication of *Letters & Sounds* (DfES, 2007), which is a SSP programme (Stainthorp, 2020) that featured prominently in *Reading by Six*, Ofsted's (2010) review of best practice in teaching reading.

Teaching synthetic phonics: Evidence from *Reading by Six*

Shortly after the election of the coalition government in 2010, *Reading by Six* (Ofsted, 2010) was published and reported on the practice of 12 outstanding schools in teaching reading during Ofsted inspections. It was different from any previous HMI or Ofsted reports on teaching reading, through drawing attention to the impact of a commercial SSP programme rather than highlighting the core instructional principles underpinning effective practice in the manner of the Rose Review. The schools were judged on their overall effectiveness, the quality of teaching and learning, and the quality of leadership and management. All 12 schools had above average results at Key Stages 1 and 2 and the schools represented a spectrum in terms of ethnicity, socio-economic backgrounds and geographical location. All the schools were visited by HMI during October 2010 where over 100 sessions in phonics, reading and writing were observed.

The schools can be divided into two groups of six. The phonics programme in one group of six only used *RWInc*. (Miskin, 2020), whereas the second groups all used *Letters & Sounds* (DfES, 2007) although not always exclusively: one school only used *Letters & Sounds*, three schools used *Letters & Sounds* alongside *Jolly Phonics* (Lloyd & Wernham, 2013) and two further schools used this combination alongside, in one instance *RWInc*. and in the other *Letterland* (Wendon, 2003). Head teachers had been in post on average for just over 6 years and the percentage of pupils achieving Level 2B (the expected standard in the Key Stage 1 assessments for 7 year olds) in 2009 in both groups was broadly comparable. In the six *RWInc*. schools, 83% achieved Level 2B whereas this figure was 85% in the *Letters & Sounds* group (the national average in 2009 was 72%). However, there was a marked difference in the percentages achieving Level 3 where the figure for the *RWInc*. schools was 29% compared to 43% for the *Letters & Sounds* schools.

It is interesting to speculate on how the 12 schools were selected from the hundreds that will have been inspected between 2006 and 2010, the period over which Ofsted graded these schools as outstanding. Head teachers had been in post for long enough to bring about change where necessary and presumably ensure a degree of stability. It is, therefore, equally interesting to look at what happened to the schools after the publication of *Reading by Six* between 2010 and 2021. One school has never been reinspected and only two schools have been reinspected and continued to be outstanding. Of the remaining nine schools, three have been reinspected and downgraded to good, four schools have closed and become academies, three of which have not been reinspected. The fourth was reinspected and found to have SATs results below the national average in 2019.

The final two of the 12 schools have been reinspected on several occasions. One school was required to improve on its second inspection in 2016 and was subsequently graded good after a third inspection in 2019. The remaining school has been inspected five times. It was downgraded to satisfactory, just a year after being included in *Reading by Six* and then graded as requiring improvement twice and inadequate once between 2014 and 2017. It was finally rated as good on its fifth inspection in February 2020. Achieving success in the long term is the goal of every government and every school, yet the experiences of these once outstanding schools illustrates that this is not an easy task.

There were several significant policy and practical consequences of *Reading by Six*. The first concerned the way reading was to be taught which was through SSP and phonically regular texts. The second was the government offering matched funding to schools to use

the SSP programmes that were listed in the Phonics Catalogue (ESPO, 2011). These programmes had been vetted by an anonymous panel of experts and were seen to meet the 11 core criteria that included teaching phonics first and fast. The third significant consequence was the development of the PSC in 2012 which was designed to assess pupils' progress in mastering phonics during the first 2 years of formal education (Reception and Year 1). The fourth consequence was a revised national curriculum (DfE, 2013a) that was introduced in September 2014 where the sections on teaching reading reflected that early reading was now to be taught through SSP.

Teaching common exception words through knowledge of grapheme-phoneme-correspondences

The DfE's (2021a) core criteria are very clear about how phonically irregular words (often referred to as common exception words [CEWs]) are to be taught. The criterion states:

> Ensure that children are taught to decode and spell common exception words (sometimes called 'tricky' words), appropriate to their level of progress in the programme.

The accompanying notes provide more information about teaching CEWs through phonics:

> They include correspondences that are unusual and those that will be taught later in the programme (e.g., 'said', 'me'). Programmes should teach children to read and then spell the most common exception words, noting the part of a word that makes it an exception word.

The notes were dismissive of the methods typically associated with teaching CEWs stating that students should not be encouraged to:

> Guess unknown words from clues such as pictures or context, rather than first applying phonic knowledge and skills. It should not include lists of high frequency words or any other words for children to learn as whole shapes 'by sight'.

Colenbrander et al. (2020) reviewed the arguments and evidence for different methods of teaching CEWs. They concluded that very few studies have compared different methodologies and so it would be difficult for the DfE to claim that its promoted methodology is based on evidence. However, the research from a number of different sources does provide guidance on how phonically irregular words can be taught. One method is to teach words by sight, which involves presenting children with written words, telling them how the words are pronounced (without reference to their shape or linking them to a picture), and asking them to repeat the pronunciation, initially when the teacher provides a model and then without a model.

Ultimately the goal when teaching reading is for all words to be read by sight fluently and with understanding, irrespective of whether they are phonically regular or irregular. In the early stages of teaching reading, the obvious disadvantage of being taught to read all words by sight is the vast number of words to be mastered and the methodology does not provide pupils with the capacity to generalise their knowledge readily from the words that can be read to unfamiliar words. According to Vousden et al. (2011) whole word reading is the most effective methodology for small vocabularies of approximately 50 words. Thereafter a small unit approach is most effective. This is the rationale for teaching phonics as it provides students with a generalisable strategy for reading unknown words when they are encountered.

Colenbrander et al. report there is no evidence that teaching irregular words at a sight level alongside knowledge of GPCs is damaging.

This is reflected in the findings of Shapiro and Solity (2016). They compared the impact of two synthetic phonics programmes on early reading. One programme was *Letters & Sounds*, which teaches that certain graphemes represent multiple phonemes, as well as teaching phonically irregular words through applying knowledge of GPCs. When teaching CEWs teachers: (i) state the phoneme represented by each grapheme; (ii) explain where a phoneme does not represent a grapheme that has been taught; (iii) give pupils practice in stating the phonemes for each grapheme and blending the phonemes to pronounce words; and (iv) refer to the word regularly throughout the day so that by the end of the day the children can read the word straight away without sounding it out. The second programme was the ERR which taught: (i) only the most frequent phoneme for each grapheme; and (ii) 100 high frequency words, of which 61 were phonically irregular and were read by sight. Teachers pointed and pronounce each word, before giving pupils daily practice in reading the high frequency words. They were introduced gradually, and only after previously introduced words had been mastered and read fluently.

The programmes were equally effective for typically developing readers. However, the ERR intervention gave pupils starting schools with poor phonological awareness, a small advantage in reading phonically irregular words by the end of the first year at school. Thus, combining sight word reading of a small number of high frequency, phonically regular words (read by sight until the relevant GPCs had been taught) and irregular words together, combined with only teaching one phoneme for each grapheme, was more effective than teaching phonically irregular words through sounding out and blending the GPCs alongside being taught multiple phonemes for graphemes.

The two interventions differed in a number of other ways but perhaps the most significant, as far as learning to read irregular words was concerned, was that the ERR group were taught to read through real books whereas the *Letters & Sounds* group applied their knowledge of GPCs to phonically decodable texts. Real books provided pupils with frequent opportunities to read phonically irregular words in continuous prose which they also practised reading out of context to high fluency levels. This combined methodology reflected the two conditions researched by Stuart et al. (2000) who found that acquiring a sight vocabulary from repeated reading of text alone, is easier for the child with good phonic knowledge, but the task itself is more difficult than mastering a sight vocabulary from exposure to words in isolation presented on flash cards.

The government's approach to teaching CEWs with reference to their GPCs is conceptually flawed. When teaching phonics pupils are taught diagraphs where two letters represent a single phoneme (e.g., <sh> /ʃ/) that are different from the phonemes represented by the individual letters (<s> /s/; <h> /h/). Although the phoneme /ʃ/ in <shop> has a consistent pronunciation there are some diagraphs that can represent three (e.g., <ch> / tʃ/—chip; /k/—school; /ʃ/—chef) or four phonemes (<ou> /aʊ/—out; /u:/—soup; /c/—could; / əʊ/—soul) and so do not offer the consistency and predictability anticipated when teaching phonics. How is this different from telling students that in some instances they will not have been taught all the GPCs to decode a word phonically, so they state the whole word without blending the phonemes represented by graphemes? For example, when <b> and <e> appear together the word is pronounced /bi:/ and will always have the same pronunciation, there are no inconsistencies. Similarly when <s>, <a>, <i> and <d> appear together the word is always pronounced /s/, /e/, /d/. In both instances pupils are learning that when certain letters appear together they result in a pronunciation that is different from the pronunciation that would be the outcome from blending the phonemes represented by the individual letters.

It might be suggested that although the grapheme <ch> may represent several phonemes, there is still a single mapping between grapheme and phoneme. However, <qu>,

<ew> and <u_e> are taught as graphemes and represent two phonemes (/kw/ and /ju:/) as do the single letter graphemes <x> and <u> that represent the phonemes /k/s/ and /j/u:/ (the graphemes <ew>, <u_e> also represent /u:/ and <u> also represents /ʊ/ and /ʌ/). Ultimately it is an empirical question to establish whether teaching a small number of CEWs by sight is problematic and interferes with students' progress. At the moment however, there is no experimental evidence to support the government's way of teaching CEWs.

Teaching synthetic phonics through phonically decodable texts

As well as promoting SSP, the government has increasingly insisted that phonics is taught through phonically decodable texts. The National Curriculum (DfE, 2013a) states that pupils:

> Read aloud accurately books that are consistent with their developing phonic knowledge and that do not require them to use other strategies to work out words.
>
> (p. 20)

Similarly Ofsted's Education Inspection Framework (2019) stresses the importance of pupils reading decodable texts and claims it will develop their confidence and enjoyment of reading:

> A rigorous approach to the teaching of reading develops learners' confidence and enjoyment in reading. At the early stages of learning to read, reading materials are closely matched to learners' phonics knowledge.
>
> (p. 10)

The use of reading schemes is based on the assumption that students must be able to read complete books, from beginning to end, with 100% accuracy, even if they do not understand what they are reading. In order to do so the text must be simplified and only include, as far as possible, phonically regular words. Before being given a scheme book students are taught all the GPCs as well as the CEWs so every word can be read without error. Children progress through a scheme, one book at a time, before becoming a 'free reader' and going on to read books of their own choice. The arguments in favour of phonically decodable schemes are persuasive and have an intuitive appeal. They create the illusion of progress and imply that learning to read is a relatively straightforward matter of students mastering phonic skills and working through a scheme in a gradual and systematic manner.

However, there is evidence to suggest that decodable texts can have a negative impact on students' attitudes to reading that contrasts with the rhetoric. Levy (2009) found that children view schemes as a vehicle to teach the mechanics of decoding whereas Carter (2020) found that students did not regard phonics as a skill that supported their reading. Phonics was seen to be a separate subject with specific activities that were dislocated from their application to reading and were not talked about as being reading. These findings reflect Ofsted's description of reading in two reports on reading for pleasure (2004, 2012) where a distinction has been created between three aspects of reading—decoding, meaning and pleasure—which are articulated as three independent processes that are developed separately and sequentially.

Despite the rationale for decodable texts, and the underlying assumptions underpinning their use, there is no evidence to date to show that they have a demonstrable impact on students' progress in learning to read. The National Reading Panel (NRP & NICHD, 2000) reported on the use of decodable texts and found that, whereas some synthetic phonics

programme used them (in the way that would also be seen as consistent with current DfE requirements), others made little or no use of decodable texts. The panel concluded the section on decodable texts as follows:

> Surprisingly, very little research has attempted to determine whether the use of decodable books in systematic phonics programs has any influence on the progress that some or all children make in learning to read.
>
> (2–137)

Since the NRP report, only two experimental studies have compared phonically decodable texts with non-decodable texts. Jenkins et al. (2004) investigated the impact of decodable texts with at risk beginning readers. Participants were divided into three groups. The first read storybooks that were consistent with the GPCs taught in the programme that was being followed. The second followed the same phonics programme but read stories where the phonic content was not controlled to reflect the skills taught in the programme. The decodability for the stories read by the first groups was 85% and for the second group 11%. A third group did not receive any teaching in phonics or story reading. The two experimental groups were taught 4 days a week for 25 weeks which is equivalent to two terms in an academic year in England. The two groups significantly out-performed the control group on measures of decoding, word reading, passage reading and comprehension but did not differ from each other on any of the post-test measures.

In a more recent small-scale study, Price-Mohr and Price (2020) investigated the impact of low or high phonically decodable texts with 4–5-year-old beginning readers. Twelve instructional reading books were created for each condition that introduced the same number of words and contained identical illustrations. Similarly, parallel sets of games were developed for each condition. There was a significant effect for passage reading comprehension indicating that a higher percentage of non-decodable words in predictable text can lead to better outcomes in reading comprehension. The results were close to significance on the measure of early word reading which suggested that a higher percentage of non-decodable words in instructional texts can lead to better outcomes on word recognition. The results of the study need to be treated with a degree of caution due to the small sample size but raise the possibility that SSP can be taught through texts which are not written to contain high levels of phonically regular words.

In England there have been a number of studies that have compared reading interventions that have taught synthetic phonics through real books with those that have taught analytic and synthetic phonics using a reading scheme. A series of quasi-experimental studies by Shapiro and Solity (2008), Solity et al. (1999, 2000) and Solity and Shapiro (2008) compared: (i) a framework for teaching synthetic phonics through real books (the SP-RB intervention) with an untreated control group over 2 years with a one-year follow-up; (ii) the SP-RB intervention with the NLP over one year; and (iii) the SP-RB intervention with the NLS over 3 years with a one-year follow-up. The groups taught through the SP-RB intervention significantly outperformed the three interventions with which it was compared, both when the interventions ended but also after a one-year follow-up when all groups were taught through the NLS. The SP-RB intervention impacted positively on the attainments of higher, middle and lower achieving pupils and led to a decrease in the incidence of difficulties in the experimental group from 20% in the comparison groups to <5% in the SP-RB groups. In the study comparing the SP-RB intervention with an untreated control group, the lower achievers in the SP-RB group performed to a comparable level to the middle achievers in the untreated control group.

The SP-RB experimental intervention differed from the untreated controls, NLP and NLS comparison groups in a number of ways, making it difficult to say with confidence that it was

the presence or absence of real books that contributed to the different outcomes in these studies. The real books groups were monitored regularly and given detailed feedback on treatment fidelity. The untreated controls taught reading with less rigour than the experimental group in contrast to the NLP and NLS groups who followed the specified curricula as required. However, the earlier discussion has highlighted the limitations in the NLP and NLS, in particular the extent to which they were teaching synthetic or analytic phonics, and so the observed differences may be attributable to a range of factors. Significantly though, teaching synthetic phonics through real books did not disadvantage the experimental groups.

Petscher et al. (2020) reviewed what they identified as high quality experimental and quasi experimental research studies on teaching reading as well as observational studies, to identify aspects of practice that are supported by what they defined as compelling evidence, promising evidence or a lack of compelling evidence. They found that there was an absence of compelling evidence for the use of decodable texts and that in the USA there was no evidence to support the use of any specific classroom reading schemes. It should also be noted that Petscher et al. found no evidence for the impact of multisensory teaching approaches on teaching phonics, indicating that such methods are no more effective than comparable programmes that do not include a recognisable multisensory component. This is of considerable significance since the Rose Review and DfE (2021a) criteria for validating SSP programmes both recommend including multisensory methods.

There are two further issues to consider in debates about the role of reading schemes when teaching early reading. Solity and Vousden (2009) compared the content of two versions of the children's fairy tale, *The Three Billy Goats Gruff*. In order to read the scheme version of the book (*Billy the Kid*, Munton, 2008) students would have been taught 12 phonically irregular words as well as to blend phonically regular words with up to four graphemes (VC, VCC, CVCC, CCVC) and one diagraph, <sh>. Sixty-nine per cent of the word tokens were accounted for by phonically regular words with phonically irregular words account for the remaining 31% of word tokens. A student who could read *Billy the Kid* with 100% accuracy would only be able to read 39% of the real book version of the story, *The Three Billy Goats Gruff* (Sharratt & Tucker, 2005).

The paradox, however, is that students would have more opportunities to practise their phonic decoding skills in *The Three Billy Goats Gruff*. *Billy the Kid* contains 138 phonically regular words whereas *The Three Billy Goats Gruff* contains 160 phonically regular words. Pupils would be able to read the scheme book independently whereas the real book would have to be read through the paired reading strategy (Topping, 2014) where students read the words for which they have the appropriate skills and the listener reads the rest. Paired-reading provides a methodology through which pupils can demonstrate their knowledge of GPCs with 100% success that does not require them to read decodable texts. Requiring that SSP is taught through decodable texts rather than real books fails to appreciate the difference between form and function. In the above example the function of both the phonically decodable and real book versions of *The Three Billy Goats Gruff* is to provide practice in decoding phonically regular words through taught GPCs. There is no convincing evidence that the form has to be phonically regular texts or that pupils are disadvantaged by the use of real books.

The second consideration is that real books provide the opportunity to develop pupils' vocabulary knowledge as well as having the capacity to facilitate their phonic decoding skills (Solity & Vousden, 2009). For example, the only word used to describe the way in which the various characters speak to each other in *Billy the Kid* is 'said', which is repeated on 11 occasions. In the real book 'said' is also repeated 11 times; however, eight different words and phrases are also used (e.g., 'shouted out', 'grunted', 'said with a grin', 'replied', 'roared', snapped', 'a cry of', 'spluttered') to describe how the different characters speak to

each other. As Castles et al. (2018) explain, exposing pupils to complex words with nuanced meanings is critical in enabling them to develop their reading skills beyond phonics.

The consequences of teaching reading through decodable texts

When children are taught to read with a reading scheme it is recognised that pupils need to hear a broader and more varied vocabulary than is available in the scheme. The revised National Curriculum (2013) implicitly accepts that the purpose of reading schemes is not to foster a love of reading, but to teach pupils to decode, when it states that pupils will need to 'hear, share and discuss a wide range of high quality books to develop a love of reading and broaden their vocabulary' (p. 18). The National Curriculum and many SSP programmes suggest that stories are read to students and imply that they do not simultaneously follow the text as it will contain GPCs that they have not yet been taught or are irregular and may cause confusion. However, it appears that there is a cost in doing so.

Colenbrander et al. (2019) conducted a study to determine when, why and how the presence of a word's written form during instruction aids vocabulary learning. This is a process known as orthographic facilitation, where pupils' knowledge of spoken words and their meaning is mapped on to the written form of a word. They found strong evidence that the presence of a word's written form leads to improved learning of its spelling and spoken form and there was also some evidence that it may lead to better learning of a word's meaning. Valentini et al. (2018) investigated children's incidental word learning from stories. They compared three conditions: (i) listening to stories; (ii) reading the stories; and (iii) a combined condition. Phonological learning occurred in all three groups. In contrast, only the groups reading stories demonstrated an enhanced acquisition of orthographic information, indicating that hearing stories alone was not sufficient to establish orthographic learning. Semantic learning was greater in the combined condition than for the groups that either only heard the story or only read the story.

Vocabulary knowledge facilitates phonic decoding

There is also evidence that vocabulary knowledge facilitates pupils' phonic decoding skills. Set for Variability (SfV) (Tunmer & Chapman, 2012) refers to students determining the correct pronunciation of approximations to spoken English. This is the capacity to derive the correct pronunciation of a phonically irregular word that has been decoded accurately (e.g., *stomach* pronounced as 'stow-match') but led to an incorrect pronunciation. Crucially, Tunmer and Chapman found that vocabulary knowledge was one of the factors that contributed to pupils correcting mispronunciations of accurately decoded, phonically irregular words.

Dyson et al. (2017), conducted a training study to investigate whether pupils could be taught to correct errors in the pronunciation of irregular words. Children were randomly assigned to an intervention or control group. The intervention group participated in a four-week programme in which they were taught to correct mispronunciations of spoken words as well as being taught the meanings of those words. Children in the control group received no additional teaching. The intervention group made significant gains in their ability to correct mispronunciations and to read and define the taught words; these gains also generalised to a comparable set of untaught control words. Children can be taught to correct errors in the pronunciation of irregular words, and this may produce generalised effects on learning to read. Austin et al. (2021) also found that knowledge of word meaning alongside explicit instruction in decoding, facilitates accurate and fluent word reading.

The research on orthographic facilitation, SfV and mispronunciation correction suggests that pupils will be missing the chance to master crucial skills in learning to read when taught SSP through phonically regular texts. Pupils will not be able to expand their knowledge of written English through not seeing text when hearing stories or will not have the opportunity to correct mispronunciations when reading as they are required to decode texts with 100% accuracy. The assumption when using a reading scheme is that decoding takes priority over understanding and fails to acknowledge how understanding a word's meaning may facilitate pupils' phonic decoding. The implication instructionally is that there is potentially a balance to be struck between the time devoted to mastering GPCs, particularly those that are low frequency in children's literature, and extending pupils' vocabulary knowledge so that they can maximise the impact of their phonic knowledge through correcting mispronunciations.

Do government validated synthetic phonics programmes work?

Shanahan (2020) suggests that research in reading has generally focused on cognitive mechanisms and neural processes involved in reading together with computational models of learning to read which, although invaluable, have not always yielded effective instructional strategies that can be transferred to the classroom. He suggests that there is a need for research exploring how best to teach reading and how to use the results to inform public policy and instructional practices. The direction for future research should be towards conducting well-designed, systematic studies of reading interventions to build up a picture of what is most effective and most likely to promote equity in pupils' experiences in learning to read.

Seidenberg et al. (2020) state that many educators have rejected the premise that their policies and practices are major factors in students' poor progress in learning to read. They explore similar themes to Shanahan in recognising that there have been many studies into looking at specific components in the reading process that have not been translated into effective classroom interventions. They define this as translational research and see the current status of research in phonics as a translational issue. Despite the research in recent years, there are many unanswered questions about how to teach phonics effectively. For example, which GPCs should be taught, is there an optimal number to teach and should students be taught multiple phonemes for a single grapheme? In the absence of sufficient translational research, Seidenberg et al. make it possible for the authors of almost any programme to claim that they are effective.

The DfE (2021a) has never made it clear how they arrived at their 16 core criteria that publishers are required to meet to become a validated SSP programme. Since they were published, 19 programmes have now been approved and the inevitable question is whether there is any research to demonstrate that they impact positively on pupils reading. Evidence of impact was not one of the 16 core criteria just as it was not specified in the DfE (2010) list of 11 core criteria. However, not only was there no requirement for authors and publishers to present any research on the impact of their programmes, there was no expectation that they would undertake any research in the future to address this issue.

The research discussed so far indicates that there is little evidence to support the government's interpretation of how best to teach SSP. There is strong evidence that phonics is best taught through synthetic phonics and this has been incorporated into the practice of the majority of schools in England. However, the 19 validated programmes differ enormously in the way that they teach synthetic phonics in terms of: (i) the number of GPCs taught; (ii) the number of graphemes that represent multiple mappings; (iii) the rate at which new material is introduced; (iv) the assessment procedures through which student progress is monitored; (v) the quality and number of books in the reading scheme; (vi) the number and methodology though which phonically irregular words are taught; (vii) whether pupils are taught synthesis

and segmentation as phonological skills in the absence of print as pre-skills for teaching phonics and spelling; (viii) the role of real books when teaching reading and whether pupils can see the text when they are read to by adults; (ix) how spelling is taught; (x) the nature of the resources and materials used to teach reading; (xi) how the programme is developed to meet the needs of the lower achieving pupils; (xii) how fluency is developed; and (xiii) how the programmes develop a love of reading in children.

Despite there being several calls over the years for an experimental investigation into the phonic programmes being recommended by government, none have been undertaken to date with one exception. The Education Endowment Foundation (EEF) funded the Queen's University Belfast (QUB) to run a study in 120 schools (60 experimental and 60 control) involving an estimated 4400 pupils between September 2016 and July 2018 to evaluate the impact of *RWInc*. at the end of Key Stage 1. The American Institutes for Research (AIR) was funded to analyse the data and write a final report on the study, which was due to be submitted to the EEF in September 2018 and published in February 2019. The publication of this report was delayed to March 2020 and then again until Autumn 2021, after it was revealed that the EEF had not received the AIR report until February 2021. Then in August 2021 the EEF stated on its website that the final report would not now be published until 2023, 5 years after the evaluation ended and 4 years after it was first due to be published.

The EEF took the decision to publish the QUB/AIR report alongside the findings of a second evaluation of *RWInc*. by the NFER that was due in the Summer of 2021. However, the evaluation has been delayed until 2023 due to Covid-19. The August 2021 announcement from the EEF stated that it 'made a commitment to publish these two studies together, to provide schools with the most useful and consistent messages around the findings' but failed to explain why such a commitment had been made or to whom, or the nature of the consistent message. The NFER evaluation of *RWInc*. was funded as part of the Teaching and Leadership Innovation Fund (TLIF) with the EEF funding the NFER for an undisclosed amount.

It is reasonable to ask questions about the design and methodology adopted in these studies since *RWInc*. is used so widely and to question the delay in reporting the results. Approximately 25% of the primary schools in England now teach reading through *RWInc*. and it has been adopted by 24 out of 34 English Hubs (DfE, 2018) that were set up by the government to promote good practice in schools to raise standards of literacy. It was featured prominently in *Reading by Six*, became a validated programme in 2011 and has been validated again in 2021, yet there is no evidence from either RCTs or quasi-experimental studies on its efficacy.

The TLIF/NFER evaluation of *RWInc*. involving four cohorts of schools, was announced in April 2019, 2 months after the AIR report was due but 2 years before the report was finally available. In the past the EEF has usually only funded a second evaluation of an intervention in order to test a scalable model, under everyday conditions, in a larger number of schools than an initial evaluation and only when that initial study is construed as 'promising.' In the absence of any evidence from the QUB/AIR evaluation of *RWInc*. it is unclear why the EEF decided that there was a need to complement this study with the TLIF/NFER evaluation.

The TLIF funded a more intensive training model rather than testing *RWInc*. under everyday conditions. Teachers were trained over 2 days in the QUB/AIR evaluation and then received just 4 days of further training on implementing *RWInc*., which included devoting one day specifically to the slowest progress pupils. This can be contrasted with 18 full days further training in the TLIF/NFER evaluation, at a cost of £14,700 per school, plus 19 days of supply cover for reading leaders and a regular focus over 18 school visits on the 20% of slowest progress readers to establish a weekly coaching cycle. It is difficult to see how a consistent message can emerge about the impact of *RWInc*. from these two evaluations

given the very different post-training support provided in terms of frequency, goals, personnel and funding.

However, the major concern is the nature of the research design and the measures taken to evaluate *RWInc*. No baseline measures were taken in either study and the NFER evaluation was announced two terms after Cohort 1 in the TLIF project began implementing *RWInc*. and one term after the schools in Cohorts 2–4 implemented *RWInc*. The schools in the TLIF study were selected by an organisation closely associated with *RWInc.'s* authorship (Ruth Miskin Training) and the primary assessment measure is the Year 1 PSC. The QUB/AIR study evaluated the impact of *RWInc*. through a different primary assessment measure, the New Group Reading Test (GL Assessment, 2014).

There are two problems with the use of the PSC to evaluate the impact of *RWInc*. The first is that Ruth Miskin, the author of *RWInc*., was one of the four leading phonics experts who developed the content of the PSC (DfE, 2011), which was based in part, on what is taught in *RWInc*. The second is that *RWInc*. has developed materials to help schools prepare for the PSC (*RWInc*.: Phonics Screening Check Tips—https://bit.ly/3kbNVbv**)** that includes information for parents on how they can provide pupils with extra practice. As the comparison schools were not using *RWInc*. they would have been at a distinct disadvantage.

The EEF's failure to publish the QUB/AIR evaluation of *RWInc*. in 2021 is disappointing since the study could have potentially provided some initial evidence on the impact of the government's drive to ensure every school in England is taught SSP through one of its validated programmes. As already stated, *RWInc*. is one of the most widely used government validated SSP programmes in England and yet there will be no experimental data on its impact for at least another 2 years. Equally the EEF has yet to announce its revised research protocol for the TLIF/NFER evaluation since the original participants who should have been assessed during Key Stage 1 (aged 6 and 7) are now in Key Stage 2 (aged 8 and 9).

Teaching validated systematic synthetic phonics programmes to read through direct instruction

Gibb (2017a) acknowledged the impact of Direct Instruction (DI) (uppercase D and I) (Carnine et al., 1997; Engelmann & Carnine, 1982) in Project Follow Through (Rhine, 1981). DI has a number of distinctive features, in particular, the principles underpinning curriculum content and it is probably one of the most extensively researched synthetic phonics programmes available. Stockard et al. (2018) report a meta-analysis of a half-a-century of research in DI where 328 studies met the inclusion criteria for the analysis that had a total of 3999 calculated effects. However, none of the 19 programmes validated by the DfE reflect any DI principles in the design of their curricula for teaching SSP.

A possible response to this assertion may be that the DfE (2021a,b) make reference to direct teaching and direct instruction in relation to the teaching methodology. However, this would be to confuse DI with direct instruction (di) (lowercase d and i) (Kirschner et al., 2006; Rosenshine, 2010, 2012) which also focuses on making instruction explicit, frequently through scripted lessons where teachers model new tasks and specify learning outcomes in behavioural terms. However, despite these superficial similarities the two are fundamentally different due to the focus in DI on curriculum design, faultless instruction, teaching generalisable skills and the assumptions and principles underpinning the teaching and learning process. The more general issue is that the core criteria through which the 19 programmes have been validated focus primarily on whether SSP is taught through decodable texts and pay little attention to any broader instructional principles that relate to curriculum design or teaching methodology.

Have standards in reading improved?

In December 2017 the results of the 2016 PIRLS (Mullis et al., 2017) were published and appeared to justify the decision to teach beginning readers through synthetic phonics. The Year 5 pupils (aged 9 and 10) participating in PIRLS was the first cohort to have been assessed on the PSC in 2012 and so had been taught to read exclusively through SSP. The results were viewed by Gibb (2017b) as vindication of the government's policy:

> When we came into office in 2010 one of the first things we did was to strengthen the National Curriculum, explicitly requiring schools to teach reading using phonics. Most controversial of all, we introduced a test for all six-year-olds, called the Phonics Screening Check. Today, we received the first set of international evidence that confirms that our approach is working.
>
> (Gibb, 2017b)

Data from two sources will be examined to explore whether standards in reading have improved since the introduction of widespread teaching of SSP. The first is the results from the PSC and the second is outcomes of the PIRLS assessments.

Has the phonics screening check led to improved standards?

The PSC (DfE, 2012) was introduced in 2012 to assess pupils' phonic knowledge and blending skills through decoding 20 real words and 20 pseudo words. The aim was to reinforce the teaching of SSP through an annual test, taken at the end of Year 1 (pupils' second full year in school when aged 6). Pupils who failed to reach the pass mark were required to take the PSC again at the end of Year 2. The PSC listed 85 GPCs that could be assessed and stated: (i) the inclusion of a particular grapheme would not necessarily be in proportion to its frequency in words and (ii) over time the PSC will include all single letters of the alphabet and all GPCs listed in the assessment framework (DfE, 2011, 2012).

Table 1 shows the results of the PSC over 8 years between 2012 and 2019. The PSC was not administered in 2020 or 2021 due to the pandemic. Instead schools were required to administer the PSC from 2017 or 2018 or 2019 to the Year 2 pupils in the autumn term. The results indicate that between 2012 and 2016 there was a steady increase in the percentage passing the PSC from 58% to 81% in Year 1 and a levelling off between 2016 and 2019. The results also show that by the end of Year 2 the percentage passing the PSC increased from the first cohort in 2012 (85%) to the second cohort in 2013 (89%) but has then been stable with between 90% and 92% passing between 2014 and 2019.

Table 1 also shows the KS1 SATs results between 2012 and 2019. It is difficult to make comparisons over the 8 years since new KS1 assessments were introduced in 2016 to assess the new national curriculum and the expected standards were raised. The DfE therefore stated that the results from 2016 onwards were not comparable to previous years. The DfE made changes to the teacher assessment framework in 2018–2019, meaning that the 2019 results were not directly comparable to previous years. Nevertheless what can be concluded is that between 2016 and 2019 one in four pupils failed to meet government expectations in reading at the end of KS1 despite the fact that over 90% had passed the PSC by the end of Year 2.

The fact that between 2016 and 2019 only one in ten pupils are failing to meet expectations on the PSC by the end of Year 2 but one in four are failing to meet expectations in the reading SATs is a cause for concern. One possible explanation is that the PSC is not an adequate assessment of pupils' phonic knowledge and decoding skills. Another is that

TABLE 1 Percentage of pupils passing the PSC between 2012 and 2019 and the percentage reaching the expected level in the KS1 SATs between 2012 and 2019

	Year in which PSC and SATs were taken							
	2012	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Percentage passing in Year 1	58	69	74	77	81	81	82	82
Percentage passing in Year 2		85	89	90	91	92	92	91
Percentage reading Level 2B	76	79	81	82				
Percentage at expected level					74	76	75	75

competence in phonics is not sufficient to ensure success in reading where a broader range of skills is assessed. It could be argued that the focus on phonics and decoding is having a detrimental impact on the wider reading skills of approximately 25% of 7 year olds.

Darnell et al. (2017) analysed the content of the PSC over the first 3 years between 2012 and 2014 and found that a small number of GPCs accounted for the majority of GPCs assessed. Their analysis showed that just 15 GPCs accounted for 67% of all GPC occurrences, with 27 of the 85 specified GPCs (31.8%) not appearing at all. Where a grapheme represented more than one phoneme (e.g., <ch> represents three phonemes /tʃ/; /k/; /ʃ/) the most frequently occurring pronunciation was assessed in the majority of cases with vocabulary knowledge being required to determine the correct pronunciation where multiple pronunciations were possible. The GPCs assessed, therefore, do not reflect the full range of 85 GPCs that it is expected will be taught within a systematic synthetic phonics approach, contrary to the claim made in the PSC assessment framework (DfE, 2011, 2012). Furthermore, children's ability to decode real words is dependent on their vocabulary knowledge, not just their phonic skills (Tunmer & Chapman, 2012).

A further analysis of the PSC over 8 years between 2012 and 2019 (Darnell & Solity, in preparation) reveals a similar pattern and shows that 1213 GPCs have been included in the 160 real words and 160 pseudo words. Fifteen GPCs accounted for 67% of all the GPCs included in the PSC. Thirteen of the 85 GPCs have never been included with a further nine GPCs appearing only once and 11 GPCs have been assessed that were not specified in the assessment framework. There are 18 graphemes that have a specified multiple mapping representing a total of 39 GPCs. The higher frequency phoneme has been assessed on 94% of the occasions a grapheme representing multiple phonemes has appeared. For the grapheme <ch>, **ch** in **ch**op has appeared 15 times but the other two representations of **ch** (ch in **ch**ef, ch in e**ch**o) have not appeared at all. Thus, when pupils pass the PSC their knowledge of GPCs is only being partially assessed on a small number of the GPCs that will have been taught.

The PSC was designed to assess pupils' knowledge of the 85 GPCs that it was thought they needed to master. However, achieving the pass mark will only indicate that pupils have learned a relatively small proportion of those GPCs that could be assessed. If it is claimed that the results of the PSC are a vindication of the government's decision to promote SSP, the data indicate that pupils only need to master a small number of GPCs for there to be a marked improvement in their reading. An alternative explanation is that a number of pupils who achieve the pass mark of 32 on the PSC have an incomplete knowledge of the 85 GPCs that it is anticipated that they have been taught. This may explain why over 25% of pupils' transferring to secondary schools at age 11 have failed to meet the government standards

in literacy in the KS2 assessment (see next section). Equally the data suggest that the focus on teaching SSP means that pupils are not being taught a broader range of essential literacy skills that are required to become competent, fluent readers.

It is interesting to speculate whether the results of the PSC would be different if: (i) all 85 GPCs that were included in the assessment framework had been assessed a comparable number of times; (ii) lower frequency GPCs were assessed when a grapheme represents more than one phoneme and (iii) pupils were not prompted with visual cues when pseudo-words are presented to help distinguish them from real words. The results of the PSC would be more meaningful if baseline measures had been taken of students' phonic and vocabulary knowledge on school entry, and at the end of Reception (aged 5), so that their progress over time could be assessed.

Progress in International Reading Literacy Study

The decision to teach reading through SSP and phonically regular reading schemes appears to have been vindicated with the publication of the results of the PIRLS assessments in 2017 (McGrane et al., 2017; Mullis et al., 2017) when England performed extremely well, moving from 10th place in 2011 to 8th in 2016. The significance of these results was that the Year 5 pupils (aged 10) who took part in PIRLS 2016 were the first cohort to have been taught exclusively through SSP throughout their school careers and had been assessed on the first PSC in 2012. By the end of Year 1, 58% of the cohort had passed the PSC and by the end of Year 2, 85% of the cohort had reached the pass mark of 32.

Double et al. (2019) and Stainthorp (2020) also offered positive interpretations of the PIRLS data. Stainthorp welcomed England's improvement compared to 2006 and 2011, noted that the attainment gap had been significantly reduced and that students' improved performance in phonics as evidenced by their performance on the PSC, meant that teachers and pupils were able to focus their attention on enjoying and understanding texts. Double et al. (2019) compared the performance of pupils who passed the PSC in Year 1 with those who passed in Year 2, on their performance on the comprehension components of the 2016 PIRLS assessment. There are, however, a number of factors that need to be considered before it can be stated confidently that England's performance in the 2016 PIRLS assessment is attributable to the focus on teaching SSP since 2012.

England improved by seven points (552 to 559) in the PIRLS assessments between 2011 and 2016. However, 20 countries (out of 41 that had PIRLS scores in 2011 and 2016) made greater progress than England. Ireland, had the same points score as England in 2011, but more than doubled England's improvement (552 to 567) as they moved from tenth equal in 2011 to fourth in 2016. Significantly, Ireland does not have a curriculum in the early years based on SSP and recommends teaching strategies that are anathema to those in England recommending SSP, including using picture and context cues to read unfamiliar words. The curriculum in Northern Ireland, who finished ahead of England in both 2011 and 2016, also encourages the use of picture cues to facilitate children's reading. In PIRLS 2016, 6% of the schools in the English cohort were from the independent sector, who are not required to teach reading through SSP or to take the PSC. If their results are excluded England would have been 11th (with a score of 556) rather than eighth equal in 2016, which was one place lower than 2011 (although their score would have been four higher than in 2011). No data are available on whether schools from the private sector participated in the previous PIRLS assessments in 2011.

The Year 5 pupils who participated in the 2016 PIRLS assessments were also the second cohort to be assessed by the revised Year 6 SATs in the summer of 2017 when 28% failed to reach the expected standard in reading. In 2018 and 2019 these figures were 25% and

27%, respectively, even though 89% and 90% of these cohorts passed the PSC. Thus, between 2017 and 2019 over one in four pupils, who have all been taught through SSP, have transferred to secondary school and failed to reach the desired standards in reading. This is particularly surprising given the extensive training pupils typically receive prior to the Key Stage 2 SATs (Mansell, 2007). Schools were notified initially in September 2015, that they had provisionally been selected to participate in the PIRLS assessment of 2016 and were given final confirmation in December 2015. It seems inconceivable that these schools will not have devoted time to preparing pupils for the assessments, which unlike SATs are similar in each assessment cycle and are readily available so they can be practised and teachers can teach to the test.

There is also the question of why the pupils who passed the PSC on the first attempt performed to a higher level on the comprehension component of the PIRLS assessment compared to the students who passed a year later when they were in Year 2 (Double et al., 2019). It seems reasonable to assume that students who have reached the pass mark on the PSC, irrespective of whether this was in Year 1 or Year 2, would have comparable attainments in comprehension by the time they were in Year 5. One possible interpretation of the data relates to what pupils were taught in Year 2, after failing the PSC in Year 1, compared to those that passed. It may be that teachers in their efforts to ensure that the Year 2 pupils passed the PSC, devoted an inordinate amount of time to teaching all the GPCs that could have been assessed, many of which rarely occur on the PSC (Darnell et al., 2017).

In contrast, those students who passed PSC first time, will presumably have spent most of Year 2 on activities designed to improve their reading fluency and comprehension. The question is why those pupils who passed the PSC when in Year 2, and will presumably still have had three academic years to develop their broader literacy skills, did not perform to a comparable level in comprehension to their peers who passed the PSC in Year 1. Perhaps the explanation lies in pupils failing a high stakes assessment early in their school careers, coupled with another year of intensive teaching in SSP, having a negative impact on their overall progress in learning to read.

The DfE only published details of its approved SSP programmes in September 2011 (ESPO, 2011) and gave schools until March 2013 to accept an offer of matched funding to buy these programmes. It seems unlikely that schools would have both purchased, and implemented these programmes effectively, during the 2011–2012 academic year in time for there to be a significant impact on pupils' performance on the PSC at the end of the academic year in June 2012. This was borne out by the fact that only 58% of pupils passed the PSC in 2012. If anything, it might be expected that the Year 1 pupils who failed the PSC, but who passed the following year, might have done better in the comprehension assessment on PIRLS, if this was attributable to their phonic knowledge, as they are more likely to have been taught phonics effectively through government approved programmes after teachers had an extra year to become proficient in teaching SSP. Presumably this is why the pass rate for the Year 1 pupils rose a little in 2013 to 69%.

One further piece of evidence emerging from PIRLS 2016, that is rarely addressed in discussions about standards in reading, concerns whether pupils in England enjoy reading. Not only did the first PSC take place in 2012 but Ofsted published 'Moving English Forward' (Ofsted, 2012) in the same year that required schools to 'develop policies to promote reading for enjoyment throughout the school'. This report followed an earlier Ofsted (2004) publication, 'Reading for Purpose and Pleasure' which reported that:

> Most schools used a structured reading scheme where children were only allowed to choose freely books that appealed to them after they had completed the scheme and become a 'free reader'.
>
> (p. 12)

> Lower attaining pupils rarely read the books that they most wanted to read as these were seen as too be too difficult. Lack of competence often led to negative attitudes as they rarely selected books and they 'saw reading as a chore'.
>
> (p. 13)

Evidence from two studies investigating pupils' attitudes to reading reflected Ofsted's findings. Clark and Poulton (2011) found that 40% of the students in their survey in London had 10 or fewer books at home. These children enjoyed reading less, read fewer books, had more negative attitudes towards reading, were more likely to read only when they had to and had lower attainments than their peers, who had greater access to books at home. Clark and Teravainen-Goff (2020) reported the outcomes of a survey of 57,000 children aged between 9 and 18 and 3700 children aged 5–8, that found levels of reading enjoyment have continued to decrease since 2016 with current levels last evidenced in 2013. There is the possibility that the focus in the early years on phonics and phonically regular schemes is leading pupils to perceive that the purpose in learning to read is to decode fluently rather than enjoyment.

In three earlier PIRLS reports (Mullis et al., 2003; Mullis et al., 2007; Mullis et al., 2012) pupils in England had a poorer attitude towards reading and read less often for fun than pupils in other English-speaking countries. Although the most recent PIRLS (2016) indicated that England improved its international standing in reading accuracy and comprehension, it still only came 34th out of 50 countries for reading for pleasure with only 35% of pupils liking reading very much. Children in four out of the seven highest performing countries enjoyed reading less than pupils in England suggesting that there may be a trade-off between high levels of achievement and pupils' enjoyment of reading. One plausible interpretation of the 2016 PIRLS data is that using decodable texts to teach reading may be a contributory factor in English children not enjoying reading. It is significant that Gibb has never publicly acknowledged that many children in England do not enjoy reading despite government reforms or that 28% of the pupils taking the KS2 SATs in 2017, just a year after PIRLS 2016, failed to meet government expectations in reading.

Does the evidence support recommended government policy and practice?

The government's policy and recommended practice in teaching early reading are encapsulated in the three DfE documents published in 2021 (DfE, 2021a,b,c). The key features can be summarised as: (i) teaching SSP first and fast through a DfE validated programme; (ii) decoding CEWs through emphasising GPCs that are not consistent with what they have been taught; and (iii) students applying their knowledge of GPCs to reading phonically decodable texts. The evidence discussed in Part 2 indicates that the evidence in favour of teaching phonics through SSP is strong, although this finding is not universally accepted (for example, see Bowers, 2020, 2021 as well as Buckingham, 2020 and Fletcher et al., 2021 for detailed responses to Bowers).

It has been argued in this article that the government's approach to teaching phonics can be characterised as phonics first, fast and only. Certainly the National Curriculum and Reading Framework embrace notions of reading for pleasure and students hearing a wide range of high quality literature but the descriptions of teaching phonics focus on this being the first and only approach to decoding unknown words. There is a lack of evidence to support this specific approach to teaching SSP. Buckingham (2020) summarises the views of many who recognise the research in favour of teaching SSP when stating that instruction in synthetic phonics 'takes place alongside meaning based instruction, including morphology, vocabulary and comprehension' (p. 106). Similarly, there is no evidence to support the DfE's

recommended approach to teaching CEWs which reflects the phonics first, fast and only methodology.

A major implication of teaching SSP for the government is the use of decodable texts to ensure that pupils: (i) apply their knowledge of GPCs to continuous text; (ii) only read texts containing GPCs that they have been taught; and (iii) experience 100% success in decoding texts to become motivated to read. The assumptions, while plausible, are not supported by evidence. Two studies have investigated the respective merits of high and low decodable texts, one with a small sample that needs to be replicated, but neither found support for the exclusive use of decodable texts to teach reading. Castles et al. (2018) suggest that there is a strong case for using decodable texts in the early stages of teaching reading as they provide students with immediate opportunities to apply the phonic skills they have been taught. Thereafter, the case for decodable texts weakens. However, they do not indicate when a time might be reached to switch from one to the other. The analysis of *Billy the Kid* and *The Three Billy Goats Gruff* (Solity & Vousden, 2009) illustrates that real books can provide the same function as reading schemes and give pupils immediate opportunities to apply their phonic knowledge.

Further evidence in favour of teaching synthetic phonics can be found in a recent critique of government policy in England (Wyse & Bradbury, 2022) even through its authors claimed that 'the most robust research evidence, from randomised controlled trials with longitudinal designs suggest that a balanced instruction approach is most likely to be successful' (p. 41). Wyse and Bradbury based their claim on their analysis of eight studies that were the result of their systematic qualitative meta-synthesis (SQMS). Wyse and Bradbury classified six of these studies as teaching balanced literacy (BL) with the two other articles teaching phonics (but not synthetic phonics [SP]) and whole language (WL).

However, five of the six BL studies have been misclassified. Two of these studies taught phonological awareness (manipulating phonemes in the absence of print; Koziminsky & Kozminsky, 1995; Lie, 1991) and three studies (Vadasy & Sanders, 2011, 2012, 2013) taught what would be recognised as SP and were defined by Vadasy and Sanders as teaching phonics. They stated that the aim of their study was to 'examine the efficacy of 20 weeks of phonics-based instruction' (p. 471). The one remaining BL study (Gunn et al., 2011) was not significant in terms of the standardised measures taken. The Lyster (2002) article, which Wyse and Bradbury claim teaches phonics but not SP, taught neither and instead compared two methods for teaching phonological awareness, one that focused on phonemes and the other on morphemes. So a more reasonable interpretation of the outcomes from the eight studies included in the SQMS is that they support teaching phonological awareness and SP, not BL.

Finally, there is no evidence from RCTs or quasi-experimental research that any of the 19 SSP programmes validated by the DfE in December 2021 have a positive impact on students' progress in learning to read. The EEF evaluations of *RWInc*. are compromised through having no baseline data on any measures and the primary outcome in the TLIF/NFER evaluation is the PSC, that the author of *RWInc*. helped to develop and is in part based on the content of *RWInc*. Equally, it has been suggested that data from the PSC do not demonstrate that students have mastered all the GPCs that schools are required to teach. There are also reasons to question the claim that the 2016 PIRLS assessments have demonstrated the impact of SSP or had a positive impact on students' attitudes to reading.

PART 3: INSTRUCTIONAL PSYCHOLOGY AND TEACHING READING

Instructional psychology researches the impact of psychological theories on effective teaching. Whereas most applications of psychology in the field of education have researched how children learn and how to teach through analysing aspects of children's cognitive

development, instructional psychology instead attempts to understand the teaching and learning process through analysing and understanding the learning environment (Brown & Chater, 2004). Instructional psychology draws on 'rational analysis', a theory within the field of cognitive psychology that shifts the emphasis from inducing what happens 'in the mind' to looking at the structure of the environment and how it influences cognition (Anderson, 1990).

Brown (1998) explains that within rational analysis, cognitive behaviour can be seen as an adaptive reflection of the structure of the environment. This suggests that skilled reading is statistically optimal with respect to the structure of the English spelling to sound mapping system. Brown describes the ROAR model (Rational, Optimal, Adaptive, Reading) which is based on the Optimal Reading Hypothesis, to explain how fluent readers become sensitive to the statistical regularities in written English. They read with maximum efficiency given the statistical structure of mappings from orthography to phonology and so the process of learning to read is seen as a statistical process (also see Arciuli, 2018; Sawi & Rueckil, 2019; Seidenberg, 2017).

The fundamental idea is that learners represent the environment in a maximally efficient way. Thus, competent readers develop skills so that what they learn will be statistically optimal (i.e., the rate at which features of written English are remembered or forgotten will reflect their frequency of occurrence) in relation to the structure of written English and the frequency of occurrence of graphemes and the phonemes that they represent. We remember what we need to know and forget what is less relevant. This being the case it follows that a key task in teaching reading is how best to characterise the (statistical) regularities in written English and represent them to beginning readers so that their performance becomes statistically optimal.

Chater (2018) explains that what distinguishes high performers in any field from lower performers is not a function of their hypothesised cognitive differences and brain functioning but the consequence of their perceptions of past experiences. So the extent to which pupils learn to match phonemes to graphemes depends on how information has been interpreted in the past. We are more likely to remember items that are relevant and to which we can attach meaning. The implications for teaching reading are two-fold. The first is that when using phonic knowledge to decode, pupils are likely to remember strings of GPCs that convey meaning. The second implication is that the rationale for what and how pupils are taught should be less about cognitive functioning, for example the working of short-term memory, cognitive load and executive functioning, and focus instead on structuring all aspects of the learning environment in an optimal manner.

Vousden (2008) provides a rationale for choosing how many and what types of spelling-to-sound units of English to teach. This is based on an analysis of the frequency with which various units of spelling-to-sound mapping occur in monosyllabic words (54,670 word types [number of different words] and 15,542,299 word tokens [total number of words]) in the English language. Analyses of spelling-to-sound mappings at three levels (whole words, onsets and rimes and graphemes) revealed that the distribution of these mappings in English text approximates to Zipf's Law (Zipf, 1949) (the frequency with which words occur starts high and then tapers off rapidly) and Pareto's 80/20 principle (a minority of causes, inputs or effort usually leads to the majority of results, outputs or rewards; Koch, 1998). Further analyses reveal that a substantial proportion of text can be read if knowledge of the most frequent mappings at each level is assumed.

Solity and Vousden (2009) explored the assumptions that have been made about the use of real books and reading schemes, which have tended to polarise arguments about their respective strengths and limitations. They analysed the structures of adult literature, children's real books and reading schemes, and examined the demands that they make on children's sight vocabulary and phonic skills and found that Zipf's Law again applied to the frequency of whole words and GPCs. Overall, approximately 90% of the monosyllabic words in all

three text types could be read by combining 89 high frequency, monosyllabic words and 64 GPCs. This outcome was counter intuitive as it was assumed, given the rhetoric, that a far higher percentage of words would be read in the schemes than either real books or adult literature. Solity and Vousden also calculated the number of onsets and rimes that would be required to teach the same percentage of words that can be read through 64 GPCs. For the adult literature this was 534 and for the children's real books 334. Solity and Vousden (2009) and Vousden (2008) provide a convincing case, in terms of the information that has to be remembered, for teaching reading through SSP rather than onset-rimes and analytic phonics when teaching an optimal level of information.

The original database developed by Solity and Vousden contained 66 real books and the analyses were conducted on monosyllabic words. Solity (2020) explains how this database (known as the *Real Books Database* [*RBD*]) has increased considerably and now contains 1500 books, 21,733 word types and 1,512,711 word tokens and the analyses have been extended to include both mono and polysyllabic words. The analyses reported by Solity (2020) identified: (i) the most frequently occurring words in children's real books as well as (ii) the most frequently occurring GPCs. Sixteen words accounted for 29.05% of the words on the *RBD* and 100 accounted for 54.16% but the next 50 words (101–150) only accounted for just over 5.46%. This is consistent with the findings of Solity and Vousden and Zipf's Law.

The percentage of high frequency words on the *RBD* has been compared to two databases derived largely from reading schemes: the *Children's Early Reading Vocabulary* (*CERV*) (Stuart et al., 2003) and the *Children's Printed Word Database* (*CPWD*) (Masterson et al., 2010) that was used in *Letters & Sounds* (DfES, 2007). The percentage of words on each database that could in theory be read by the 100 most frequently occurring words within each database was calculated. The most striking outcome of the analysis is that the high frequency words in the *RBD* accounted for a higher percentage of the total number of words than the high frequency words on the *CERV* and *CPWD* databases that included books that were specifically written to include a high proportion of CEWs. These data suggest that it is difficult to write anything, whether it be adult literature, children's real books or reading schemes, without including a high percentage of a relatively small number of core words.

Solity (2020) also reported the analyses of the GPCs contained in the phonically regular words on the *RBD* that indicated that a small number of GPCs enable pupils to read the majority of phonically regular words in children's literature, an outcome that also reflects Zipf's Law and Pareto's 80/20 principle. The ten most frequently occurring GPCs account for 56.39% of all the GPCs in phonically regular words and the 20 most frequently occurring GPCs account for 81.07%. Single letter GPCs account for 80.49% of the GPCs in phonically regular words whereas multiple letter GPCs account for the remaining 19.51%. The analyses indicate that there are certain GPCs that appear frequently in children's real books and others that are far less frequent. For example, the 25 lowest frequency GPCs in *Jolly Phonics* (Lloyd & Wernham, 2013), *Letters & Sounds* (DfES, 2007) and *RWInc.* (Miskin 2020) accounted for just 5.39%, 1.15% and 1.84% respectively of the phonically regular words on the *RBD*.

There is one additional factor that needs to be addressed when teaching GPCs and that concerns the implications of teaching more than one phoneme for each grapheme. The main consequence is that a significant number of words that are decoded phonically will have an alternative, phonically correct pronunciation. For example, as stated previously the grapheme <ch> can represent three phonemes, (e.g., /tʃ/—chip; /k/—school; /ʃ/—chef) and the grapheme 'i' can represent two phonemes in p͟ip and p͟ipe so <chip> can represent six phonically plausible pronunciations. It is therefore of interest to know: (i) the number of graphemes with multiple pronunciations and (ii) the percentage of words decoded phonically that would have an alternative pronunciation within three SSP programmes ([*Jolly Phonics*, Lloyd & Wernham, 2013]; [*Letters & Sounds*, DfE, 2007]; [*RWInc.*, Miskin, 2020]).

Solity (2020) reported that in *Jolly Phonics* five phonemes represent multiple graphemes, whereas the figures for *Letters & Sounds* and *RWInc*. are 34 and 13 respectively. The percentage of words on the *RBD* that could be decoded accurately but have multiple pronunciation are: 16.08% for *Jolly Phonics*, 94% for *Letters & Sounds* and 67% for *RWInc*. Thus, when pupils are taught multiple phonemes for graphemes many of the words that can be decoded will have an alternative pronunciation. When this occurs it is pupils' vocabulary knowledge and use of context that leads them to the correct pronunciation. This is an example of Set for Variability and mispronunciation correction rather than a case of pupils using context to guess the correct pronunciation. Pupils would not be able to pronounce words containing multiple pronunciations accurately without resorting to the context in which they appear.

It is questionable whether there are any benefits in teaching either lower frequency GPCs or multiple phonemes for graphemes given the potential confusion that they cause. However, this is a requirement in the National Curriculum (DfE, 2013a) and is a feature of most commercial SSP programmes. This is a significant issue particularly for lower achieving pupils. A considerable amount of instructional time is devoted to teaching GPCs that must be mastered in order that pupils progress through the scheme but which will not be useful when children start to read real books. Potentially the instructional time would be better spent focusing on the small number of GPCs that occur most frequently in real books and devoting time to explicitly developing pupils' language skills and vocabulary knowledge. This will enable them to maximise the impact of their phonic knowledge through mispronunciation correction (Dyson et al., 2017) and SfV (Tunmer & Chapman, 2012).

A second issue that relates to teaching lower achieving pupils is that the goal for many children following a reading scheme is to become a 'free reader' and to progress to chapter books. However, it will take lower achieving pupils much longer to reach this stage than their peers since they are only given a new scheme book when it is thought that they have mastered the GPCs and CEWs required to read a book with 100% accuracy. So, they will read fewer books and have less practice in reading continuous prose than their peers. This is the opposite of what they require. Finally, there is a troubling consequence of having to read books from a scheme with 100% accuracy. Before being given a book pupils are taught to decode all the phonically regular words that will appear in the book out of context. As a result, they do not encounter unfamiliar words when reading continuous text and so do not get the opportunity to practise their decoding skills.

The research in instructional psychology has demonstrated that, despite the perceived irregular nature of written English, a large proportion can be read through being taught the most frequently occurring GPCs and CEWs. The consequence of failing to do so is that instructional time will be devoted to being taught GPCs that are largely redundant as they occur so infrequently. The value in teaching a small number of phonically irregular high frequency words is that since they occur so frequently the process of statistical learning, together with explicit instruction, will facilitate pupils' acquisition of these words. The limitations in teaching SSP through decodable texts, discussed in Part 3, are potentially a contributory factor in over 25% of pupils transferring to secondary school having failed to reach the required standards in the Year 6 SATs despite having passed the PSC.

CONCLUSION

This article has reviewed the research on teaching reading in England over the last 50 years and its impact on the policies and the approved practice of successive governments. Whereas initially it was generally recommended that pupils were taught to develop a sight vocabulary before being introduced to phonics, the advice has changed so that current government policy appears to promote a version of phonics first, fast and only as the initial

route to learning to read, where the initial focus is on accurate decoding rather than understanding. Children are then taught to read CEWs through their knowledge of GPCs, and the exceptions to taught GPCs, rather than by sight.

Teaching SSP has been recognised as an effective way of teaching pupils to decode for many years and it is perhaps the one area of current government policy that is well supported by research. However, there are many unanswered questions about how best to teach SSP since government validated programmes vary considerably in terms of: (i) the number of GPCs taught; (ii) the number of graphemes that represent multiple phonemes; (iii) whether pupils are taught to synthesise phonemes and segment words into phonemes as phonological skills (i.e. no print); (iv) teaching methodology; (v) assessment arrangements; and (vi) the broader literacy skills taught. These differences are significant barriers to understanding and researching the essential components of SSP programmes. The government's 16 core criteria represent just one way of teaching synthetic phonics but there is no evidence to demonstrate that it is more effective than other ways of teaching synthetic phonics.

There is also no evidence to demonstrate that synthetic phonics is best taught through decodable texts and their use is challenged by research derived from instructional psychology and a detailed understanding of the GPCs and CEWs that are required to read real books. The evidence also indicates that current policies are failing a significant number of children who are not reaching the levels expected when they transfer from primary to secondary school despite over 10 years of government policy favouring teaching SSP through decodable texts. The research in instructional psychology suggests that this level of failure could potentially be addressed through teaching an optimal number of GPCs through synthetic phonics and real books.

In the past the tension has been between advocates of whole language and phonics based methods of teaching reading. The research reviewed in this article suggests that the debate, and research, should now move on and instead focus on: (i) how much phonics to teach; (ii) the books that will not only facilitate pupils' progress but also lead them to develop a positive attitude towards reading; (iii) how best to develop pupils' reading skills beyond phonics; and (iv) the instructional methods that prevent reading difficulties.

CONFLICT OF INTEREST

The author is not aware of any conflicts of interest in relation to this article.

ETHICAL APPROVAL

This article is based on a review of published studies. Since no participants were involved, an ethical review was not applicable.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analysed.

REFERENCES

Adams, M. J. (1990). *Beginning to read*. MIT Press.

Anderson, J. R. (1990). *The adaptive character of thought*. LEA Publishers.

Arciuli, J. (2018). Reading as statistical learning. *Language, Speech and Hearing Difficulties in Schools*, *49*, 634–643. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0135

Austin, C. R., Vaughn, S., Clemens, N. H., Pustejovsky, J. E., & Boucher, A. N. (2021). The relative effects of instruction linking word reading and word meaning compared to word reading instruction alone on the accuracy, fluency, and word meaning knowledge of 4th-5th grade students with dyslexia. *Scientific Studies of Reading*, 1–19. https://doi.org/10.1080/10888438.2021.1947294

Ball, E. W., & Blachman, B. A. (1991). Does phoneme segmentation training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, *26*, 49–66.

Bowers, J. S. (2020). Reconsidering the evidence that systematic phonics is more effective than alternative methods of reading instruction. *Educational Psychology Review*, *32*(3), 681–705. https://doi.org/10.1007/s10648-019-09515-y

Bowers, J. S. (2021). Yes children need to learn their GPCS but there really is little or no evidence that systematic or explicit phonics is effective – A response to Fletcher, Savage and Sharon (2020). *Educational Psychology Review*, *33*, 1965–1979. https://doi.org/10.1007/s10648-021-09602-z

Bradley, L., & Bryant, P. E. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, *271*, 746–747. https://doi.org/10.1038/271746a0

Bradley, L., & Bryant, P. E. (1983). Categorising sounds and learning to read – A causal connection. *Nature*, *301*, 419–421.

Brown, G. D. A. (1997). Connectionism, phonology, reading and regularity in developmental dyslexia. *Brain and Language*, *59*, 207–235. https://doi.org/10.1006/brln.1997.1817

Brown, G. D. A. (1998). The endpoint of skilled word recognition: The ROAR model. In J. L. Metsala, & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 121–138). Lawrence Erlbaum.

Brown, G. D. A., & Chater, N. (2004). Connectionist models of children's reading. In T. Nunes, & P. E. Bryant (Eds.), *Handbook of children's literacy* (pp. 67–89). Kluwer Academic Publishers.

Brown, G. D. A., & Deavers, R. (1999). Units of analysis in nonword reading: Evidence from children and adults. *Journal of Experimental Child Psychology*, *73*, 208–242. https://doi.org/10.1006/jecp.1999.2502

Buckingham, J. (2020). Systematic phonics instruction belongs in evidence-based reading programs: A response to bowers. *The Educational and Developmental Psychologist*, *37*, 105–113. https://doi.org/10.1017/edp.2020.12

Carnine, D. W., Silbert, J., & Kame'enui, E. J. (1997). *Direct instruction reading*. Prentice Hall.

Carter, J. (2020). The assessment has become the curriculum: Teachers' views on the Phonics Screening Check in England. *British Journal of Education*, *46*(3), 593–609. https://doi.org/10.1002/berj.3598

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Central Advisory Council for Education (CACE) (1967). *Children and their primary schools ('The Plowden Report')*. HMSO.

Chater, N. (2018). *The mind is flat: The illusion of mental depth and the improvised mind*. Allen Lane.

Christensen, C. A., & Bowey, J. A. (2005). The efficacy of orthographic rime, grapheme-phoneme-correspondences, and implicit phonics approaches to teaching decoding skills. *Scientific Studies of Reading*, *9*, 327–349.

Clark, C., & Poulton, L. (2011). *Book ownership and its relation to reading enjoyment, attitudes behaviour and attainment*. National Literacy Trust.

Clark, C., & Teravainen-Goff, A. (2020). *Children and young people's reading in 2019: Findings from our annual literacy survey*. National Literacy Trust.

Clark, M. (2018). How justified is it to dominate government policy on early literacy with synthetic phonics and the phonics screening check? Evidence, effects and expenditure. Part I: Recent evidence on effects and expenditure. *Education Journal*, *351*, 18–21.

Colenbrander, D., Miles, K. P., & Ricketts, J. (2019). To see or not to see: How does seeing spellings support vocabulary learning? *Language, Speech and Hearing Services in Schools*, *50*, 609–628. https://doi.org/10.1044/2019_LSHSS-VOIA-18-0135

Colenbrander, D., Wang, H. C., Arrow, T., & Castles, A. (2020). Teaching irregular words: What we know, what we don't, and where we can go from here. *The Educational and Developmental Psychologist*, *37*, 97–104.

Darnell, C. A., & Solity, J. E. (in preparation). *Decoding the phonics screening check: Revisited (2012–2019)*.

Darnell, C. A., Solity, J. E., & Wall, H. (2017). Decoding the phonics screening check. *British Educational Research Journal*, *43*(3), 505–527. https://doi.org/10.1002/berj.3269

Deavers, R., Solity, J. E., & Kerfoot, S. (2000). The effect of instruction on early nonword reading strategies. *Journal of Research in Reading*, *23*(3), 267–286.

Department for Education (DfE) (2010). *Phonics teaching materials: Core criteria and the self-assessment process*. HMSO.

Department for Education (DfE) (2011). *Year 1 phonics screening check: Pilot 2011-technical report*. STA.

Department for Education (DfE) (2012). *Assessment framework for the development of the year 1 phonics screening check*. STA.

Department for Education (DfE) (2013a). *The National Curriculum in England: Framework document*. HMSO.

Department for Education (DfE) (2013b). *English programmes of study: Key stages 1 and 2 National Curriculum in England*. HMSO.

Department for Education (DfE) (2015). *Phonics partnership grants*. https://bit.ly/3Gdk62Y

Department for Education (DfE) (2018). *English hubs: List of primary schools*. https://bit.ly/3JGfnb0http://www.englishhubs.org

Department for Education (DfE) (2021a). *Validation of systematic synthetic phonics programmes: Supporting documentation*. HMSO.

Department for Education (DfE) (2021b). *The reading framework: Teaching the foundations of literacy*. HMSO.

Department for Education (DfE) (2021c). *National Professional Qualification (NPQ): Leading literacy framework*. HMSO.
Department for Education (DfE) (2022). *Choosing a phonics teaching programme: January 2022 update to the phonics validation process*. https://www.gov.uk/government/publications/choosing-a-phonics-teaching-programme/list-of-phonics-teaching-programmes#list-of-validated-systematic-synthetic-phonics-ssp-programmes
Department for Education & Employment (DfEE) (1996). *The national literacy project: Framework for teaching*. HMSO.
Department for Education & Employment (DfEE) (1998). *The national literacy strategy: Framework for teaching*. HMSO.
Department for Education & Employment (DfEE) (1999a). *National literacy strategy: Review of research and other related evidence*. HMSO.
Department for Education & Employment (DfEE) (1999b). *Progression in phonics*. HMSO.
Department for Education & Employment (DfEE) (1999c). *Additional literacy support*. HMSO.
Department for Education & Employment (DfEE) (2000). *Watching and learning: OISE/UT evaluation of the implementation of the National Literacy and Numeracy Strategies*. DfEE.
Department for Education & Employment (DfEE) (2001). *Watching and learning 2: OISE/UT evaluation of the implementation of the National Literacy and Numeracy Strategies*. DfEE.
Department for Education and Science (DES) (1975). *A language for life (The Bullock Report)*. HMSO.
Department for Education and Science (DES) (1978). *Primary education in England: A survey of HMI*. HMSO.
Department for Education and Science (DES) (1988). *Education reform act*. HMSO.
Department for Education and Science (DES) (1990). *The teaching and learning of reading in primary schools*. HMSO.
Department for Education & Skills (DfES) (2001). *Early literacy support*. HMSO.
Department for Education & Skills (DfES) (2003). *Watching and learning 3: OISE/UT evaluation of the implementation of the National Literacy and Numeracy Strategies*. DfES.
Department for Education & Skills (DfES) (2004). *Playing with sounds*. HMSO.
Department for Education & Skills (DfES) (2007). *Letters & sounds: Principles and practice of high quality phonics*. HMSO.
Double, K. S., McGrane, J. A., Stiff, J. C., & Hopfenbeck, T. N. (2019). The importance of early phonics improvements for predicting later comprehension. *British Educational Research Journal*, *45*(6), 1220–1234.
Duncan, L. G., Seymour, P. H. K., & Hill, S. (1997). How important are rhyme and analogy in beginning reading? *Cognition*, *63*, 171–208. https://doi.org/10.1016/S0010-0277(97)00001-2
Duncan, L. G., Seymour, P. H. K., & Hill, S. (2000). Small to large unit progression in metaphonological awareness and reading. *The Quarterly Journal of Experimental Psychology*, *53a*(4), 1081–1104.
Dyson, H., Best, W., Solity, J. E., & Hulme, C. (2017). Training mispronunciation correction and word meanings improves children's ability to learn to read words. *Scientific Studies of Reading*, *21*(5), 392–407. https://doi.org/10.1080/10888438.2017.1315424
Eastern Shires Purchasing Organisation (ESPO) (2011). *The importance of phonics: A catalogue of systematic synthetic phonics products and training*. ESPO.
Eastern Shires Purchasing Organisation (ESPO) (2012). *The importance of phonics: A supplementary catalogue of systematic synthetic phonics products and training*. ESPO.
Education and Skills Committee (2005). *Teaching children to read*. H.M.S.O.
Ellis, S. (2009). Policy and research: lessons from the Clackmannanshire synthetic phonics initiative. In F. Fletcher-Campbell, J. Soler, & G. Reid (Eds.), *Approaching difficulties in literacy development: Assessment, pedagogy and programmes* (pp. 39–51). Sage.
Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and practice*. Irvington.
Fletcher, J. M., Savage, R., & Vaughn, S. (2021). A commentary on Bowers (2020) and the role of phonics instruction in reading. *Educational Psychology Review*, *33*, 1249–1274. https://doi.org/10.1007/s10648-020-09580-8
Gibb, N. (2017a). *The evidence in favour of teacher-led instruction*. Speech given to the Education World Forum about Teacher-Led Instruction, London: 24th January.
Gibb, N. (2017b). *Reading is the key to unlocking human potential*. Speech given at a presentation of England's successful Progress in International Reading Literacy Study results, British Library: 5th December.
G.L. Assessment (2014). *New group reading test*. G.L. Assessment.
Goodman, K. S. (1976). Reading: A psycholinguistic guessing game. In H. Singer, & R. B. Ruddell (Eds.), *Theoretical models and process of reading* (pp. 497–508). International Reading Association.
Goodman, K. S. (1986). *What's the whole in whole language*. Scholastic Press.
Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology*, *42*, 73–83. https://doi.org/10.1016/0022-0965(86)90016-0

Goswami, U. (1988). Orthographic analogies and reading development. *Quarterly Journal of Experimental Psychology*, *40A*, 239–268. https://doi.org/10.1080/02724988843000113
Goswami, U. (1990a). A special link between rhyming skill and the use of orthographic analogies by beginning readers. *Journal of Child Psychology and Psychiatry*, *31*, 301–311. https://doi.org/10.1111/j.1469-7610.1990.tb01568.x
Goswami, U. (1990b). Phonological priming and orthographic analogies in reading. *Journal of Experimental Child Psychology*, *49*, 323–340. https://doi.org/10.1016/0022-0965(90)90062-D
Goswami, U. (1991). Learning about spelling sequences: The role of onsets and rimes in analogies in reading. *Child Development*, *62*, 1110–1123. https://doi.org/10.2307/1131156
Goswami, U. (1993). Towards an interactive analogy model of reading development: Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology*, *56*, 443–475.
Goswami, U., & Bryant, P. E. (1990). *Phonological skills and learning to read*. Lawrence Erlbaum Associates.
Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*, 6–10.
Gunn, B., Smolkowski, K., & Vadasy, P. (2011). Evaluating the effectiveness of read well kindergarten. *Journal of Research on Educational Effectiveness*, *4*, 53–86.
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 Psychology and Psychiatry*, *45*, 338–358. https://doi.org/10.1111/j.1469-7610.2004.00225.x
Hulme, C., Muter, V., & Snowling, M. (1998). Segmentation does predict early progress in learning to read better than rhyme: A reply to Bryant. *Journal of Experimental Child Psychology*, *71*, 39–44. https://doi.org/10.1006/jecp.1998.2456
Jenkins, J. R., Peyton, J. A., Sanders, E. A., & Vadasy, P. F. (2004). Effects of reading decodable texts in supplemental first grade tutoring. *Scientific Studies of Reading*, *8*(1), 53–85. https://doi.org/10.1207/s1532799xssr0801_4
Johnston, R. S., & Watson, J. E. (2004). Accelerating the development of reading, spelling and phonemic awareness skills in initial readers. *Reading and Writing*, *17*, 327–357. https://doi.org/10.1023/B:READ.0000032666.66359.62
Kirschner, P., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem based, experiential and inquiry based teaching. *Educational Psychologist*, *41*(2), 75–86.
Koch, R. (1998). *The 80/20 principle: The secret to success by achieving more with less*. Doubleday.
Kozminsky, L., & Kozminsky, E. (1995). The effects of early phonological awareness training on reading success. *Learning and Instruction*, *5*, 187–201. https://doi.org/10.1016/0959-4752(95)00004-M
Levy, R. (2009). Children's perceptions of reading and the use of reading scheme texts. *Cambridge Journal of Education*, *39*(3), 361–377. https://doi.org/10.1080/03057640903103769
Lie, A. (1991). Effects of a training program for stimulating skills in word analysis in first-grade child. *Reading Research Quarterly*, *26*(3), 234–250.
Literacy Task Force (1997). *A reading revolution: How we can teach every child to read well*. Literacy Task Force.
Lloyd, S., & Wernham, S. (2013). *The phonic handbook*. Jolly Learning Ltd.
Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in pre-school children. *Reading Research Quarterly*, *23*, 262–284. https://doi.org/10.1598/RRQ.23.3.1
Lyster, S. A. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. *Reading and Writing*, *15*, 261–294.
Mansell, W. (2007). *Education by numbers: The tyranny of testing*. Methuen.
Masterson, M., Stuart, M., Dixon, M., & Lovejoy, J. (2010). Children's printed word database: Continuities and changes over time in children's early reading vocabulary. *British Journal of Educational Psychology*, *101*, 221–242. https://doi.org/10.1348/000712608X371744
McGrane, J., Stiff, J., Baird, J., Lenkeit, J., & Hopfenbeck, T. (2017). *Progress in international reading literacy study: National report for England*. DfE.
Miller, F. (2016). Forty years after the Ruskin speech, education needs another moment, *The Guardian*, 13th December.
Miskin, R. (2020). *Read, Write Inc., phonics handbook 1*. Oxford University Press.
Morris, J. M. (1958). The relative effectiveness of different methods of teaching: The place and value of phonics. *Educational Research*, *1*(1), 38–49.
Morris, J. M. (1959). *Reading in the primary school*. Newness Educational Publishing.
Morris, J. M. (1984). Focus on phonics: Phonics 44 for initial literacy in English. *Reading*, *18*(1), 13–24. https://doi.org/10.1111/j.1467-9345.1984.tb00784.x
Morris, J. M. (1994). Phonicsphobia. *Journal of the Simplified Spelling Society*, *17*(2), 3–12.

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *International results in reading 2011*. TIMMS & PIRLS International Study Centre.
Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *International results in reading 2016*. TIMMS & PIRLS International Study Centre.
Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001, international report*. TIMMS & PIRLS International Study Centre.
Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006, international report*. TIMMS & PIRLS International Study Centre.
Munton, G. (2008). *Billy the kid*. Oxford University Press.
Murray, W. (1969). *Teaching reading: Handbook of the ladybird key words reading scheme*. Ladybird Books Ltd.
Muter, V., Snowling, M., & Taylor, S. (1994). Orthographic analogies and phonological awareness: Their role and significance in early reading development. *Journal of Child Psychology and Psychiatry*, *35*(2), 293–310. https://doi.org/10.1111/j.1469-7610.1994.tb01163.x
Ofsted (1996). *The teaching of reading in 45 inner London primary schools*. Ofsted.
Ofsted (1998). *The National Literacy Project: An HMI evaluation*. Ofsted.
Ofsted (1999). *The National Literacy Strategy: An interim evaluation*. Ofsted.
Ofsted (2000). *The National Literacy Strategy: The second year*. Ofsted.
Ofsted (2001). *The National Literacy Strategy: The third year*. Ofsted.
Ofsted (2002). *The National Literacy Strategy: The first four years*. Ofsted.
Ofsted (2004). *Reading for purpose and pleasure: An evaluation of the teaching of reading in primary schools*. Ofsted.
Ofsted (2010). *Reading by six: How the best schools do it*. Ofsted.
Ofsted (2012). *Moving English forward: Action to raise standards in English*. Ofsted.
Ofsted (2019). *The education inspection framework*. Ofsted.
Petscher, Y., Caball, S. Q., Catts, H. W., Compton, D. L., Foorman, B. R., Hart, S. A., Lonigan, C. J., Phillips, B. M., Schatschneider, C., Steacy, L. M., Terry, N. P., & Wagner, R. (2020). How the science of reading informs 21st century education. *Reading Research Quarterly*, *55*(S1), 5267–5282. https://doi.org/10.1002/rrq.352
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115. https://doi.org/10.1037/0033-295X.103.1.56
Price-Mohr, R., & Price, C. (2020). A comparison of children aged 4–5 years learning to read through instructional texts containing either a high or a low proportion of phonically-decodable words. *Early Childhood Education Journal*, *48*, 39–47.
Rhine, W. R. (Ed.) (1981). *Making schools more effective: New directions from follow through*. Academic Press.
Rose, J. (2006). *Independent review of the teaching of early reading (The Rose Report)*. DfES.
Rosenshine, B. (2010). *Principles of instruction*. International Academy of Education.
Rosenshine, B. (2012). Principles of instruction: Research based strategies that all teachers should know. *American Educator*, *36*, 12–39.
Sainsbury, M., Schagen, I., Whetton, C., Hagues, N., & Minnis, M. (1998). *Evaluation of the national literacy project*. DfEE.
Sawi, O., & Rueckil, J. (2019). Reading and the neurocognitive basis of statistical learning. *Scientific Studies of Reading*, *23*(1), 8–23.
Seidenberg, M. (2017). *Language at the speed of sight: How we read, why so many can't, and what can be done about it*. Basic Books.
Seidenberg, M. S., Borkenhagen, M. C., & Kearn, D. M. (2020). Lost in translation? Challenges in connecting reading science and educational practice. *Reading Research Quarterly*, *55*(S1), S119–S130. https://doi.org/10.1002/rrq.341
Shanahan, T. (2020). What constitutes a science of reading instruction? *Reading Research Quarterly*, *55*(S1), S235–S247.
Shapiro, L., & Solity, J. E. (2008). Delivering phonological and phonics training within whole-class teaching. *British Journal of Educational Psychology*, *78*(4), 597–620. https://doi.org/10.1348/000709908X293850
Shapiro, L., & Solity, J. E. (2016). Differing effects of two synthetic phonics programmes on early reading development. *British Journal of Educational Psychology*, *86*(2), 182–203. https://doi.org/10.1111/bjep.12097
Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education: Contributions from Educational Psychology*, *1*, 1–57.
Sharratt, N., & Tucker, N. S. (2005). *The three billy goats gruff*. Macmillan.
Skailand, D. B. (1971). *A comparison of four language units in teaching beginning reading*. Paper presented at the meeting of the American Educational Research Association, New York, USA, 4–7 February 1971.
Smith, F. (1973). *Psycholinguistics and reading*. Holt, Rinehart, & Winston.
Smith, F. (1978). *Understanding reading: A psycholinguistic analysis of reading and learning to read* (2nd ed.). Holt, Rinehart, & Winston.

Solity, J. E. (2003). *Teaching phonics in context: A critique of the National Literacy Strategy*. Phonics Seminar: DfES.
Solity, J. E. (2020). Instructional psychology and teaching reading: Ending the reading wars. *The Educational and Developmental Psychologist*, *37*, 123–132. https://doi.org/10.1017/edp.2020.18
Solity, J. E., Deavers, R., Kerfoot, S., Crane, G., & Cannon, K. (1999). Raising literacy attainments in the early years: The impact of instructional psychology. *Educational Psychology*, *19*(4), 373–395. https://doi.org/10.1080/0144341990190401
Solity, J. E., Deavers, R., Kerfoot, S., Crane, G., & Cannon, K. (2000). The Early Reading research: The impact of instructional psychology. *Educational Psychology in Practice*, *16*(2), 109–129. https://doi.org/10.1080/02667360050122190
Solity, J. E., & Shapiro, L. R. (2008). Developing the practice of educational psychologists through theory and research. *Educational and Child Psychology*, *25*(3), 123–149.
Solity, J. E., & Vousden, J. (2009). Real books and reading schemes: A perspective from instructional psychology. *Educational Psychology*, *29*(4), 469–511.
Stainthorp, R. (2020). A National Intervention in Teaching Phonics: A case study from England. *The Educational and Developmental Psychologist*, *37*(2), 114–122. https://doi.org/10.1017/edp.2020.14
Stannard, J., & Huxford, L. (2007). *The literacy game: The story of the National Literacy Strategy*. Routledge.
Stockard, J., Wood, T. W., Coughlin, C., & Rasplica Khoury, C. (2018). The effectiveness of direct instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research*, *88*(4), 479–507. https://doi.org/10.3102/0034654317751919
Stuart, M. (1999). Getting ready for reading: Early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology*, *69*, 587–605. https://doi.org/10.1348/000709999157914
Stuart, M. (2003). *Fine tuning the National Literacy Strategy to ensure continuing progress in improving standards of reading in the UK: Some suggestions for change*. Phonics Seminar, London: DfES
Stuart, M., Dixon, M., Masterson, J., & Gray, B. (2003). Children's early reading vocabulary: Description and word frequency Lists. *British Journal of Educational Psychology*, *73*, 585–598.
Stuart, M., Masterson, J., & Dixon, M. (2000). Spongelike acquisition of sight vocabulary in beginning readers? *Journal of Research in Reading*, *23*(1), 12–27. https://doi.org/10.1111/1467-9817.00099
Stuart, M., & Stainthorp, R. (2016). *Reading development and teaching*. Sage.
The National Reading Panel (NRP), & NICHD (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. National Institute for Child Health and Human Development Clearinghouse.
Topping, K. (2014). Paired reading and related methods for improving fluency. *International Electronic Journal of Elementary Education*, *7*(1), 57–70.
Torgerson, C., Brooks, G., Gascoine, L., & Higgins, S. (2018). Phonics: Reading policy and the evidence of effectiveness from a systematic 'tertiary' review. *Research Papers in Education*, *34*, 208–238. https://doi.org/10.1080/02671522.2017.1420816
Torgerson, C., Brooks, G., & Hall, J. (2006). *A systematic review of the research literature on the use of phonics in the teaching of reading and spelling*, (Research Report RR711). DfES.
Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, *34*(1), 33–58. https://doi.org/10.1177/002221940103400104
Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, *16*, 122–140. https://doi.org/10.1080/10888438.2010.542527
Turner, M. (1990). *Sponsored reading failure*. IPSET Education Unit.
Vadasy, P., & Sanders, E. A. (2011). Efficacy of supplemental phonics based instruction for low skilled first graders how language minority status and pretest characteristics moderate. *Scientific Studies of Reading*, *15*(6), 471–497.
Vadasy, P., & 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*(4), 987–1005. https://doi.org/10.1037/a0028163
Vadasy, P., & Sanders, E. A. (2013). Two-year follow-up of a code-oriented intervention for lower-skilled first-graders – The influence of language status and word reading skills on third-grade literacy outcomes. *Reading and Writing*, *26*, 821–843. https://doi.org/10.1007/s11145-012-9393-4
Valentini, A., Ricketts, J., Pye, R. E., & Housten-Price, C. (2018). Listening while reading promotes word learning from stories. *Journal of Experimental Child Psychology*, *167*, 10–31. https://doi.org/10.1016/j.jecp.2017.09.022
Vousden, J. (2008). Units of English spelling-to-sound mapping: A rational approach to reading instruction. *Applied Cognitive Psychology*, *22*, 247–272. https://doi.org/10.1002/acp.1371

Vousden, J. I., Ellefson, M. R., Solity, J. E., & Chater, N. (2011). Simplifying reading: Applying the simplicity principle to reading. *Cognitive Science*, *35*, 34–78.
Wendon, L. (2003). *Letterland teacher's guide*. Collins.
Wyse, D., & Bradbury, A. (2022). Reading wars or reading reconciliation? A critical examination of robust research evidence, curriculum policy and teachers' practices for teaching phonics and reading. *Review of Education*, *10*(1), 1–53. https://doi.org/10.1002/rev3.3314
Wyse, D., & Goswami, U. (2008). Synthetic phonics and the teaching of reading. *British Education Research Journal*, *34*(6), 691–710. https://doi.org/10.1080/01411920802268912
Wyse, D., & Styles, M. (2007). Synthetic phonics and the teaching of reading: The debate surrounding England's 'rose report'. *Literacy*, *41*(1), 35–42. https://doi.org/10.1111/j.1467-9345.2007.00455.x
Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.

How to cite this article: Solity, J. (2022). Instructional Psychology and Teaching Reading: An analysis of the evidence underpinning government policy and practice. *Review of Education*, *10*, e3349. https://doi.org/10.1002/rev3.3349