



A systematic literature review of decodable and levelled reading books for reading instruction in primary school contexts: an evaluation of quality research evidence

UNIVERSITY OF NEWCASTLE
RESEARCH ALLIANCE FOR LANGUAGE, LITERATURE, AND LITERACY

Rachel Birch

Associate Lecturer, School of Education
University of Newcastle

Associate Professor Heather Sharp

Deputy Head of School, Research
Training
University of Newcastle

Dr Drew Miller

Senior Lecturer, School of Education
University of Newcastle

Denyse Ritchie

Literacy Fellow, Murdoch University

Professor Susan Ledger

Head of School, Dean of Education
University of Newcastle

TEACHERS
& TEACHING
RESEARCH CENTRE

© 2022 by Rachel Birch, Heather Sharp, Drew Miller, Denyse Ritchie, Susan Ledger is licensed under [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/).

ISBN 978-0-7259-0196-

THE REVIEW TEAM

Rachel Birch

Rachel is an associate lecturer at the University of Newcastle, Australia, where she coordinates undergraduate core courses in the Bachelor of Education degree program. She holds a Bachelor of Teaching/Arts with a major in literacy, and has over a decade of teaching experience in primary school and tertiary settings. Rachel teaches literacies with a focus on the English curriculum and children's literature. Her research interests include: literacies for the 21st century; student engagement and learners' self-efficacy; and best practices in reading and writing instruction.

Associate Professor Heather Sharp

Heather is an associate professor in the School of Education, University of Newcastle and Deputy Head of School Research Training. After working as a secondary school teacher (English and History) for several years, Heather began working in the tertiary sector in 2006. She has taught and coordinated Education courses in the areas of transformative pedagogies, philosophies of education, History curriculum and pedagogies, literacies, education as a professional career, assessment, and professional experience. Her research expertise is in qualitative approaches to curriculum analysis.

Dr Andrew Miller

Andrew (Drew) is the Deputy Director at the Teachers and Teaching Research Centre and a senior lecturer in the School of Education, University of Newcastle. Drew's research is focused on securing optimal development of children and young people in both physical and academic settings and finding innovative and practical ways to improve these outcomes at-scale. Drew provides significant depth of understanding and expertise in high quality evidence production and quantitative data analysis for the Teachers and Teaching Research Centre. His goal is to make the complexities of rigorous research more comprehensible for teachers and school leaders, enabling schools to successfully implement high quality evidence-based strategies into ongoing practice.

Denyse Ritchie

Honorary Chair of Literacy and Fellow, Murdoch University, Western Australia, and co-author and developer of THRASS, a pedagogical practice for teaching literacy acquisition skills. Denyse has worked extensively in teaching and teacher training and delivers professional development in English linguistics, phonetics, and orthography. Denyse has worked in mainstream teaching, remote indigenous communities, and internationally to promote the importance of teacher knowledge in phonological awareness, phonetics, phonics, and orthography as the foundation to literacy teaching and learning. Denyse has also worked extensively in ESOL teaching.

Professor Susan Ledger

Susan is the Head of School/Dean of Education at the University of Newcastle. She is an advocate for the teaching profession and explores education policies and practices related to teaching, and preparing to teach, diverse students and diverse contexts. Her broad experiential base teaching and researching in rural, remote, and international contexts informs her belief that all educators need to be confident teachers of literacy, literature and language acquisition.

EXECUTIVE SUMMARY

This project reports on a systematic literature review of published research on the topic of using decodable and levelled reading books¹ in primary school classrooms. Articles from academic, peer-reviewed journals published between 2000 and 2021 were selected for analysis. The central focus of the selected articles was on teaching reading or reading instruction using decodable and/or levelled reading books. The report that is written from the larger project conducted, aims to provide evidence to support education decision makers—particularly classroom teachers—to make informed decisions to meet the needs of students within a mainstream primary school setting. The systematic literature review focused on three critical research questions:

1. What does current published research (2000–2021) indicate on the topic of using decodable and/or levelled reading books in primary school classrooms?
2. In a primary school classroom does the use of a reading program/s with decodable or levelled reading books increase the reading outcomes (reading proficiency) of students?
3. Is there empirical research evidence to recommend the use of decodable and/or levelled and/or another type of book over others in classroom reading instruction?

Through the systematic literature review, from a total of 1865 articles initially identified, 91 studies were deemed relevant (see Section 8: Methodology for the criteria for inclusion and exclusion) and included for full-text screening. After closer examination, including a full-text reading of each of the identified articles, 19 studies were considered relevant for the project.

No longitudinal studies were identified that provide a randomised controlled trial of various approaches to teaching reading in the early primary years in Australia. Rather, a series of small-scale research projects have been conducted that *do not* focus on

¹ This report uses the term ‘decodable and levelled reading books’. Elsewhere, they are also referred to as *texts* and *readers*. Here, to avoid confusion with other terms we use throughout the report, *reading books* is consistently used (unless in a direct quote) when referring to those books used in classrooms to support students’ reading proficiency.

reading proficiency using comparison groups, thus making the case for evidence-based reading practice, somewhat wanting.

Key findings in this report include:

- The influence of reading books is diverse (Mesmer, 2010) and consideration should be given to how they develop context, interest, and engagement for students.
- Individual students at various stages and ages should be offered a variety of texts, including reading books, for their reading development (Brown, 2000; Jenkins et al., 2003; Mesmer, 1999, Beverley et al., 2009).
- Quality instructional reading materials for students need to include multiple features and components to accommodate the complexity of reading and to provide opportunities for the explicit teaching of specific components (Mesmer, 2010; Cheatham et al, 2014).
- Exclusive use of one type of book has a detrimental effect. Rather, carefully selected and varied reading materials leads to more positive outcomes and attitudes to reading (Mesmer, 2005; Wood, 2005)
- Students need multiple experiences, strategies, and support when learning to read (Ankrum, 2020).
- Teacher knowledge and understanding about instructional approaches and the components that make up the complex process of reading cannot be underestimated (Hofman & Pearson, 2015; Price-Mohr & Price, 2020).
- The use of reading books or associated models or programs of beginning reading instruction are impacted by the classroom integration and pedagogical approaches used by the teacher (Hofman & Pearson, 2015; Price-Mohr & Price, 2020).
- The teacher's expertise and judgment are just as important as the model, approach, or text (Rightmyer et al., 2006).

CONTENTS

The review team	3
Executive summary	4
Introduction	7
Overview and background	9
Influence of international reports and assessment	10
Literacy teaching and learning contexts	12
Reading approaches	15
Using books to teach reading in schools	17
Defining key terms	18
Methodology	20
Characteristics of the included random controlled trials and quasi-experimental studies	26
Key results and discussion of results	36
Conclusions	42
Acknowledgments	47
Glossary	48
Reference list	53

INTRODUCTION

As schooling has become progressively more complex, so have the curriculum elements which make up a student's education. Increasingly, there are greater demands on teachers to not only fit more content into an already crowded curriculum, but to also incorporate the latest information communication technologies and learning innovations. Literacy is an area of the curriculum which has not been exempt from these complexities but rather, along with numeracy (Sellars, 2017), has received the majority of attention and criticism from those seeking to identify failures in the schooling system. The very nature of the literacy debate has been an intensely contested issue for many decades, and in Australia the debate was incorporated into the Culture Wars of the early 2000s (see, for example, Sawyer, 2004; Maiden & Harris, 2005; Wilson, 2005; Wiltshire, 2006; Ferrari, 2007; Bantick, 2006; Livingstone, 2006; Lawrence 2006). There are many contemporary debates about the nature of literacy and the place it has in the modern classroom, often discussed in binary terms, including the *science of reading*, *phonics*, *whole language*, *balanced*, and *genre* approaches. Stakeholders such as media commentators, government education departments, schools, educators, parents, and students have, or want to have, influence over what is taught in the school classroom. As part of larger debates about literacy teaching, effective reading instruction has been debated and re-evaluated for decades.

One of the fallouts associated with ongoing literacy debates is the immense pressure for educators to 'choose a side' and to express exclusive support for one position, such as *whole language* (often associated with the *balanced literacy approach*) or the use of *phonics* (whether synthetic, analogy, analytic, or embedded), in many cases with the various perspectives in opposition to each other. Anecdotally, school-based decisions are made in primary schools to teach literacy using only one specified program (often commercially available) or approach. Sometimes teachers are mandated to use this school-wide approach at the exclusion of any other approach. Often these literacy choices involve purchasing reading books to support the chosen program. Subsequently, when schools have made an economic investment in resources, there can be pressure to use them to ensure that money is not wasted.

There is also pressure on teachers across the same year level to teach using the same approach and material resources.

As teachers are frequently expected to apply the best evidence-based practice to guide their pedagogical decision making, it is critically important that quality empirical research is evaluated to guide this decision making. It is therefore essential that research outputs are clear and accessible for teachers, principals, and other education stakeholders, to inform the implementation of literacy programs and practices in the school setting. The following critical research questions guided the systematic literature review that informed this report:

1. What does current published research (2000-2021) indicate on the topic of using decodable and/or levelled reading books in primary school classrooms?
2. In a primary school classroom does the use of a reading program/s with decodable or levelled or reading books increase the reading outcomes (reading proficiency) of students?
3. Is there empirical research evidence to recommend the use of decodable and/or levelled and/or another type of book over others in classroom reading instruction?

This report provides a review of articles about the use of levelled and decodable reading books. Articles from academic, peer-reviewed journals published between 2000 and 2021 (June) were selected for analysis. The central focus of the selected articles was on teaching reading or reading instruction using decodable and/or levelled reading books. This report, the outcome of a systematic review, provides an evaluation of the quality of research evidence in reading instruction with decodable and/or levelled reading books in the classroom.

This report analyses research on the benefits of decodable and/or levelled reading books and in turn provides evidence to support education decision makers—particularly classroom teachers—to make informed decisions to meet the reading proficiency needs of students within mainstream primary school setting

OVERVIEW AND BACKGROUND

Reflecting on the history and progress of literacy instruction over time provides context to understand the current state of ambiguity and debate around the teaching of reading (Barry, 2008). History has shown that the teaching of reading is a complex process that relies on multiple elements to work together cooperatively. The process of reading has been likened to that of operating a car (Adams, 1990) in which we rely on many systems and parts operating simultaneously to get to our destination. The components of reading are described as “not discrete. We cannot proceed by completing each individual sub-system and then fastening it to one another. Rather, the parts of the reading system must grow together. They must grow to one another and from one another” (Adams et al., 1990, pp. 20–21). This metaphor points to the view that reading is multifaceted (Clay, 1991). This idea is extended by Compton-Lilly et al. (2020) through their discussion of the multidimensionality of reading that guides young students’ trajectories and meets their unique needs as they grow into competent readers. Castles et al. (2018) discuss this in relation to the *Simple View of Reading* as posited by Gough and Tunmer (1986) and Hoover and Gough (1990) and confirm this concept by asserting that no single model or method can function in isolation, writing that a “child who can decode print but cannot comprehend, is not reading; likewise, regardless of the level of linguistic comprehension, reading cannot happen without decoding.” (p. 27)

Another approach, the *Reading Rope Model* (Scarborough, 2001) also supports the interconnected components of reading as a framework combining linguistic/language comprehension to include background knowledge, vocabulary knowledge, language structures, verbal reasoning, and literacy knowledge. The importance of phonological awareness, decoding and sight word recognition are noted as key factors in this framework (Scarborough, 2001). This concept is further supported in reports such as the *National Reading Panel* (2000), Rose (2006), and Rowe (2005) who contend that skills in word recognition and language comprehension are of equal importance to develop a skilled reader. Being skilled in one element does not offset the need for the skill in another (Scarborough, 2001). Studies over decades have shown that overusing or removing a particular element or strategy leaves learners vulnerable to disengagement in the process. Despite research consensus on the clear identification

of the complex and interrelated components of the nature of reading, the reading wars persist (Compton-Lilly, Mitra, Guay & Spence, 2020), and debates continue over how to best support students learning to read.

INFLUENCE OF INTERNATIONAL REPORTS AND ASSESSMENT

Data from the National Inquiry into the Teaching of Literacy (Rowe, 2005) highlights areas of concern on the teaching of literacy in schools for governments, policy makers, educators, and the general public. Concern over literacy achievement is also noticeable across other jurisdictions, evidenced by the large number of reports commissioned over the past 20 years. These include: the National Reading Panel, USA (2000); What Works Clearinghouse, USA; No Child Left Behind, USA (2001); The Science of Reading; Closing the Gap (Literacy and Numeracy), Australia (2008); National Inquiry into the Teaching of Literacy (The Rowe Report), Australia (2005); The Independent Review of the Teaching of Early Reading (The Rose Report), UK (2006); and more recently, the National Early Language and Literacy Strategy, Australia, (2021).

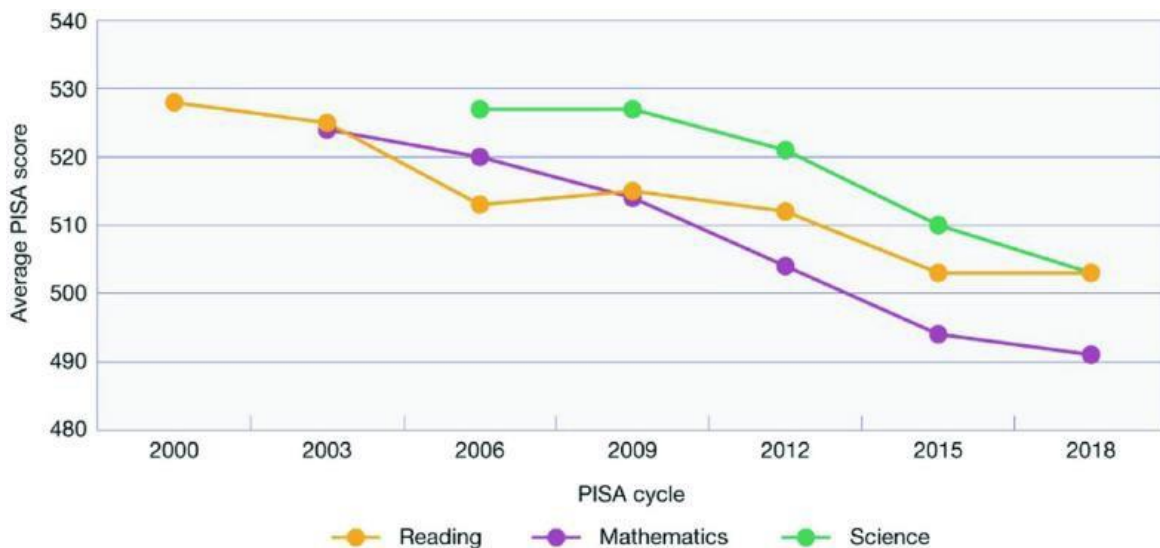
Two decades ago, Australian students scored highly on the Programme for International Student Assessment (PISA), and evidence-based approaches to teaching of reading in schools were being developed, including First Steps Literacy (Government of Western Australia, n.d); Teaching Handwriting, Reading and Spelling Strategies (THRASS Institute, 2019); Reading Recovery (Reading Recovery Australia, 2021); and Mt Gravatt Development Language Reading Program (Hart, 1977). However, over the last decade, literacy education in Australia has been more influenced by the UK, and systematic synthetic phonics (SSP) programs have become increasingly dominant in the teaching of phonics in Australian schools. Key findings from the Independent Review of Early Reading (Rose, 2006) from a longitudinal study conducted in Clackmannanshire, Scotland has become a focus for many Australian researchers. Recommendations from The Rose Report, as it is colloquially known, were instrumental in the introduction of mandated synthetic phonics teaching in the UK, and the exclusion of other phonics teaching methods. In parallel with the UK, special interest groups in Australia have pressured education ministers and policy

makers to exclusively incorporate SSP and decodable reading books into primary school classrooms.

However, as the research that informed this report indicates, there have been no longitudinal studies conducted in Australia that provide a randomised controlled trial of various approaches to teaching reading in the early primary years. Rather, a series of small-scale research projects have been conducted that *do not* focus on reading proficiency using comparison groups. Australia's PISA results have also fallen during this time of policy borrowing (that is, taking ideas from other education jurisdictions and applying them to another context, in this case Australian primary schools), suggesting that current literacy program implementation is not ideal for Australian school students.

Recent research suggests that literacy levels in Australia are declining—see for example PISA results from 2000 compared to results in 2018 (Figure 1, OECD, 2019). The mean data shows a steady decline since Australia first participated in the programme. In 2000, Australian students achieved a mean score of 528 points in reading literacy, which was higher than the OECD average of 500 (Thomson, 2021), and the 2018 results showed a mean score in reading literacy of 503 points – a decline which is indicative of about three-quarters of a school year in terms of literacy skills (Thomson, 2021). Results in mathematical and scientific literacy have also followed this downward trend.

Figure 1 – Source OECD, 2019



LITERACY TEACHING AND LEARNING CONTEXTS

The role of decodable and levelled reading books needs to be considered in terms of how we teach reading and the purpose of teaching reading. The Organisation for Economic Co-operation and Development (OECD) considers reading skills as key to educational success. According to the agreed OECD PISA definition, “reading literacy is understanding, using and reflecting on written texts, in order to achieve one’s goals, to develop one’s knowledge and potential, and to participate in society” (OECD, 2000, p. 3). The OECD analysed data from the International Adult Literacy Survey (IALS) and its own Program for International Student Assessment (PISA) and found students with the lowest PISA reading scores are at risk in adulthood, and face increased chances of unemployment, reduced prospects of having a well-paid job, and a limited likelihood of engaging in future learning (OECD, n.d).

As expressed by the OECD the importance of understanding, using, and reflecting on texts (including reading books) is fundamental in the design of reading programs. However, the development of reading for enjoyment or leisure is equally as important, with research suggesting that time spent reading for pleasure could be a more important indicator of a child’s future success than is their family’s socio-economic status (Awah, 2020). Educators have long discussed the importance of *enjoyment* of learning as a key to success; for example, as Fischetti states, “we want to prepare children to be successful in their futures and to do that they need knowledge, skills and dispositions to be passionate, vibrant, dynamic, curious, open-minded, engaged (and literate and numerate) participants in their own journeys. We can’t assembly-line assess that” (Fischetti, 2016, para. 12).

The interaction between text and reader is central to any approach to teaching reading. The Four Resources Model relates to the reader becoming a text decoder, text participant, text user, and text analyst (Freebody & Luke, 1990). Read-aloud books, shared reading, and guided reading remain core pedagogical strategies in classrooms around the globe. A ‘read aloud’ (where the teacher reads a text aloud to students) encourages fluency and models the process of reading. Guided reading strategies introduce the student to reading using levelled reading books, conversation and thinking tasks. Shared reading incorporates modelling, questions, and responses.

Whatever approach or strategy is employed, students need to learn how to make meaning of or comprehend texts. This is done primarily through understanding of pragmatics, semantics, syntactic, and graphophonic knowledge (Government of Western Australia, n.d.).

The Canadian National Institute for Literacy recently presented building blocks for teaching students to read, which include instruction in phonemic awareness, phonics, fluency, vocabulary, and text comprehension. Although surrounded by theories, approaches, instructional strategies, and programs for the teaching of reading, teachers are distracted by ongoing choice and challenges. Adding to this is the array of decodable and levelled reading books to choose from.

Literacy teaching and learning in Australia

The three-dimensional Australian curriculum (key learning areas, general capabilities, and cross-curriculum priorities) underpin the language, literature, and literacy strands of the Australian Curriculum: English. Reading is addressed within the key learning area of *English* and the general capability of *literacy*. Teachers from early childhood, primary, and secondary are responsible for implementing these three dimensions with an understanding that learning to read has many facets that are needed to make a whole and/or achieve the outcome of literate students exiting our schools (Ledger & Merga, 2018).

Each state and territory has adapted the Australian Curriculum to suit their needs (some have incorporated the Curriculum into their own syllabus while others use it as a syllabus in and of itself). However, the elements of the three strand English program remain, which also includes the teaching of speaking and listening, writing, reading, and viewing. Reading includes the sub-elements of phonological awareness, fluency, phonemic knowledge, word recognition, and understanding texts. These remain essential elements recognised by all states and territories, with importance often placed on the teaching of reading, literacy learning progression and understanding print, aural and visual texts, as described by ACARA:

Understanding texts describes how a student becomes increasingly proficient in decoding, using, interacting with, analysing and evaluating

texts to build meaning. Texts include components of print, image, sound, animated movements and symbolic representations. This sub-element is organised into three subheadings: comprehension, processes and vocabulary. (ACARA, n.d., para 1)

English as defined by the Australian Curriculum, Assessment and Reporting Authority (ACARA)

Literacy as a general capability is organised into two overarching processes: comprehending texts through listening, reading and viewing; and composing texts through speaking, writing and creating. The following areas of knowledge apply to both processes: text knowledge, grammar knowledge, word knowledge and visual knowledge.

Literacy development involves conscious attention and focused learning. It involves skills and knowledge that need guidance, time and support to develop. These skills include:

- the ability to decode and encode from sound (phoneme) to written systems (graphemes);
- the learning of grammatical, orthographic and textual conventions; and
- the development of semantic, pragmatic and interpretative, critical and reflective literacy skills.

In 2015, The Australian Literacy Educators' Association (ALEA) released Literacy in 21st Century Australia: ALEA Declaration, drawing together best practices for teaching literacy in the classroom. It outlined how students progress from learning *how* to read when they first enter (five to eight years of age) in fluency, comprehension; and learning vocabulary through processes of modelled, shared, guided and independent reading (Riddle & Honan, 2016). ALEA declared:

No one method of reading/writing instruction will ever meet the needs of all children at all times. Therefore, educators need to be discerning practitioners as they draw on research that is contemporary, valid and rigorously conducted to inform their practice. (ALEA, 2015, p. 2)

The Primary English Teaching Association Australia (PETAA), the peak English teachers' group, advocate, support, and champion teachers of English and literacies, and the importance of evidence-based pedagogies. They also highlight that:

teachers develop expertise as they respond to the needs of the students in their classroom; different students will need support in different aspects of reading and as teachers respond to these needs, they will add to their repertoire of teaching strategies. Teachers also pay attention to building comprehension skills, fluency and vocabulary knowledge in developing readers. (PETAA, 2021a, p. 1)

Leaders and experts in English and literacy support the structure of the Australian curriculum, but debate exists on the implementation of approach, strategies and texts:

Text selection is an important and complex aspect of teachers' work; there are many factors which they must consider before making their selection. Factors include: purposes for reading; student motivation, interest and enjoyment; community, parent and school expectations; student diversity and planning for increasing complexity and challenge. (PETAA, 2021b, p. 1)

READING APPROACHES

A recent influx of 'how to teach reading' approaches has impacted literacy reform in Australia. Driven by a need to 'improve' literacy levels, a mix of newly constructed, packaged reading initiatives and commercial products, programs, and ideals developed by speech pathologists, psychologists and special needs experts have surfaced. More recently in Australia the phrase *the science of reading* has entered the lexicon of teaching students to read. On the one hand, the science of reading refers to the body of research evidence on teaching reading to young readers. The science of reading in its authentic form is the effective integration of all the essential strategies needed to master reading. The science of reading states unequivocally that reading is a complex process that from the very beginning requires effective instruction from educators who are trained to support students. In this case, the science of reading

does not align to one specific program, reading strategy, nor instructional reading text. This term should not be confused with groups labelling themselves as Science of Reading (SoR) or with science-based evidence. The science of reading's supporters argue that it is *the* 'scientific' (and by default, correct and researched) method to teach reading, describing it as "representing the accumulated knowledge about reading, reading development, and best practices for reading instruction obtained by the use of the scientific method" (Petscher et al, 2020, p. 268).

Similar to the debates that led to the Culture Wars, the science of reading topic has also attracted controversy and created a divide amongst educators regarding the interpretation of what constitutes science of reading (Castles et al., 2018) and the practical implications of this evidence for the teaching of reading (Solari et al, 2020). There are many interpretations that make up different approaches to reading (Yaden et al., 2021). A current approach underpinned by the science of reading draws upon *The Simple View of Reading* (SVR) (Gough & Tunmer, 1986) as a basis for a pronounced focus on decoding and word reading as the most essential element in early reading instruction. This view differs from the intended aim of the SVR as it advocates reading comprehension as the product of decoding and listening comprehension, and it gives equal importance to both components (Cervetti et al., 2020).

This interpretation points to a concern that evidence can misrepresented and/or oversimplified to endorse practices that are loosely coupled to the science of reading (Cervetti et al., 2020, Seidenberg et al., 2021). This cautions the field to be wary of unjustified interpretations and inferences using the science of reading as sole justification for an approach. Additionally, media portrayals of the science of reading sometimes perpetuate misinterpretations, which then impacts on education decision makers (Dewitz & Graves, 2021; MacPhee et al., 2021). Yaden et al. (2021, p. S120) express caution over this narrow view of the science of reading and call for educators in the field to consider:

- (1) too heavy a reliance on a narrow conception of science claimed to be authoritative and monolithic,
- (2) too little accounting for environmental factors that complicate the idea that the brain functions identically across the whole of the human population,
- (3) an exclusive

view that experimental designs and replicability are the gold standard of scientific research when other approaches have generated many useful insights, and (4) dismissal of all other conceptions of reading as unscientific and, therefore, of marginal value in generating knowledge about reading and how to teach it.

This section on the science of reading is not contesting or discounting the rigorous scientific evidence that forms the approach, but rather points to the need to assess the quality and merit of such evidence and to consider how each study can be (mis)interpreted and utilised to support the argument being made about the teaching of reading. It also seeks to affirm the need for additional translational research to link the evidence to classroom practice and pedagogical approaches as recommended by Seidenberg et al. (2021).

USING BOOKS TO TEACH READING IN SCHOOLS

Instructional reading materials, specifically instructional reading books, have been elevated to the centre of literacy debates through questions about pedagogical approaches (including *phonics*), with decodable and levelled reading books often pitted against each other. The momentum for undertaking the research outlined in this report was to ‘fact find’ on whether quality research evidence provides definitive and concrete information to support teachers and schools in making an informed decision about which reading books provide the best conditions to foster effective reading instruction. While there has been extensive research into teaching methods, there has been less research on text usage. Empirical – especially quantitative – research on using decodable and/or levelled reading books with beginning students as the focus is sparse in comparison to research on other areas of reading instruction. This could be accounted for by the complexity involved in segregating texts (including reading books) and/or methods used from the pedagogical approach of the teacher, along with other variables that inevitably arise in the field of education. The cost of conducting randomised controlled trials and the ethical complexities when conducting research with primary school students cannot be ignored either. While there has been extensive research into teaching methods or approaches, teacher beliefs about reading, and

pedagogical preferences, empirical research on text selection has received far less attention. It is important to note that the definition of *decodable reading books* and *levelled reading books* often differ from the textual features that characterise these texts (Mesmer, 2010). For the purposes of this report, and in an endeavour to provide clarity for this review and for teachers and other stakeholders, key focus areas are defined below.

DEFINING KEY TERMS

This section defines and describes the features of two types of texts, decodable and levelled reading books.

Decodable reading books

Decodable reading books are instructional texts that are carefully designed and subsequently set according to a specific phonics teaching sequence. Decodables direct their focus to a code-based approach to reading (Pogorzelski et al., 2021) that supports a linear method aligned to Systematic Synthetic Phonics (SSP). Decodability can be defined as drawing on two key elements:

- The first element is the proportion of words with regular phoneme to grapheme correspondences.
- The second element is the amount of letter/sound relationships that have been taught (Mesmer, 2010).

Decodable reading books are designed to provide opportunities for beginning readers (students) to apply phonological skills progressively as texts are composed at varying degrees of decodable words based on the systematic progression of taught skills (Cheatham & Allor, 2012; Buckingham, 2018). Decodable reading books emphasise the repetition of taught phonics patterns to build orthographic knowledge and fluency. Most beginning decodable reading books focus only on a one-letter to one-sound code with words written to match. They follow the approach that students learn to “convert a string of letters (our written code) into sounds before blending them to produce a spoken word” (Pogorzelski & Wheldall, 2018, para. 8). Decodable reading books focus

on phonic coding; comprehension and vocabulary development are secondary considerations. As more sophisticated orthography becomes necessary to incorporate the naturalness of language, the structure of decodable reading books changes. At this stage, the decodable reading book mirrors the characteristics of a levelled reading book.

Critics of decodable reading books assert that they have restrictive vocabulary, are syntactically awkward, do not allow for the naturalness of language, provide minimal storyline or narrative, and provide less engaging context for students. There is also inconclusive data about the optimal level of decodability for different reading books at different stages which makes focusing on only decodability as the key criterion to be a cause for concern (Mesmer, 2010).

Levelled reading books

Levelled reading books are decodable instructional books that are carefully designed and composed according to specific literacy criteria. Levelled reading books are composed as meaningful whole texts with consideration given to the naturalness of language (Fountas & Pinnell, 2013). The reading books increase in difficulty as the student progresses, having acquired the literacy skills scaffolded at each level or stage. Levelled reading books rely on the synergy of multiple elements in their composition to allow for explicit teaching. They are composed using a gradient of identified difficulty to include multiple, essential strategies for reading, including the introduction, beginning and on-going scaffolds of orthography, decoding at individual word level (phonics), fluency, high frequency words, vocabulary, and comprehension. These reading books can then be analysed to support choices for differentiated reading instruction (Fountas and Pinnell, 2013; Hiebert, 2012; Mesmer, 2010; Ankrum, 2021).

Beginner levels combine decodable, predictable, and repetitive text in the context of topics that are age appropriate and familiar to young students (Cunningham et al., 2005). The inclusion of illustrations provides additional support and instructional scaffold through building context (Fountas & Pinnell, 2013; Cunningham et al., 2005). The multiple modes provided in levelled reading books include a variety of cues for students to make meaning from and to decode unfamiliar words underpinning the

written text. For levelled reading books, multiple criteria are important characteristics for developing students to be successful readers (Cheatham and Allor, 2012).

Levelled reading books support the theory that students learn language and subsequently learn to read through making meaning using the three-cueing system of visual (graphophonic), meaning making (semantic context), and structure (syntactic context; see for example, Clay, 1991). Students need multiple strategies to consider whether a word looks right, sounds right and makes sense (Fountas & Pinnell, 2006).

With the push to teach phonics through decodable reading books in recent years, levelled reading books have been criticised. Detractors have claimed that they do not require students to attend to the phonics detail in printed words nor provide opportunities for explicit direct phonics instruction; that they encourage students to 'guess' words based on multiple scaffolds or to overuse the context of the reading books, such as illustrations; and that by allocating students to a set level, students are restricted from reading more complex texts.

METHODOLOGY

This section provides detail on the methodological approach used to identify, source, and analyse the research that formed the systematic literature review data.

Search strategy

The search strategy restricted studies to:

- double blind peer reviewed journals;²
- studies written in and about the English language; and
- published between January 2000 and June 2021.

Databases

To locate relevant published research that responded to the critical research questions identified above, a systematic search of electronic databases was conducted.

² Double-blind peer review process means that both the author and reviewer are not known to the other. This is to ensure that an objective review takes place. In academic journals, an expert in the research area is asked to review a manuscript and to provide feedback to the (anonymised) author including advice to the Journal editor on whether to accept or reject the manuscript for publication.

- Databases included: Academic Search Ultimate, Education Research Complete, Education Database, SCOPUS, ERIC, Linguist and Language Behaviour Abstracts, Literature Online, Proquest Central, A + Education and Emerald.

Keywords

Given the wide international interest in literacy research across multiple types of educational jurisdictions, multiple terms were used in keyword searches with an expansive search thread to ensure a comprehensive investigation. In collaboration with research librarians the search strategy included:

- “Level* read*” OR “level* text*” OR “level* book*” OR “Levell* read*” OR “levell* text*” OR “levell* book*” OR “guide* text*” OR “guide* read*” OR “guide* book*” OR “predicable* text*” OR “predicable* read*” OR “predictable book*” OR “instructional book*” OR “instructional text*” OR “instructional* read*” AND “decod* book*” OR “decod* read*” OR “decod* text*” OR “ phonic* book*” OR “phonic text*” OR “phonic read*” AND “primary school” OR “elementary school” OR “school” AND “reading instruction*”

Inclusion and exclusion criteria

Cochrane Review Methods (Chandler & Hopewell, 2013; Higgins et al., 2019) were employed to develop the protocol for the inclusion and exclusion criteria prior to screening articles that would be included in the systematic literature review (Higgins et al., 2019). Studies that were determined suitable for inclusion met the following criteria:

- a) L1 English language instruction;
- b) Mainstream students;
- c) Primary (or equivalent, such as elementary) school students;
- d) Directly addressed use of levelled reading books, instructional, predictable, guided reading books, or other equivalent; *and/or* decodables or phonic reading books and their use or impact on reading instruction in primary classrooms;
- e) Included a measurable primary outcome of reading achievement;
- f) Focus was the reading book and not the teacher’s pedagogy, whereby the reading book itself influenced the outcome; and

- g) Study design was experimental (randomised controlled trial (RCT) or quasi-experimental (cohort, case-control, pre-post or repeated measures)).

Meta-analysis and systematic reviews were initially included to search within the text and reference list for additional studies that may have been missed in the search strategy that could be relevant for inclusion. Studies were excluded from the systematic literature review if they fell into the following categories:

- a) Conducted in a high/secondary school (or equivalent), tertiary, tutoring, pre-school, or early childhood context;
- b) Focused on specific learning difficulties and/or disabilities such as autism or dyslexia; and
- c) Cross-sectional or descriptive studies.

Initial screening

For transparency in citations from the studies identified in the initial search, articles were imported into Covidence (a systematic review management software program that enables and acts as a screening and data extraction tool) for further selection to take place.

Once articles had been collected and incorporated into Covidence, the initial stage of the research involved removing duplicates and assessing the study relevance of the articles based on titles and abstract only. Studies were screened independently by two researchers based on the above-mentioned inclusion and exclusion criteria. Using this method, from a total of 3025 articles, 91 studies were deemed relevant and included for further screening.

Full-text screening

The full text of the 91 identified articles was collected and assessed for eligibility in Covidence using the identified inclusion and exclusion criteria (a reason for exclusion needed to be identified). Independently, two researchers for each article—from a pool of four—read the articles to ascertain their relevance to the study. For studies in which the two researchers did not agree, a third researcher undertook adjudication through the Covidence data management system. To ensure selection integrity, the adjudicator's reason for exclusion needed to match the decision of the researcher who

first excluded the article, otherwise it was returned by Covidence’s automation to the database for re-selection or exclusion. At this stage, meta-analyses were excluded. Subsequently, meta-analysis³ studies were reviewed to identify additional articles for identification of possible additional studies not captured in the database search (see Figure 2). At the completion of this stage, 19 studies were considered relevant.

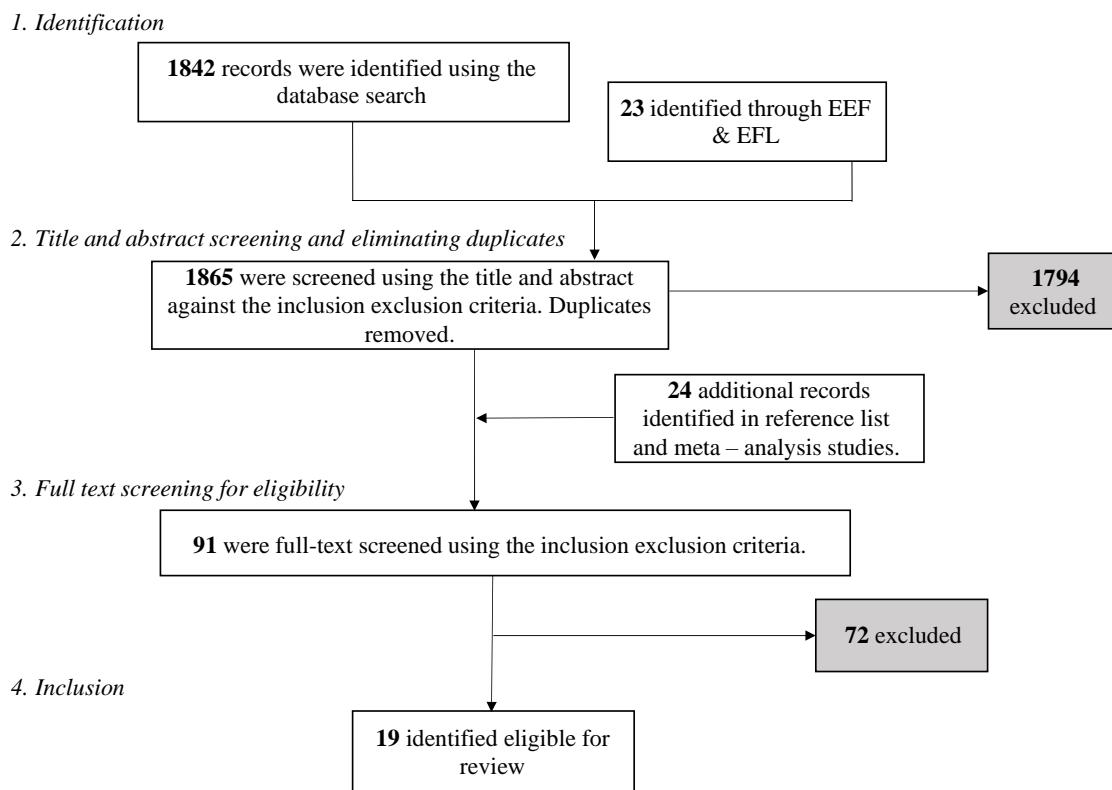


Figure 2: Systematic Literature Review Flowchart

Data extraction

The four subject level experts who undertook full-text screening were responsible for the extraction of data from each of the included studies. Examples of data extracted from each article included:

- Country of study
- Study design (RCT, Quasi, Pre-post)
- Type of comparison (e.g., control vs intervention)

³ A meta-analysis refers to a study that combines results of other research conducted on the same topic. It performs statistical analysis on research results and does not include new or original empirical/field work research.

- Sample details (sample size, reading proficiency, grade, age)
- Intervention details (time [weeks], volume [hours], primary text type [levelled, decodable, mixed], pedagogical focus [levelled, decodable or mixed instruction])
- Comparison details (text type and pedagogy)

Outcomes (assessment used and outcome category [Fluency, Comprehension, Phonemic Awareness, Word Recognition, Decoding skills, Accuracy, Spelling, Vocabulary, Reading rate and Silent Reading Fluency]) (refer to glossary for definitions of key terms)

Effect size was calculated for randomised controlled trials and quasi-experimental studies only. Calculation utilized the Cohen's *d* formula:

$$d = (M1 - M2)/SD_c$$

where M1 and M2 are the change scores (follow-up – baseline) for the intervention and control/comparison conditions respectively, and SD_c is the pooled or common standard deviation, calculated as the weighted average of the standard deviation of the two groups (across pre and post time points). Effect sizes were interpreted using the benchmarks defined by Cohen (1988), with a small effect size defined as $d = 0.2$, a medium effect size as $d = 0.5$, and a large effect size as $d = 0.8$.

Where studies included outcomes at only one time point (e.g., a follow-up outcome using a different baseline outcome as a model baseline covariate), an effect size was not calculated. If not enough information was included to calculate an effect size (e.g., only change scores are reported), these outcomes were tagged as NEI (Not Enough Information) in all tables. If a study included multiple assessments within a single outcome category (e.g., two comprehension tests), these effect sizes were averaged for inclusion in the summary tables.

While the quantitative data extraction was being undertaken, qualitative description and analysis was also taking place using a close reading of text by two of the researchers who had also been involved in the identification and selection stages of the research. As the data consisted of a large quantity of articles, a summative content analysis was used. This method allows for a combined quantitative and qualitative

analysis, which is helpful when large data is analysed. The choice of methodology provides a general image of patterns and connections, the outcome may be of interest as a first point to understand the literacy research being undertaken involving primary school students.

Risk of bias assessment

Articles were reviewed to assess their methodological quality using the Risk of Bias tools from Cochrane Methods (Cochrane Collaboration, 2021). The Risk of Bias 2 (ROB 2) tool was used to assess randomised controlled trials using the parallel trial and cluster trial tools for individually and cluster (class/school) randomised trials respectively. The Risk of Bias in Non-randomized Studies – of Interventions (ROBINS-I) assessment tool was used to assess the quasi-experimental trials. Two researchers who were not involved in the earlier selection stages undertook the methodological analysis.

Results overview

From the 19 studies included for review, there were eight (8) randomised controlled trials⁴ (42%), three (3) quasi-experimental⁵ (16%) and eight (8) pre-post⁶ (42%). All eligible studies were included in the narrative review to give coverage of the field of reading research in relation to the inclusion criteria; however, only data from the randomised controlled trials (n = 8) and quasi-experimental trials (n = 3) were included in the quantitative synthesis of results (e.g., data extraction, risk of bias assessment, and calculation of effect sizes). This was undertaken due to the concerns about the external validity of findings from studies that do not include a comparison group to give an indication of intervention effects in relation to ‘normal’ (regular, expected) growth in student attainment across time (for example, a school year). Pre-post studies (n = 8)

4 A randomised controlled trial (RCT) is an experimental form of impact evaluation in which the population receiving the programme or policy intervention is chosen at random from the eligible population, and a control group is also chosen at random from the same eligible population. It tests the extent to which specific, planned impacts are being achieved. The distinguishing feature of an RCT is the random assignment of units (e.g. people, schools, villages, etc.) to the intervention or control groups.” (UNICEF, n.d. para. 1)

5 A quasi-experimental study aims “to evaluate interventions but that do not use randomization... quasi-experiments aim to demonstrate causality between an intervention and an outcome [and]...can use both preintervention and postintervention measurements as well as nonrandomly selected control groups.” (Harris et al., 2005, p. 17)

6 A pre-post study “measures the occurrence of an outcome before and again after a particular intervention is implemented.” (Thiese, 2014, p. 205)

provide useful insights that need to be taken into consideration (Yaden et al.,2021). As such, these studies were examined for their insights and included within the qualitative discussions of this report.

CHARACTERISTICS OF THE INCLUDED RANDOM CONTROLLED TRIALS AND QUASI-EXPERIMENTAL STUDIES

Sampling characteristics

Of the included studies on which data extraction was undertaken (n = 11), nine (9) were conducted in the United States and the remaining two (2) in the United Kingdom (Table 1). Sampled grade levels ranged from K to grade-6, with six (6) studies of K to grade-1 students and five (5) with grade-2 or above. Most of the interventions (n = 8, 73%) were directed at students identified as having lower reading proficiency levels. The average sample was 190 students, ranging from 24 and 1101 participants. A sample of approximately 210 students is required to detect a statistically significant moderate effect size ($d = 0.5$ – which we consider a realistic aim for studies using proximal outcomes⁷). This sample grows when participants are not individually randomised/allocated to an intervention or comparison condition due to the influence of clustering at the class level. Only three (3) studies had a sample size over 200 participants.

Intervention characteristics

Across the 11 studies, there were 15 interventions tested, with four (4) studies testing several interventions against a comparison group simultaneously (Table 2). The average length of interventions was 42 hours with the smallest duration being 5 hours and the longest 117 hours. Of the 15 interventions, nine (9) focused on participants engaging with levelled text, either solely or the vast majority. Five (5) interventions focused on engagement with decodable text in the majority, and one (1) specifically used a combination of levelled and decodable text. The majority (n = 12, 80%) of interventions utilised a mixed pedagogical strategy (i.e. instruction focused on skills of decoding [with either decodable or levelled text] whilst also employing comprehension focused strategies [using levelled or authentic text]). Several interventions (n = 2, 13%)

⁷ A proximal outcome is one that can be seen to change in the short-term and is usually detailed and specific (e.g., testing fluency before and after a fluency intervention).

Table 1. Characteristics of included experimental and quasi-experimental studies

Study	Study Design (Research Type) / Specific population / Participants: Grade; Age range, Mean age	Sample size: Total (Control; Intervention/s) / Intervention volume: Weeks (hours) / Intervention: Control/ intervention/s	Focus of intervention texts	Outcomes (measure)	Outcomes - Levelled text	Outcomes - Decodable text
Cheatham et al (2014); USA	RCT (CON/INT) / All students / 2; 6-7; 7.2	62 (28; 34) / 10wk (5hr) / Read children's books without intentionally imbedded scaffolds vs Mixed: Levelled readers instructed using a phonic decoding strategy	Authentic; Levelled	Decoding skills (PDE OF TOWRE), Word recognition (WRMTR - SUBTEST SWE)	Decoding skills: 0.09 Word recognition: 0.13	
Denton et al (2014); USA	RCT (CON/INT) / Low proficiency / 1; NI; NI	218 (53; 59; 50) / 23 - 25wk (46hr) / Business as usual - no small group intervention vs Mixed: Guided reading with small phonics component vs Mixed: Explicit phonics instruction with comprehension instruction using authentic text	Levelled; Decodable	Decoding skills (WJ III - Letter word ID), Decoding skills (WJ III - Word Attack), Comprehension (WJ III - Passage Comprehension), Comprehension (GMRC), Comprehension (Passage Comprehension), Fluency (TPRI), Silent Reading Fluency (Test of Silent Reading Efficiency)	Fluency: 0.22 Comprehension: 0.09 Decoding skills: 0.16 Silent Read Fluency: 0.20	Fluency: 0.54 Comprehension: 0.38 Decoding skills: 0.39 Silent Read Fluency: 0.42

Study	Study Design (Research Type) / Specific population / Participants: Grade; Age range, Mean age	Sample size: Total (Control; Intervention/s) / Intervention volume: Weeks (hours) / Intervention: Control/ intervention/s	Focus of intervention texts	Outcomes (measure)	Outcomes - Levelled text	Outcomes - Decodable text
Jenkins et al (2004); USA	RCT (CON/INT) / Low proficiency / 1; 6-7; 6.5	121 (20; 39; 40) / 25wk (50hr) / NI vs Mixed: Phonics instruction and guided reading with more decodable text (Bob Books; Get Ready, Get Set, Read! Wrights Skills set) vs Phonics instruction and guided reading with less decodable (levelled) text (Story Box; Vision Series; Sunshine Books)	Decodable; Levelled	Decoding skills (WRMT-R), Word recognition (WRAT-R Reading), Word recognition (WRAT-R Word Identification), Spelling (WRAT-R Spelling words correctly)	Decoding skills: 1.25 Spelling: 0.64 Word recognition: 0.96	Decoding skills: 1.47 Spelling: 0.72 Word recognition: 1.06
Kim et al (2010); USA	RCT (CON/INT) / Low proficiency / 4, 5, 6; NI; NI	264 (131; 133) / 23wk (92hr) / Levelled readers with no small group intervention (KidzLiz) vs Mixed: Levelled readers with a focus on computer assisted word study and teacher directed word reading, fluency, vocabulary, and comprehension activities (Read 180)	Levelled; Levelled	Word recognition (TOWRE), Decoding skills (TOWRE), Comprehension (GRADE), Vocabulary (GRADE), Fluency (DORF)	Fluency: 0.12 Comprehension: -0.09 Decoding skills: -0.04 Vocabulary: 0.04 Word recognition: 0.11	
Mesmer (2005); USA	RCT (CON/INT) / Moderate proficiency (low/high excluded) / 1; NI; NI	24 (12; 12) / 2wk (4.6hr) / Decoding instruction with levelled Reader vs Decoding instruction with phonics readers (Sundance)	Levelled; Decodable	Accuracy, Other - specify (Self Correction Rate)		Accuracy: NEI Other: NEI
Price-Mohr and Price (2020); UK	RCT (INT/INT) / All students / K; 4-5; NI	36 (N/A; 18; 18) / 30wk (45hr) / N/A vs Mixed: Decoding exposure (orthographic level) with progressive emphasis on comprehension with low-percentage decodable vs Mixed: Decoding exposure (orthographic level) with progressive emphasis on comprehension with high-percentage decodable	Levelled (high decodable); Levelled (low decodable)	Vocabulary (BPVS), Word Recognition (YARC), Phonemic Awareness (Sound isolation - YARC), Phonemic Awareness (Sound deletion - YARC), Comprehension (YARC)	Comprehension: NEI Phonemic Awareness: 0.11 Word Recognition: 0.16 Vocabulary: 0.18	

Study	Study Design (Research Type) / Specific population / Participants: Grade; Age range, Mean age	Sample size: Total (Control; Intervention/s) / Intervention volume: Weeks (hours) / Intervention: Control/ intervention/s	Focus of intervention texts	Outcomes (measure)	Outcomes - Levelled text	Outcomes - Decodable text
Ring et al (2012); USA	RCT (CON/INT) / Low proficiency / 2,3,4,5; 7-11; 9.46	86 (N/A; 37; 49) / 12wk (22hr) / N/A vs Re-reading instruction with levelled text (Read Naturally) vs Mixed: Phonics and sight words using a repeated reading strategy (Rite Flight)	Levelled; Decodable (with some levelled)	Reading rate (GORT), Accuracy (GORT), Fluency (GORT), Comprehension (GORT), Word recognition (TOWRE), Decoding skills (TOWRE), Word Recognition (WRMT), Decoding skills (WRMT Word Attack), Comprehension (WRMT)		Fluency: 0.16 Comprehension: -0.12 Word Recognition: 0.12 Decoding skills: 0.06 Accuracy: 0.24 Reading rate: 0.04
Vadasay and Sanders (2008); USA	RCT (CON/INT) / Low proficiency / 4,5; NI; NI	119 (65; 54) / 20wk (24hr) / Classroom instruction only vs Levelled strategy: Word recognition, comprehension and reading rate (Quick Reads as Dyad Reading)	NI; Levelled	Accuracy (Word ID - WRMT-R), Word recognition (TOWRE), Vocabulary (Quick reads MC test), Comprehension (Comprehension - WRMT-R), Fluency (DIBELS ORF benchmarks), Comprehension (WRMT-R/NU Passage Comprehension)	Fluency: 0.16 Comprehension: 0.23 Accuracy: 0.35 Vocabulary: 0.15 Word recognition: 0.00	
Beverley et al (2009); USA	Quasi (CON/INT) / Low proficiency / 1; 6-7; 6.9	32 (NI; NI; NI) / 8wk (8hr) / Read aloud to from authentic literature vs Systematic phonics instruction with individual reading practice using decodable texts vs Phonics instruction without individual reading practice	Authentic; Decodable; Decodable	Fluency (GORT), Comprehension (GORT), Accuracy (GORT), Reading rate (Rate)		Fluency: NEI Comprehension: NEI Accuracy: NEI Reading rate: NEI
Price-Mohr & Price (2018); UK	Quasi (CON/INT) / Low proficiency / 1; 4-6; NI	28 (16; 6; 6) / 26wk (50hr) / Synthetic phonics only and phonically decodable vocabulary vs Mixed: Decoding exposure (orthographic level) with progressive emphasis on comprehension with high % decodable vocabulary vs Mixed: Decoding exposure (orthographic level) with progressive emphasis on comprehension with low % non-decodable vocabulary	Levelled (low decodable); Levelled (high decodable)	Vocabulary (BPVS), Phonemic Awareness (YARC), Word Recognition (YARC), Phonemic Awareness (Sound isolation - YARC), Phonemic Awareness (Sound deletion - YARC), Comprehension (YARC), Reading rate (Reading speed - words per minute)	<i>Low decodable:</i> Comprehension: NEI Phonemic Awareness: 0.92 Word Recognition: 0.90 Vocabulary: 0.25 Reading rate: NEI <i>High decodable:</i> Comprehension: NEI Phonemic Awareness: 0.81 Word Recognition: 0.45 Vocabulary: 0.72 Reading rate: NEI	

Study	Study Design (Research Type) / Specific population / Participants: Grade; Age range, Mean age	Sample size: Total (Control; Intervention/s) / Intervention volume: Weeks (hours) / Intervention: Control/ intervention/s	Focus of intervention texts	Outcomes (measure)	Outcomes - Levelled text	Outcomes - Decodable text
Walpole et al (2017); USA	Quasi (CON/INT) / Low proficiency / 3, 4, 5; 8-10; NI	1101 (507; 594) / 52wk (117hr) / Guided reading and specific individual instruction) - Identified as Balanced Literacy vs Mixed: guided reading with small group differentiation (phonics using decodable, fluency and vocabulary with authentic text) - (Book Worms)	Levelled; Levelled (high complexity)	Fluency (DORF - DIBELS SUBTEST), Comprehension (Scholastic Reading Inventory)	Fluency: 0.27 Comprehension: 0.18	

BPVS = British Picture Vocabulary Scale; DORF = Dynamic indicators of basic early literacy skills oral reading fluency; GORT = Gray Oral Reading Test; GRADE = Group Reading Assessment and Diagnostic Evaluation; TOWRE = Test of Word Reading Efficiency; TPRI = Texas Primary Reading Inventory; WRMT-R = Woodcock Reading Mastery Test; WJ-III = Woodcock-Johnson III diagnostic reading battery; WRAT = Wide Range Achievement Test; YARC = York Assessment of Reading Comprehension.

focused solely on decodable instruction (i.e., only phonics-based instruction, no comprehension instruction) and one (1) solely on comprehension-based instruction (i.e., no phonic decoding utilised). With regard to the comparison/control conditions, interventions were mostly compared to an alternate text type (i.e., comparison of levelled and decodable text) with a different pedagogical approach (n = 8, 53%) or 'business as usual' instruction with limited description of the text type or pedagogy used (n = 5, 33%). One (1) intervention focusing on levelled text was compared against a more decodable version of levelled text with the same pedagogy used across groups.

Risk of bias

Risk of bias assessment is outlined in Table 2 and detailed in Table 4. High risk of bias means that there is a high likelihood that the results observed in a study may not be a 'true' representation of the effects that occurred during an intervention. Of the 11 studies assessed, the majority were considered to have high risk of bias (n = 7, 64%), with three (3) studies considered at moderate risk, and only one (1) study displaying low risk of bias. The majority of studies (n = 9, 82%) displayed low levels of bias in the identification and randomisation of participants to create conditions that were close to even at a studies baseline assessment. Controlling for unintended or unknown exposure to interventions, especially around an academic skill set that is a focus at home and school, is very difficult in education-based interventions (compared with, for example, health-based ones). Risk of bias was variable across studies for intervention deviation, with programs focused on individual instruction of students (as opposed to cluster trials focused on the class group) appearing to control better for this form of bias. The way that data was collected (i.e., collection by personnel separate to the research team), missingness (i.e., proportion of follow-up data) and the reporting of results (i.e., addition of outcomes at post-test) all display variable, but mostly moderate to high, risk of bias.

Outcomes and effect sizes

Outcome effect sizes are detailed in Table 2 and summary effect sizes (average of available study outcomes) are detailed in Table 3. Effect sizes were calculated for studies that provided sufficient information to calculate them. Two (2) studies provided insufficient information for all outcomes assessed and two (2) studies provided insufficient information for some outcomes (i.e., baseline or follow-up

Table 2. Data extraction for included experimental and quasi-experimental studies

Study	Characteristics								Coding		Outcome effect sizes									
	Design	Grade	Reading level	Volume (hrs)	Sample	Risk of bias	Control text (pedagogy)	Intervention text (pedagogy)	Text focus	Text (low only)	Fluency	Comprehension	Phonemic Awareness	Word Recognition	Decoding skills	Accuracy	Spelling	Vocabulary	Reading rate	Silent Reading Fluency
Cheatham (2014)	RCT (C/I)	2	All	5	62	High	Auth (none)	Lev (Mixed)	Lev	NA				0.13	0.09					
Denton (2014)	RCT (C/I)	1	Low	46	218	High	NI (as usual)	Lev (Mixed)	Lev	Lev	0.22	0.09			0.16					0.20
							NI (as usual)	Dec (Mixed)	Dec	Dec	0.54	0.38			0.39					
Jenkins (2004)	RCT (C/I)	1	Low	50	121	High	NI (as usual)	Dec (Mixed)	Dec	Dec				1.06	1.47				0.72	
							NI (as usual)	Lev (Mixed)	Lev	Lev				0.96	1.25					
Kim (2010)	RCT (C/I)	4-6	Low	92	264	Mod	Lev (none)	Lev (Mixed)	NA	NA	0.12	-0.09		0.11	-0.04				0.04	
Mesmer (2005)	RCT (I/I)	1	Mod	5	24	High	Lev (NA*)	Dec (Dec)	NA	NA						NEI				
Price-Mohr (2020)	RCT (I/I)	K	All	45	36	High	Lev - high (NA*)	Lev - low (Mixed)	Lev	NA			NEI	0.11	0.16				0.18	
Ring (2012)	RCT (I/I)	2-5	Low	22	86	High	Lev (re-read)	Mixed (Mixed)	Dec	Dec	0.16	-0.12		0.12	0.06	0.24			0.04	
Vadasay (2008)	RCT (C/I)	4-5	Low	24	119	High	NI (as usual)	Lev (Lev)	Lev	Lev	0.16	0.23		0.00		0.35		0.15		
Beverley (2009)	Quas (C/I)	1	Low	8	32	Mod	Auth (read aloud)	Dec (Dec)	NA	NA	NEI	NEI				NEI			NEI	
							Auth (read aloud)	Dec (Mixed)	NA	NA	NEI	NEI				NEI				
Price-Mohr (2018)	Quas (C/I)	1	Low	50	28	High	Dec (phonics)	Lev - low (Mixed)	Lev	Lev			NEI	0.92	0.90			0.25	NEI	
							Dec (phonics)	Lev - high (Mixed)	Lev	Lev			NEI	0.81	0.45				0.72	NEI
Walpole (2017)	Quas (C/I)	3-5	Low	117	1101	High	Lev (balanced)	Lev (Mixed)	Lev	Lev	0.27	0.18								

Note. Coding is included to indicate grouping of studies for calculation of summary statistics. RCT = Randomised Controlled Trial; Quas = Quasi-experimental trial; NI = Not included (information not provided in publication); NEI = Not enough information presented to calculate effect sizes; NA = Not applicable; (C/I) = Control vs Intervention study; (I/I) = Intervention vs intervention study; Lev = Levelled; Dec = Decodable; (n) = number of studies summarized; * Pedagogy is the same as the intervention (only the text exposure is modified); Mod = moderate.

Table 3. Summary of effects for included experimental and quasi-experimental studies

Analysis level	Average of outcome effect sizes									
	Fluency	Comprehension	Phonemic Awareness	Word Recognition	Decoding skills	Accuracy	Spelling	Vocabulary	Reading rate	Silent Reading Fluency
Studies (n)	6	6	3	9	7	2	2	5	1	2
Average effect size	0.25	0.11	0.61	0.43	0.48	0.30	0.68	0.27	0.04	0.31
min	0.12	-0.12	0.11	0.00	-0.04	0.24	0.64	0.04	0.04	0.20
max	0.54	0.38	0.92	1.06	1.47	0.35	0.72	0.72	0.04	0.42
Text focus										
Levelled (n)	0.22 ⁽³⁾	0.17 ⁽³⁾	0.61 ⁽³⁾	0.43 ⁽⁶⁾	0.50 ⁽³⁾	0.35 ⁽¹⁾	0.64 ⁽¹⁾	0.33 ⁽⁴⁾		0.20 ⁽¹⁾
Decodable (n)	0.35 ⁽²⁾	0.13 ⁽²⁾		0.59 ⁽²⁾	0.64 ⁽³⁾	0.24 ⁽¹⁾	0.72 ⁽¹⁾		0.04 ⁽¹⁾	0.42 ⁽¹⁾
Text focus – Low competence										
Levelled (n)	0.22 ⁽³⁾	0.17 ⁽³⁾	0.87 ⁽²⁾	0.58 ⁽⁴⁾	0.71 ⁽²⁾	0.35 ⁽¹⁾	0.64 ⁽¹⁾	0.37 ⁽³⁾		0.20 ⁽¹⁾
Decodable (n)	0.35 ⁽²⁾	0.13 ⁽²⁾		0.59 ⁽²⁾	0.64 ⁽³⁾	0.24 ⁽¹⁾	0.72 ⁽¹⁾		0.04 ⁽¹⁾	0.42 ⁽¹⁾
Study design										
RCT (C/I) (n)	0.26 ⁽⁴⁾	0.15 ⁽⁴⁾		0.45 ⁽⁵⁾	0.55 ⁽⁶⁾	0.35 ⁽¹⁾	0.68 ⁽²⁾	0.10 ⁽²⁾		0.31 ⁽²⁾
RCT (I/I) (n)	0.16 ⁽¹⁾	-0.12 ⁽¹⁾	0.11 ⁽¹⁾	0.14 ⁽²⁾	0.06 ⁽¹⁾	0.24 ⁽¹⁾		0.18 ⁽¹⁾	0.04 ⁽¹⁾	
Quasi (n)	0.27 ⁽¹⁾	0.18 ⁽¹⁾	0.87 ⁽²⁾	0.68 ⁽²⁾				0.49 ⁽²⁾		

RCT = Randomised Controlled Trial; Quasi = Quasi-experimental trial; (C/I) = Control vs Intervention study; (I/I) = Intervention vs intervention study; (n) = number of studies summarized.

Table 4. Risk of Bias for included experimental and quasi-experimental studies

Experimental								
Study	Type	Randomization process	Timing of identification / recruitment	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall risk of bias rating
Cheatham et al (2014)	Individual	Low	N/A	High	Low	High	High	High
Denton et al (2014)	Individual	Low	N/A	Moderate	Low	High	Moderate	High
Jenkins et al (2004)	Individual	Moderate	N/A	Moderate	High	Moderate	Moderate	High
Kim et al (2010)	Individual	Low	N/A	Low	Low	Moderate	Moderate	Moderate
Mesmer (2005)	Cluster	Low	Low	High	High	High	Moderate	High
Price-Mohr and Price (2020)	Cluster	Moderate	Low	High	Low	Moderate	Moderate	High
Ring et al (2012)	Cluster	Low	Low	Low	High	Low	Moderate	High
Vadasay and Sanders (2008)	Individual	Low	N/A	Moderate	High	Low	Moderate	High
Quasi-experimental								
Study	Confounding	Selection of participants	Intervention classification	Deviations from intended interventions	Missing data	Measurement of the outcome	Selection of the reported result	Overall risk of bias rating
Beverley et al (2009)	Low	Low	Low	Low	Low	Low	Serious	Moderate
Price-Mohr & Price (2018)	Low	Low	Low	Low	Low	Moderate	Low	Moderate
Walpole et al (2017)	Low	Low	Low	Low	Low	Low	Low	Low

assessment only). The most commonly measured outcome was Word Recognition (n = 9) followed by Decoding Skills (n = 7 studies), Fluency (n = 6 studies), Comprehension (n = 6 studies), and Vocabulary (n = 5 studies).

The overall impact of intervening on student reading (regardless of the text type used) displayed average effects ranging from $d = 0.04$ (n = 1) for Reading Rate to $d = 0.68$ (n = 2) for Spelling. Of the outcomes with greater than four (4) effect sizes contributing to their calculation, the average effect observed for the reading interventions included here was small for Comprehension ($d = 0.11$), Fluency ($d = 0.25$) and Vocabulary ($d = 0.27$), and moderate for Word Recognition ($d = 0.43$) and Decoding Skills ($d = 0.48$).

When investigated by the focus of the text used in the intervention, for the majority of outcomes, the difference in the average effect of intervention using a specific text type was marginal between the two text groups ($-d \pm 0.10$). Of the outcomes with greater than $d = 0.10$ difference in the average effect and more than a single study contributing to the effect calculation, Fluency, Word Recognition and Decoding Skills outcomes were stronger among the decodable interventions.

When comparison of text types was made including only studies that sampled participants with low reading proficiency and more than a single study for computation of effect, Fluency was the only outcome to display a difference greater than $d = 0.10$ between the text groupings, with greater outcomes for exposure to decodable text ($d = 0.22$ [Lev]; 0.35 [Dec]). Effect sizes overall for intervention among low proficiency readers using either text type were considered low for Comprehension ($d = 0.17$ [Lev]; 0.13 [Dec]), and moderate for Word Recognition ($d = 0.58$ [Lev]; 0.59 [Dec]) and Decoding Skills ($d = 0.71$ [Lev]; 0.64 [Dec]). Additionally, interventions focused on levelled text exposure produced a small effect on Vocabulary ($d = 0.37$) and a large effect on Phonemic Awareness ($d = 0.87$), although both effects used in this estimate of Phonemic Awareness were from a single study and should be interpreted with caution.

When a 'grand average' effect was calculated for each text type using only the outcomes that have effects for both text types (Fluency, Comprehension, Word Recognition, Decoding, Accuracy, Spelling and Silent Reading Fluency), the difference was trivial between the text types ($d = 0.41$ [Lev]; 0.44 [Dec]).

With reference to study design, randomised controlled trials produced lower effect size estimates than quasi experimental for all outcomes, and randomised controlled trials comparing multiple interventions (as opposed to comparison to a control condition) produced lower effect estimates for all but one outcome category.

KEY RESULTS AND DISCUSSION OF RESULTS

This discussion cumulatively addresses the three critical research questions posed at the beginning of this report. The findings reported here reflect the limitations of the quality of the evidence identified in the systematic literature review, including the fact that, as previously mentioned, no randomised controlled trials that measure the educational outcomes of using decodable and/or levelled reading books in primary school classrooms, the strength and the significance of the effect, have been conducted in Australia in at least the past 20 years (Torgerson et al., 2006). The best currently available evidence has been used to compile this report.

Key messages from the studies analysed include:

- The majority of studies focused on lower proficiency readers in lower to middle primary grade levels.
- Interventions, while focused on exposure to a majority text type and corresponding pedagogical intervention, mostly involved minority components of the opposing text type and corresponding pedagogical intervention to be seen as a 'mixed' intervention. This appears to correspond with the desire to obtain a range of reading related outcomes (i.e., decoding and comprehension capabilities) from a single intervention.
- The research methods utilised, while robust in their design, left many studies with greater than desirable risk of bias. This reduces the trustworthiness of the reported impacts of the interventions being investigated.
- The measurement of outcomes consisted of proximal outcomes of varying complexity (i.e., individual component skills of reading competence), with no measurement of the impact on distal assessment (i.e., standardised testing – for example, NAPLAN reading).

- For students of lower reading proficiency, there appears to be no clear advantage in focusing specifically on levelled or decodable text when intervening on reading development.

The debates surrounding the beneficial practice of using decodable or levelled reading books from empirical studies was an impetus for this research. While there is a lot of politicking about phonics and how students should be taught to read, it was identified through the conduct of this review that these arguments often draw on little, if any, empirical research. As was identified in the research, often the studies are descriptive, based on unsubstantiated opinions, self-referenced, and/or referenced other research studies that had no empirical grounding, but reported on as though it did. Definitions of decodable and levelled reading books varied, and the terms are often used interchangeably in the various articles. Both decodable and levelled reading books have strengths and limitations in teaching beginning reading (See Table 4). Key concepts that were evident in both the qualitative and quantitative data are summarised below.

1. Teachers and educational practitioners need to look at *features of reading books* rather than a type or label.

As it stands, the evidence suggests that overall, mixed interventions that focus on decoding and comprehension level and phonics based readers produce moderate effects across the board. With crossover of features between decodable and levelled reading books, Hiebert (2005) supports the notion that the features found within reading books need to be considered by looking at each text for how they will best serve the learning needs of individual students and their distinctive stages of development. The influence of texts (including reading books) is diverse (Mesmer, 2010) and should also be looked at in terms of their functionality in how they develop context, interest, and engagement for students. This supports the held philosophy of the value of varied reading books for individual students at various stages and ages in reading development (Brown, 2000; Jenkins et al., 2003; Mesmer, 1999, Beverley et al., 2009). Quality instructional reading materials for students need to be comprised of multiple features and components to accommodate the complexity of reading and to provide opportunities for explicit teaching of the various components required reading proficiency. The evidence from this systematic literature review supports the

Table 4: Advantages and disadvantages of decodable and levelled reading books as identified in the systematic literature review

Advantages of Levelled Reading Books

- Allows practice of high frequency words (Mesmer, 2010).
- Percentage of high frequency words and syntactic repetition will lead to greater gains (Kuhn and Stahl, 2003 as cited in Mesmer, 2010), for example two times the high frequency words than decodables (Mesmer, 2010).
- Encourages students to use syntactic and memory strategies (Mesmer, 2010).
- Can produce higher reading rate (Mesmer 2010) and more fluency across time and with more practice.

Disadvantages of Levelled Reading Books

- Evidence that reading was less accurate (Mesmer, 2010).
- Results support the view that students are motivated to read above their grade level when the content is enjoyable. In this way they gain a better understanding of the text as opposed to reading books that restrict text and vocabulary limited to their graphophonemic ability (Price-Mohr & Price, 2020).

Advantages of Decodable Reading Books

- Usefulness of decodables on word recognition strategy (Mesmer, 2010).

Disadvantages of Decodable Reading Books

- Usefulness of decodable text is highly conditional in only specific ways, specific times of the year, specific strategies or with one-on-one contexts (Mesmer, 2010).
- Systematic, synthetic phonics instruction failed to improve meaning construction and detracted attention away from the purpose of obtaining meaning when reading (Altwerger et al., 2004).
- Negative impact on comprehension beyond the literal level (Altwerger et al., 2004).
- No benefits to using phonics in isolation (Altwerger et al., 2004).
- Decodable text not necessarily easier to read (Price-Mohr, 2020).
- Inconclusive evidence if decodable reading books improve accuracy (Mesmer, 2010).

view that students can benefit from just one element (for example, decodability) in reading books. However, reading books that draw upon multiple criteria and elements such as sound letter relationships, high frequency words, vocabulary, meaningfulness, levels of predictability, are of greater benefit in improving reading achievement (Cheatham et al., 2014, Mesmer, 2010). Cheatham et al (2014) extend this concept further to call for researchers, educators and publishers in education to consider multiple criteria when developing and implementing early reading materials, particularly for students with low literacy proficiency. Studies draw attention to the benefits of multiple features of reading books to foster reading development including authentic literature (Beverley et al., 2009; Lyons et al., 2012). Mesmer (2010) calls on educators to reflect on the degree that selected reading books fulfil their purpose and to apply this accordingly in their classrooms.

Research data was inconclusive as to whether the amount of time spent on classroom reading instruction had a measurable impact on reading proficiency development (Mesmer, 2010). This finding furthers the argument that text features are important considerations for teachers, arguably pointing to a balanced approach (see next preliminary finding) as being of value. When examining decodable and levelled reading books side-by-side, the featured differences between each diminished in Year 1, suggesting that eventually textual scaffolds surpass their usefulness; as students' literacy proficiency developed, they grasped more complex reading books with fewer scaffolds. At the earliest levels, it could be argued that reading books with highly, or a significant level of, decodable features outperformed reading books with high frequency words. However, as Mesmer (2010) reported, reading books with more high frequency words increased fluency and students had lower reading rates in decodable than in levelled reading books.

2. Balance is not inferior

A recurrent theme in all the papers included in the systematic literature review is a matter of equilibrium, both decodable and levelled reading books and their features develop aspects of reading if considered from a holistic lens. Beverley et al. (2009) uphold the theory that reading is a complicated and multifaceted process that cannot be simply deconstructed and reduced to a few components (drawing on Adams, 1990;

Scarborough, 2001; Compton-Lilly et al., 2020; Castles et al., 2018). To accommodate the multidimensional elements of reading, reading instructional materials should be multifaceted to meet their stated purpose. One model or book may suit some students (and subsequently a school's approach), but this one model will not produce equitable reading achievement opportunities for all (Rightmyer et al., 2006). Rather, research indicated that exclusive use of one model or approach has a detrimental effect, and that carefully selected and varied reading materials leads to a more positive outcome and attitude to reading (Mesmer, 2005; Wood, 2005). In their study, Rightmyer et al. (2006) emphasized the reminder from the National Reading Panel (2000) which cautioned educational leaders to not give blanket endorsements of any phonics programs. Teachers need to incorporate many elements to reading instruction and White (2009) supports the idea of a 'balanced' approach in his study of systematic and analogy-strategic based approach to phonics. Students need multiple experiences, strategies, and support within the teaching of reading (Ankrum, 2020). This concept of balance should also be applied to the reading books that support this varied experience and approaches. Furthermore, experiences need to be created that foster "critical reflection and intellectual engagement" (Altwerger et al., 2004, p. 129). This process cannot be simplified to one method, strategy, or type of text.

3. The teacher and their pedagogy are the critical component—surpassing the type of text used in reading instructions.

There is no debate on the impact of an excellent teacher (Hanushek, 2014; Hattie, 2003), also supported by Wood (2005) in the context of reading. The impact of the teacher, both their instructional approaches and their attitude towards reading is crucial. Poor practice and a negative attitude from the teacher will adversely affect reading regardless of text level (Morgan et al., 2000, Rightmyer et al., 2006). Teachers need to carefully consider the reading books and reading materials for their students and be pedagogically diverse in their use of reading books to meet the needs of their students in supporting them to new levels of achievement (Wood, 2005; Beverley et al., 2009; Mesmer, 2012). Importantly, the type of reading books used in reading instruction is impacted significantly by their integration with pedagogical approaches used by the teacher (Price-Mohr & Price, 2020).

An additional finding from the systematic literature review is that more effort and attention is needed to support teachers in their decision-making to form different choices depending on the individual profiles of their students' learning. In Hiebert's study (2005), the greatest gains (reading fluency) were from a teacher who did not systematically implement the intervention to all the students. This could be viewed as a limitation in the use of either type of reading book (decodable or levelled) and the often linked approaches to the teaching of reading. This finding shows that teacher knowledge and understanding about instructional approaches and the components that make up the complex process of reading cannot be underestimated. Reflecting on this idea in relation to fluency, Hiebert (2005) states, "Teachers' knowledge about fluency, their choices about fluency instruction and the effects of these choices on student outcomes, and discussions with teachers about these choices need to be a focus of future research on fluency" (p. 206).

Moreover, the type of reading book or associated models or programs of beginning reading instruction are impacted by the classroom integration and pedagogical approaches used by the teacher (Hofman & Pearson, 2015; Price-Mohr & Price, 2020). The teacher is just as important as the model, approach, or text (Rightmyer et al., 2006). While approaches, books, and models can be recommended and even mandated by schools, institutions, or education departments, without extensive teacher knowledge and professional development of these programs, how they each contribute to reading (Hiebert, 2005; Rightmyer et al., 2006), and the skills of the teacher to implement the programs and materials effectively, they are likely to fail and either be of no benefit to students, or worse, have a detrimental effect. The pedagogical application therefore is just as or more powerful than the text (Jenkins et al., 2004; Johnston, 2001; Mesmer, 2010).

CONCLUSIONS

This section addresses the role of the teacher, limitations of the studies and findings, interesting observations, and provides a discussion on the findings, including identifying possible future research endeavours to investigate literacy teaching and learning in Australian primary school classrooms.

The role of the teacher

One of the most significant outcomes from the systematic literature review that this report draws from is the idea that the teacher is at the centre of, or at least is a leading motivator for, student learning success. Effective teachers understand that reading is a complex process and use an integrated approach for setting the foundation to reading through decoding and a toolbox of skills and strategies that cater for all their learners' needs. Teachers therefore need to be trusted to work with the students in their classroom, as they best know their students' needs and how they learn. Equipping teachers with relevant resources will support them to carry out their professional duties. Several studies (see, for example, Ankrum, 2021) include this observation in their findings. For example,

It has been stated that only the teacher, not a specific program or material, makes the difference in literacy instruction (Hofman & Pearson, 2015). Teaching reading is complex and calls for high levels of teacher knowledge and decision making. It is imperative for teachers to make thoughtful decisions about which books to use, when, and for what purposes (Mesmer et al., 2012). (Ankrum, 2021, p. 6)

This points to the need to support teachers, to trust their professional judgement and pedagogical skills, and to provide them with the material resources to be able to teach literacy and, more specifically, reading skills using a repertoire of practices that suits the students, classroom context, and school environment of each individual teacher. Altwerger et al. (2004) proffer the following:

Our goal is to create intellectual spaces in which we can all think critically about what is happening in schools and classrooms, by drawing on

empirical evidence, our own and our students' learning, and the ideas of others. We strive to empower our students as knowledgeable decision-makers able to resist a system that silences them and their children through mandated programs and standardized testing. As teacher educators, we also must resist falling victim to the pressure to align ourselves with programs that silence teachers and their students, and instead, re-envision ourselves as catalysts for reclaiming professionalism in education (Altwerger et al., 2004, p 131).

Limitations of the studies and findings

Several limitations were clear when looking at the corpus of identified studies, identifying a clear gap in research that could be undertaken as randomised controlled trials in mixed ability, mainstream, primary school classrooms in Australia. These limitations are noted here:

- The majority of the programs (17) identified in the systematic literature review were focused and conducted on remediation of identified students with low literacy proficiency. This taints the findings, as the majority of published research into the best strategies and/or teaching methods in reading acquisition is overwhelmingly conducted on the remediation of students who already have low literacy proficiency rather than comparative studies on students beginning to read at the inception of the process.
- There was a limited number of only nine (9) randomised controlled trials (RCTs; others included five (5) quasi experimental designs and seven (7) pre-post designs). More studies are needed to compare across a control.
 - Connected to this, there were too few studies to form an effect size.
- It can be difficult to pinpoint the outcomes as multiple measures are similar, for example word attack, word accuracy, word recognition, and reading rate.
- In the initial screening of the studies (refer to Figure 2), which identified 3025, few articles were included post-first stage as they were not grounded in any empirical research, with many containing wholly theoretical research, including drawing on other descriptive studies. Others made claims that were not supported by their own empirical research, for example studies that surveyed

teachers for perspectives rather than research in the classroom, yet claims were made for student literacy attainment despite not being able to be substantiated, and importantly, not being the purpose of the initial research.

- The identified studies that were analysed for the quantitative data had high and moderate risk of bias. Oftentimes the researchers engaged in the study had a stake or other interest (such as being an author or creator) in the educational product (for example, the reader or reading program) being used as part of the study being reported on. This provides an opportunity for further scholarly research (for example, a randomised controlled trial) to take place that reduces the risk of bias evident in current studies.
- None of the identified studies that used a decodable reading book measured phonemic awareness.
- All the identified studies suffer from reactivity, also known as ‘the observer effect.’
- Testing a commercial product, as many of the identified studies did, is not in and of itself a problem. However the studies need to be set up well to do this properly, and this has not been the case (as can be seen, for example, through moderate and high risk or bias)
- Reading age was not consistently used as a measure across studies limiting the reliability of their results in terms of whether or not it was the intervention that made the difference or whether it would have occurred in any case as part of student growth in a school year.

Interesting observations

Across the canon of identified studies, several interesting observations were made, which in some ways speak to the complexity of undertaking research in a school setting:

- It can be difficult to differentiate the *literacy program* from the *teacher’s method of instruction* or *pedagogy* which is impacted by many variables including teacher knowledge, experiences, and skills. It was found that teacher presence plays a not insignificant role in student achievement.
- The length of time of the study is not significant in making a notable difference in student achievement; whether that achievement gain is maintained post-

intervention, or after the research project in which the students participated is worthy of further consideration and research. This study found that interventions for students who are *struggling* or *poor* readers or who are *belowgrade level* need to continue for more than just one year, and need to be followed up in subsequent years to ensure that their gains remain.

- Levelled reading books are more engaging for students and encourage them to read for enjoyment.
- Students need to have the opportunity to read and re-read texts to increase fluency, word recognition and associated literacy skills.
- A mixed-methods approach to literacy instruction (see, for example Kim et al., 2011) appears to be a quality approach to achieve excellent results in improving literacy proficiency.
- Any type of intervention with control groups who received no special treatment and no intervention at all often performed well under the intervention groups regardless of what they were (see, for example, Jenkins et al., 2004). However, the balanced approach and those using levelled reading books often fared better than those with an overt focus on phonics/decodables.

To conclude, the research included in this review has highlighted several essential elements and teaching practices that provide opportunities for all students to acquire reading skills. The daily learning experiences and difficulties of many vulnerable groups in our society is often not considered in these debates: social, emotional, and economic factors along with language barriers are often overlooked by groups presenting a pitch for one particular method (often wrapped in a commercial product) over another. Freebody (2007) discussed this in his *Literacy Education in School* report, writing: “Literacy education has become the scapegoat of choice for the economic, social, moral and intellectual fragilities and failings of our society, or at least its immediately impending fragilities and failings, or, at the very least, the fragilities and failings of some groups within the society.” (p. 70)

Ultimately, in the current schooling context, external based assessments, such as NAPLAN and PISA testing are considered imperative in shaping educational decision makers and subsequently impacts the teaching of reading and the programs and approaches that are adopted at both school and government levels. Studies need to

be conducted that measure distal outcomes (that is research with long term goals over an extended period) to ascertain the impact of any intervention. As recommendations around the teaching of reading are in the main aimed at lower proficiency students, but in the future more empirical research is needed to ascertain the impact on all students regardless of their level of reading proficiency.

The central focus of the study was on teaching reading or reading instruction using decodable and/or levelled reading books. This review has evaluated the evidence needed to build effective teaching of reading for all. As with most aspects of effective teaching and learning in primary school classrooms, there is not one simple answer to the question of which material is best for teaching students to read and improving their literacy proficiency.

ACKNOWLEDGEMENTS

The authors would like to thank and acknowledge Research Assistant Sean Groth for his work with the Risk of Bias Assessments for the included studies. We are also grateful for the support and guidance of the Library Services, Academic Division at the University of Newcastle, specifically Research Liaison Librarian Jennie Skulander for her assistance with conducting the database search and strategies when developing the protocol, and locating papers required for the review.

We acknowledge the financial support provided by Cengage Learning Australia for this review to occur.

GLOSSARY

Definitions of terms used in conversations, research, and in-practice situations can be confusing as various educators, researchers, and education jurisdictions use the same term in different ways, and there can be some overlap and sometimes contradictions between them. Key terms included in this report and identified in this glossary of terms have been developed from the research articles included in the study, from literacy expert understandings, education department definitions, and from curriculum materials such as syllabuses and other support documents. Where relevant, especially where multiple understandings/definitions can be readily found for the same term, referencing has been provided. Given a significant proportion of the research articles included in this study are US based there are many terms included here that are not widely used in the Australian context (for example, trade books) and these are important to define as they will enable this report to be read in context. The definitions provided are applicable to their use in this report only and are offered as a clarification of general terms, specific terms, and those that can have multiple meanings.

Authentic texts

Books which are not written specifically for the teaching of reading but are often used in the classroom. Examples include, websites, picture books (for example Roald Dahl books), newspaper articles, and novels. In short, they are written for 'real world' contexts.

Balanced literacy

A philosophy of teaching literacy that encompasses the full and complex science of reading.

Comprehension

The ability to process and understand meaning—spoken, written, visual.

Cueing systems

The integrated use of more than one specific decoding strategy.

Decodability

Decoding skills are the skills required to read and write. This is generally more associated with skills linked to the phoneme and grapheme correspondences. It involves the process to recognise and identify letter(s), matching the correct letter(s)

to the sound correlation, retain the sounds in sequence to blend the word, pronounce the word and then apply meaning.

Decodable text/reader

- *See definition in report*

Direct instruction

A teacher-centred pedagogy that originates from the 1960s with Siegfried Engelmann in the special education sphere. It places emphasis on scripted lessons and materials, organised around clearly defined learning increments. Teacher creativity and autonomy are advised to 'give way' to follow carefully prescribed instructional practices (National Institute for Direct Instruction, 2015).

Encoding

The process of identifying the phonemes to graphemes for spelling words.

Explicit instruction/explicit teaching

A teacher-centred approach or pedagogy that involves the teacher delivering knowledge of concepts through explicitly teaching strategies and skills to students, it involves clearly showing students what to do and how to do it rather than leaving students to construct and organise information on their own (Goulding, 2021, p. 81). It combines instructional practices to produce clearly defined content, goals and outcomes for individual lessons with continuous checking for understanding.

Graphemes

The written letter patterns used to represent phonemes.

Graphophonemic

The awareness of the connection between phonemes and graphemes.

Guided reading

A small group teacher-centred approach to reading where explicit instruction is focused on strategies to assist decoding and constructing meaning, to meet the needs of each individual learner.

Fluency

The ability to read with accuracy, speed, rhythm, intonation, and expression. This can be reading out-loud or silently.

High frequency words

Words most commonly used in reading and writing. These are the words relied on for the flow of syntactic structure in English. If these words are unknown, a sentence can't be 'glued' together. These are mostly function (or structure) words such as 'to' and 'of'

which have little meaning on their own and include conjunctions, pronouns, articles, and prepositions.

Instructional books/readers

Non-fiction, informational or expository reading books.

Levelled reader/text

- See definition in report.

Little books

Books designed for the teaching of reading and interaction between the reader and teacher. They are written using: predictable text with high frequency words; on familiar topics engaging to students so they can draw context and utilise the connected text and pictures.

Morphemes

The smallest unit of meaning in language (a word or part of a word).

Morphography

Understanding written morphemes.

Phonological awareness

Encompasses broad aspects of spoken language and is strongly connected to early reading and spelling efficiency due to its association with phonics. It involves the ability to recognise larger aspects of language than individual phonemes in words, such as sentences and whole words. This includes broader phonological patterns such as onset, rhyme and alliteration, syllables, and segmenting a sentence and words.

Phonemes

The smallest units of spoken sound that distinguish meaning in words.

Phonemic awareness

The ability to hear, identify, and manipulate phonemes.

Phonics

The explicit teaching of the connection between spoken sounds (phonemes) and written symbols (graphemes). There are many methods within explicit teaching of phonics, three examples are:

- **Analytic phonics**
 - A phonics teaching method that teaches the phonics relationship of words by analysing the phonics patterns in context which also includes whole word recognition. Example – examining a word as a whole and breaking into individual sounds to deconstruct a word.

- **Synthetic phonics**
 - A phonics teaching method that teaches grapheme to phoneme correspondences and the blending of taught grapheme to phoneme correspondences to produce words. This involves focusing on the individual sounds to make the whole word – identifying phoneme, applying the grapheme and then synthesising to produce a word.
- **Systematic phonics**
 - The teaching of phonics using a clearly defined sequence. This could be phonics in all its forms; systematic synthetic (SSP), analytic phonics (Torgerson et al., 2006) and in context.

Predictable books/reader

Beginning reading books that contain repetitive words, phrases, and rhymes.

Reading rate

The speed at which someone reads with accuracy.

Semantics

Concerned with word meaning.

Sight words

Learning words without decoding; words that are learnt by recognition, memory, whole word recognition. Visual sight vocabulary and words with less transparent orthography.

Sight word efficiency

Correct identification of sight words in a timed process.

Trade books

Books which are published to sell to the general public, and like *authentic texts* may be used in the classroom to teach literacy, for example *The Very Hungry Caterpillar* by Eric Carle.

Word accuracy

The ability to read words correctly.

Word accuracy (tested)

The percentage of attempted words read correctly.

Word recognition

Words that are recognised regardless of what cues/strategies are used.

Whole language

A philosophy of teaching literacy that is meaning centred and is focussed on strategies to show that language (oral and written) is a system of parts that work together to create meaning.

REFERENCE LIST

- Altwerger, B., Arya, P., Jin, L., Jordan, N. L., Laster, B., Martens, P., Wilson, G. P., & Wiltz, N. (2004). When research and mandates collide: The challenges and dilemmas of teacher education in the era of NCLB. *English Education*, 36(2), 119–133.
- Ankrum, J. W. (2021). Complex texts or leveled readers for the primary grades? Yes and yes! *Early Childhood Education Journal*. <https://doi.org/10.1007/s10643-021-01178-1>
- Australian Curriculum, Assessment and Reporting Authority (n.d.). National literacy learning progression: Reading and viewing. Understanding texts description. Retrieved from <https://www.australiancurriculum.edu.au/resources/national-literacy-and-numeracy-learning-progressions/national-literacy-learning-progression/reading-and-viewing/?subElementId=50915&scaleId=0>
- Australian Literacy Educators' Association (2015). Literacy in 21st Century Australia: The ALEA Declaration. Retrieved from <https://www.alea.edu.au/about/alea-literacy-declaration/>
- Awah, I. (2020, December 18). *Five ways to help your child develop a love for reading*. The Conversation. Retrieved from <https://theconversation.com/five-ways-to-help-your-child-develop-a-love-for-reading-150669>.
- Bantick, C. (2006, September 24). Why our kids must learn to love the comma. *The Sunday Mail*, p.52.
- Barry, A.L. (2008). Reading the past: Historical antecedents to contemporary reading methods and materials. *Reading Horizons*, 49(1), 31-52.
- Beverly, B. L., Giles, R. M., & Buck, K. L. (2009). First-grade reading gains following enrichment: Phonics plus decodable texts compared to authentic literature read aloud. *Reading Improvement*, 46(4), 191-205.
- Brown, C. S. (2000). Language and literacy development in the early years: Foundational skills that support emergent readers. *The Language and Literacy Spectrum*, 24, 35-49. <https://files.eric.ed.gov/fulltext/EJ1034914.pdf>
- Buckingham, J. (2018) The fight for phonics in early years reading. Research Ed. <https://researched.org.uk/2018/07/04/the-fight-for-phonics-in-early-years-reading/>
- 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>
- Cervetti, G. N., Pearson, P. D., Palincsar, A. S., Afflerbach, P., Kendeou, P., Biancarosa, G., Higgs, J., Fitzgerald, M. S., & Berman, A. I. (2020). How the reading for understanding initiative's research complicates the simple view of reading invoked in the science of reading. *Reading Research Quarterly*, 55(S1), S161–S172.
- Chandler, J. & Hopewell, S. (2013). Cochrane methods: Twenty years experience in developing systematic review methods. *Systematic Reviews*, 2(76). <https://doi.org/10.1186/2046-4053-2-76>

- Cheatham, J. P. & Allor, J. H. (2012). The influence of decodability in early reading text on reading achievement: A review of the evidence. *Reading and Writing* 25(9) 2223–2246
- Clay, M. (1991). *Becoming literate: The construction of inner control*. Heinemann.
- Cheatham, J. P. & Allor, J. H. (2012). The influence of decodability in early reading text on reading achievement: a review of the evidence. *Reading and Writing*, 25(9), 2223-2246.
- Cheatham, J. P., Allor, J. H., & Roberts, J. K. (2014). How does independent practice of multiple-criteria text influence the reading performance and development of second graders? *Learning Disability Quarterly*, 37(1), 3–14.
- Cochrane Collaboration (2021). Cochrane methods: Bias: Resources. Retrieved from https://www.researchgate.net/publication/265619755_A_Systematic_Review_of_the_Research_Literature_on_the_Use_of_Phonics_in_the_Teaching_of_Reading_and_Spelling
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.
- Compton-Lilly, C. F., Mitra, A., Guay, M., & Spence, L. K. (2020). A confluence of complexity: Intersections among reading theory, neuroscience, and observations of young readers. *Reading Research Quarterly*, 55(1), 185-195.
- Cunningham, J. W., Spadorcia, S. A., Erickson, K. A., Koppenhaver, D. A., Sturm, J. M., & Yoder, D. E. (2005). Investigating the instructional supportiveness of levelled texts. *Reading Research Quarterly*, 40(4), 410-427.
- Dewitz, P. & Graves, M.F. (2021). The science of reading: Four forces that modified, distorted, or ignored the research finding on reading comprehension. *Reading Research Quarterly*. <https://doi.org/10.1002/rrq.389>
- Fletcher, J. M., Savage, R., & Vaughn, S. (2021). A commentary on bowers and the role of phonics instruction in reading. *Educational Psychology Review*, 33(3), 1249.
- Ferrari, J. (2007, March 1). Labor plan focuses on four key subjects. *The Australian*, p. 4.
- Fountas, I. C. & Pinnell, G. S. (2006). *Teaching for comprehending and fluency: Thinking, talking and writing about reading K-8*. Heinemann.
- Fountas, I. C., Pinnell, G. S. (2013) Guided Reading: The Romance and the Reality. *The Reading Teacher*, 66(4), 268–268.
- Freebody, P. (2007). *Literacy education in school: Research perspectives from the past, for the future*. Camberwell, Victoria: Australian Council for Educational Research. <https://research.acer.edu.au/cqi/viewcontent.cgi?article=1001&context=aer>
- Freebody, P., & Luke, A. (1990). Literacies programs: Debates and demands in cultural context. *Prospect: An Australian journal of TESOL*, 5(3), pp. 7-16.
- Gough, P. & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6-10.

- Goulding, J. (2021). Approaches to historical thinking and learning. In H. Sharp, J. Dallimore, A. Bedford, M. Kerby, J. Goulding, T. Heath, D. Von Guttner, & L. Zarmati, *Teaching Secondary History* (pp. 79-99). Cambridge University Press.
- Government of Western Australia (n.d). Department of Education, First Steps Literacy. Retrieved from <http://det.wa.edu.au/stepsresources/detcms/navigation/first-steps-literacy/>
- Halliday, M. A. K. (1975). *Learning to Mean: Explorations in the Development of Language*. Edward Arnold.
- Hanushek, E. A. (2014). Boosting teacher effectiveness. In C. E. Finn, & R. Sousa (Eds.), *What Lies Ahead for America's Children and Their Schools*. (pp. 23–35). Hoover Institution Press.
- Harris, A. D., McGregor, J. C., Perencevich, E. N., Furuno, J. P., Zhu, J., Peterson, D. E., & Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *Journal of the American Medical Informatics Association*, 13(1), 16-23.
- Hart, N. W. M. (1977). *The Mount Gravatt Developmental Language Reading Program. Teachers Manual. Level 2*. Addison-Wesley.
- Hattie, J. (2003, October). Teachers make a difference: What is the research evidence? Retrieved from http://research.acer.edu.au/research_conference_2003/4/
- Hiebert, E. H., Samuels, S.J., & Rasinski, T. (2012). Comprehension-based silent reading rates: What do we know? What do we need to know? *Literacy Research and Instruction*, 51(2), 110-124.
- Hiebert, E. H. (2005). The effects of text difficulty on second graders' fluency development. *Reading Psychology*, 26(2), 183-209.
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds). (2019). *Cochrane Handbook for Systematic Reviews of Interventions* (2nd ed.). John Wiley & Sons.
- Hofman, J. V. & Pearson, P. D. (2015). Teachers or programs? A historical perspective on where trust is placed in teaching reading. In P. D. Pearson & E. H. Hiebert (Eds.), *Research-based Practices for Teaching Common Core Literacy* (pp. 237–258). Teachers College Press.
- Hoover, W. A. & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160
- Jenkins, J.R., Peyton, J.A., Sanders, E, A.& Vandasy, P.F.(2004) Effects of Reading Decodable Texts in Supplemental First-Grade Tutoring. *Scientific Studies of Reading*, 8(1),53-85
- Johnston, F. R. (2000). Word learning in predictable text. *Journal of Educational Psychology*, 92(2), 248–255.
- Kim, J. S., Capotosto, L. C., Hartry, A., & Fitzgerald, R. (2011). Can a mixed-method literacy intervention improve the reading achievement of low-performing elementary school students in an after-school program? Results from a randomized controlled trial of READ 180 Enterprise. *Educational Evaluation and Policy Analysis*, 33(2), 183-201.

- Lawrence, J. (2006, February 12). Shocking result of our spelling survey, *The Sunday Mail*, p. 11.
- Ledger, S. & Merga, M. K. (2018). Reading aloud: Children's attitudes toward being read to at home and at school. *Australian Journal of Teacher Education*, 43(3). <http://dx.doi.org/10.14221/ajte.2018v43n3.8>
- Livingstone, T. (2006, March 11–12). Reading skills in dramatic slump. *The Courier Mail*, p. 1.
- Lyons, W. & Thompson, S. A. (2012). Guided reading in inclusive middle years classrooms. *Intervention in School and Clinic*, 47(3), 158-158.
- MacPhee, D., Handsfield, L. J., & Paugh, P. (2021) Conflict or Conversation? Media Portrayals of the Science of Reading. *Reading Research Quarterly*. Volume 56(S1), S145–S155.
- Maiden, S., & Harris, T. (2005, February 10). Howard hits back with broadside at 'radical teachers'. *The Australian*, p. 3.
- Mesmer, H. A. (1999). Scaffolding a crucial transition using texts with some decodability. *The Reading Teacher*, 53, 130–142.
- Mesmer, H. A. E. & Griffith, P. L. (2005). Everybody's selling it – But just what is explicit, systematic phonics instruction? *Reading Teacher* 259(4), 366–376.
- Mesmer, H. A. E. (2010). Textual scaffolds for developing fluency in beginning readers: Accuracy and reading rate in qualitatively and levelled decodable text. *Literacy Research and Instruction*, 49(1), 20–39.
- Mesmer, H. A., Cunningham, J. W., Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly*, 47 (3), 235–258
- Morgan, A., Wilcox, B. R., & Eldredge, J. L. (2000). Effect of difficulty levels on second-grade delayed readers using dyad reading. *Journal of Educational Research*, 94(2), 113-119.
- National Institute for Direct Instruction. (2015). *Basic Philosophy of Direct Instruction (DI)* Retrieved from <https://www.nifdi.org/what-is-di/basic-philosophy.html>
- National Institute for Literacy. (2006). *The research building blocks for teaching children to read: Put reading first: Kindergarten through Grade 3*. Retrieved from <https://lincs.ed.gov/publications/pdf/PRFbooklet.pdf>
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Retrieved from <https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf>
- OECD (n.d.). Adult literacy. Retrieved from <https://www.oecd.org/education/innovationeducation/adultliteracy.htm>
- OECD (2000). Reading for Change: Performance and Engagement across countries. Retrieved from <https://www.oecd.org/education/school/programme-for-international-student-assessment-pisa/33690986.pdf>

- OECD (2019). PISA 2018 database. Retrieved from <https://www.oecd.org/pisa/data/2018database/>
- Petscher, Y., Cabell, 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. K. How the science of reading informs 21st-century education. *Reading Research Quarterly*, 0(0), 1–16. <https://doi.org/10.1002/rrq.352>
- Pogorzelski, S. & Wheldall, R. (2018, November 12). Explainer: what's the difference between decodable and predictable books, and when should they be used? Retrieved from <https://theconversation.com/explainer-whats-the-difference-between-decodable-and-predictable-books-and-when-should-they-be-used-106531>
- 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(1) ,39-47
- Primary English Teaching Association Australia (2021a). Approaches to early reading instruction. Retrieved from https://www.petaa.edu.au/w/About/PETAA_position_papers.aspx/#earlyreading
- Primary English Teaching Association Australia (2021b). Selecting texts for the English classrooms. Retrieved from https://www.petaa.edu.au/w/About/PETAA_position_papers.aspx/#selecting
- Rightmyer E. C., McIntyre, E., & Petrosko, J. M. (2006). Instruction, development, and achievement of struggling primary grade readers. *Reading Research & Instruction*, 45(3), 209–242.
- Reading Recovery Australia (2021). Reading recovery Australia: Working in partnership with schools. Retrieved from <https://www.readingrecovery.org.au/>
- Rose, J. (2006, March). Independent review of the teaching of early reading. Retrieved from <https://dera.ioe.ac.uk/5551/2/report.pdf>
- Rowe, K. (2005). *Teaching reading: National inquiry into the teaching of literacy*. Commonwealth of Australia, Department of Education, Science, and Training.
- Sawyer, W. (2004, Spring). Editorial. *English in Australia*, 141(3), 8.
- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for Research in Early Literacy* (pp. 97–110). Guilford Press.
- Seidenberg, M. S., Borkenhagen, M. C., & Kearns, D. M. (2021). Lost in translation? Challenges in connecting reading science and educational practice. *Reading Research Quarterly*, 55(S1), S119–S130.
- Sellars, M. (2017). Schools as institutes of acculturation: A question of belonging. *Turkish Online Journal of Educational Technology*, 917-921.
- Solari, E. J., Terry, N. P., Gaab, N., Hogan, T. P., Nelson, N. J., Pentimonti, J. M., Petscher, Y., & Sayko, S. (2020) Translational Science: A Road Map for the Science of Reading. *Reading Research Quarterly* 55(S1), S347–S360.

- Stuebing, K. K., Barth, A. E., Cirino, P. T., Francis, D. J., & Fletcher, J. M. (2008). A response to recent reanalyses of the national reading panel report: Effects of systematic phonics instruction are practically significant. *Journal of Educational Psychology, 100*(1), 123-134.
- Thiese, M. S. (2014). Observational and interventional study design types; an overview. *Biochemia Medica, 24*(2), 199–210.
- Thomson S. (2021) Australia: PISA Australia—Excellence and Equity? In: Crato N. (eds) *Improving a Country's Education*. Springer, Cham. https://doi.org/10.1007/978-3-030-59031-4_2
- Wilson, E.A. (2007, July 3-6). Critical literacy and inquiring minds: More critical than ever in our text-saturated world. In J. Kiggins, L. Kervin, & J. Mantei (Eds.), *Quality in teacher education: Considering different perspectives and agendas* (pp. 666–675). Proceedings of the 2007 Australian Teacher Education Association National Conference, University of Wollongong.
- THRASS Institute (2019). What is THRASS? Retrieved from <https://www.thrass.com.au/about-thrass/what-is-thrass/>
- Torgerson, C. J., Brooks, G. & Hall, J. (2006). *A Systematic Review of the Research Literature on the Use of Phonics in the Teaching of Reading and Spelling*. Department for Education and Skills, The University of Sheffield. https://www.researchgate.net/publication/265619755_A_Systematic_Review_of_the_Research_Literature_on_the_Use_of_Phonics_in_the_Teaching_of_Reading_and_Spelling
- UNICEF (n.d.). Randomized controlled trials (RCTs). https://www.unicef-irc.org/KM/IE/impact_7.php
- White, C.L. (2009). "What he wanted was real stories, but no one would listen": A child's literacy, a mother's understandings. *Language Arts 86*(6), 431-439.
- Wilson, J. (2005, February 12-13). Kids force-fed a diet of indoctrination. *The Weekend Australian*, p. 18.
- Wiltshire, K. (2006, September 23–24). In defence of the true values of learning. *The Weekend Australian*, pp. 22
- Wood, M. (2005). Progress with pleasure: Success with struggling beginning readers. *New England Reading Association Journal, 41*(2), 30–36, 68.
- Yaden, D. B., Reinking, D., & Smagorinsky, P. (2021). The trouble with binaries: A perspective on the science of reading. *Reading Research Quarterly 56*(S1), S119–S129.