INVESTIGATING EARLY READING IN PRIMARY SCHOOLS SINCE THE INCREASED EMPHASIS ON PHONICS TEACHING

NATALIE JOANNE WALSH

Doctor of Philosophy

ASTON UNIVERSITY

March 2022

Aston University

Investigating Early Reading in Primary Schools since the Increased Emphasis on Phonics Teaching

Natalie Joanne Walsh

Doctor of Philosophy

2022

This thesis addresses three outstanding areas of inquiry regarding the impact of systematic, synthetic phonics teaching following the introduction of the Year 1 Phonics Screening Check in England. These areas of inquiry concern the following key questions: (i) has this increased emphasis on phonics teaching impacted early reading performance and emerging reading difficulties (ii) do phonics-taught early readers have the skills to "self-teach" grapheme-to-phoneme-correspondences? The first study compared two secondary datasets, to determine whether reading performance for different word types, indicated differences between a pre-phonics and post-phonics sample. The post-phonics sample was longitudinal which also enabled an examination of changes in performance for different words over time. The second study utilised the same data, to determine whether emerging profiles of reading difficulty differed between the pre-phonics and post-phonics samples. Stability of these profiles of reading difficulty were also examined. The third study was a novel experimental training study with Reception and Year 1 children (n = 126). The aim was to determine whether beginner readers who were already receiving synthetic phonics teaching could "self-teach" unfamiliar grapheme-to-phoneme-correspondences and generalise these to novel words. Literacy-related skills and the role of context on this learning ability were also examined. Overall, the key findings were that firstly, the post-phonics sample outperformed the pre-phonics sample longitudinally when reading both word types. Additionally, the post-phonics sample displayed fewer profiles of reading difficulty than found within the pre-phonics sample, when allocated with the same measures. Limited stability was found for early reading difficulties, with Exception and Mixed profiles displaying more stability than Nonword profiles. Finally, early readers taught through phonics and with adequate phoneme awareness, letter-sound knowledge and vocabulary, were able to "self-teach" novel grapheme-to-phoneme-correspondences. In conclusion, this thesis contributes new evidence of the beneficial impact of phonics teaching in reducing difficulties and increasing independence in reading.

Keywords: *decoding, phonics, reading, grapheme-to-phoneme-correspondence, self-teaching, reading-development, reading-difficulties, letter-sound-knowledge, phoneme-awareness*

Dedication

This thesis is dedicated to my mother, in loving memory.

Personal Acknowledgements

Firstly, I'd like to acknowledge and give my warmest thanks to my supervisor, Dr Laura Shapiro. Their consistent support, guidance and encouragement over the last six years has made completing this thesis and its included works both possible and an enjoyable experience. From beginning as a Research Assistant in my final undergraduate year and starting to learn the ropes, to finishing this thesis, they have always been willing to offer me their time and expertise. Thank you for all your help, this work would not have been possible without you.

I'd also like to thank my associate supervisor, Professor Adrian Burgess for sharing their knowledge and feedback.

This work would also not have been possible if not for the help of those at Aston University. Including all the Research Assistants who helped during the training study data collection, all the Research Assistants who I worked alongside on the ALP project and everyone from the PGR office who has helped with any questions I've had along the way. Thank you to all of you.

I am also enormously grateful to all the primary schools who agreed to take part in the training study, including all of the wonderful children who I was fortunate enough to work with. Every day of data collection was exciting and rewarding and it was a pleasure to see reading come to life within these primary schools.

Collaborator Acknowledgements

I would also like to thank everyone who collaborated on the research presented in this thesis. In Chapters 2 and 3, data from the Year 1 and Year 4 phases of the Aston Literacy Project (ALP) are reported, a project led by: Dr Anna J. Cunningham, Professor Adrian P. Burgess, Dr Caroline Witton, Professor Joel B. Talcott and Dr Laura R. Shapiro. I was not involved with all phases of data collection, however contributed to the Year 4 data collection and supervision as a Research Assistant. Whilst the ALP dataset has been presented in other works (Cunningham et al., 2021; Cunningham et al., 2015), the research questions addressed and analyses presented within this thesis are original works conducted solely for this thesis and not presented anywhere else.

This thesis also presents published data from the Diagnostic Test of Word Reading Processes (DTWRP) from the Forum for Research in Literacy and Language, of which I was not involved with data collection. I would especially like to thank Dr Jessie Ricketts and Dr Sue Stothard for their help with trying to assemble the original DTWRP dataset, although this was not possible with the time constraints of the programme.

I would also like to thank Dr Jo Taylor for their feedback and guidance regarding the Chapter 4 training study stimuli and design.

Thank you everyone and I hope you enjoy reading the thesis.

Natalie

List of Contents

1 Introduction 14

1.1 Models of Reading 17

1.1.1 The Simple View of Reading 18

1.1.2 Share's Self-Teaching Hypothesis 20

1.1.3 The Dual Route Cascaded Model 21

1.1.4 The Triangle Model 23

1.1.5 Summary of Reading Theories 26

1.2 Measuring Reading Ability 29

1.2.1 Dual-Routes to Reading 29

1.2.2 Word Types 30

1.2.3 Predictions about Word Types according to Reading Models 32

1.3 Important Skills for Reading 35

1.3.1 Phoneme Awareness 37

1.3.2 Letter-sound Knowledge 39

1.3.3 Vocabulary 41

1.3.4 Home Literacy Environment 44

1.4 Reading Difficulties 45

1.4.1 Exception Word Reading Difficulties 49

1.4.2 Reading Interventions 50

1.5 How Reading is Taught in the UK 52

1.5.1 Introduction of the Phonics Screening Check 53

1.5.2 The increased focus on Phonics Teaching 55

1.6 Overview of Thesis 60

2 Study 1a: Reading Performance of Early Readers Post-Phonics Introduction 62

2.1 Introduction and Rationale 62

2.1.1 The Year 1 Phonics Screening Check and Phonics Teaching in England 62

2.1.2 Predictions about how Systematic, Synthetic Phonics Teaching Influences Reading of Different Word Types 64

2.1.3 Changes in Performance for Different Word Types as Reading Develops 65

2.1.4 Research Questions 67

2.2 Method 68

2.2.1 Participants 69
2.2.1.1 ALP Sample 69
2.2.1.2 DTWRP Sample 70
2.2.1.3 Power Calculations 71
2.2.2 Design 71
2.2.3 Measures 72
2.2.3.1 DTWRP Assessment Package 72
2.2.3.2 ALP Assessments 73
2.2.3.3 Validity Check for Nonword versus Exception word measures on the DTWRP (ALP Data) 73
2.2.4 Procedure 74
2.2.4.1 ALP Procedure 74
2.2.4.2 DTWRP Sample Procedure 76
2.3 Results 76
2.3.1 Data Analysis Strategy and Exploration of Data 76
2.3.2 Question 1: ALP Word Type Reading Performance (Raw Scores) 78
2.3.3 Question 1: ALP Word Type Reading Performance (Stanine Scores) 81
2.3.4 Question 2: Word Reading Performance between Pre and Post Phonics Samples (ALP and DTWRP Sample Data) 86
2.4 Discussion 88
2.4.1 Conclusions and next steps 92
3 Study 1b: Reading Difficulty Profiles in Early Readers Post-Phonics Introduction 94
3.1 Introduction and Rationale 94
3.1.1 Reading Difficulties along Dual-Routes to Reading 94
3.1.2 Reading Difficulties and Phonics Teaching 96
3.1.3 Stability of Word Reading Difficulties 98
3.1.4 Research Questions 101
3.2 Method 102
3.2.1 Participants 103
3.2.2 Power Calculations 103
3.2.3 Design 103
3.2.4 Measures 104
3.2.4.1 Validity Check for Nonword and Exception Profiles within the ALP Sample determined through DTWRP classifications 105

3.2.5 Procedure 107

3.3 Results 107

3.3.1 Data Analysis Strategy and Exploration of Data 107

3.3.2 Question 1: Profiles of Reading Difficulty within the ALP Sample determined through DTWRP classifications 109

3.3.2.1 Question 1: Comparison of the Risk of Nonword versus Exception Profiles within the ALP Sample 110

3.3.3 Question 2: Profiles of Reading Difficulty within the ALP sample compared to the DTWRP sample using DTWRP classifications. 112

3.3.4 Question 1: Profiles of Reading Difficulty within the ALP sample determined through Z-Score classifications 113

3.3.5 Question 3: Longitudinal Stability of Reading Difficulty Profiles within the ALP sample 115

3.3.5.1 Longitudinal Stability of DTWRP Classification Profiles within the ALP Sample 115

3.3.5.2 Inferential test of the Longitudinal Stability of DTWRP-classified Profiles of Reading Difficulty 116

3.3.5.3 Longitudinal Stability of Z-Score Profiles, classified within the ALP Sample 117

3.3.5.4 Inferential test of the Longitudinal Stability of Z-score-classified Profiles of Reading Difficulty 119

3.3.5.5 Overall Stability of ALP Word Type Reading Performance over Time (ALP Z-Score Data) 120

3.4 Discussion 122

3.4.1 Conclusions and next steps 127

4 Study 2: Can Early Readers Independently Learn Novel GPCs Post-Phonics Instruction? 129

4.1 Introduction and rationale 129

4.1.1 Early Reading and Phonics Instruction 129

4.1.2 Important Skills for Self-Teaching 131

4.1.3 The Role of Context 131

4.1.4 Research Questions 133

4.2 Method 135

4.2.1 Participants 135

4.2.2 Power Calculations 135

4.2.3 Design 136

4.2.4 Measures 137

4.2.4.1 GPC Screener 137

4.2.4.2 GPC Knowledge: LeST (Letter Sound Test) ... 138
4.2.4.3 Oral Vocabulary ... 138
4.2.4.4 Home Literacy Environment ... 138
4.2.4.5 YARC Phoneme Awareness ... 139
4.2.4.6 Training Word Stimuli ... 140
4.2.4.7 Generalisation Word Stimuli ... 140
4.2.5 Procedure ... 142
4.2.5.1 Pre-Test Session ... 142
4.2.5.2 Training Sessions ... 143
4.2.5.3 Post-Test Session ... 144
4.3 Results ... 146
4.3.1 Data Analysis Strategy and Exploration of Data ... 146
4.3.2 Question 1: Does Reading Performance improve across Training Sessions? (Training Data) ... 147
4.3.3 Question 1: Did Early Readers learn GPCs and apply this knowledge to Generalisation Items? (Post-Test Data) ... 148
4.3.4 Question 2: Which literacy-related skills effect GPC learning across Training Sessions? (Training Data) ... 150
4.3.5 Question 2: Which literacy-related skills effect GPC reading and generalisation at Post-Test (Post-Test Data) ... 153
4.3.6 Question 3: Are early readers better able to generalise from real words or nonwords, using new GPCs learnt from whole words? ... 158
4.4 Discussion ... 158
4.4.1 Conclusions and next steps ... 162
5 General Discussion ... 164
5.1 Summary of findings ... 164
5.2 Theoretical implications of these findings ... 166
5.2.1 Theoretical reading development ... 174
5.3 Educational implications ... 176
5.4 Potential limitations and next steps ... 178
5.5 Conclusions ... 183
References ... 185
Appendices ... 201
Appendix 1 ... 201
Appendix 2 ... 203

Appendix 3 204
Appendix 4 207
Appendix 5 208
Appendix 6 209
Appendix 7 212

List of Abbreviations

ALP: Aston Literacy Project

DRC: Dual Route Cascaded Model

DTWRP: Diagnostic Test of Word Reading Processes

FSM: Free School Meals

GPC: Grapheme-to-Phoneme Correspondence

PSC: Phonics Screening Check

SEN: Special Educational Needs

SES: Socioeconomic Status

SSP: Systematic synthetic phonics

SVR: Simple View of Reading

List of Tables

Table 1: ALP Participants across Year 1 and Year 4 70
Table 2: Spearman's Rho Correlation Matrix for DTWRP and ALP Choice Assessments 74
Table 3: Percentage of Stanine Scores achieved within a Normal Distribution 82
Table 4: Difficulty Profile Median Scores on Orthographic and Phonological Choice at Year 4 (ALP Data) 106
Table 5: Reading Difficulty Profiles in the ALP Year 1 Sample using DTWRP Classifications 110
Table 6: Reading Difficulty Profiles in the ALP Year 4 Sample using DTWRP Classifications 110
Table 7: ALP Year 1 McNemar Input 111
Table 8: ALP Year 4 McNemar Input 111
Table 9: Reading Difficulty Profiles within the ALP sample compared with the DTWRP Standardisation Sample using DTWRP Classifications 113
Table 10: Reading Difficulty Profiles within the ALP Year 1 Sample using within-sample Classifications 114
Table 11: Reading Difficulty Profiles within the ALP Year 4 Sample using within-sample Classifications 115
Table 12: Matrix of Longitudinal Reading Difficulty Profiles within the ALP Sample with DTWRP Classifications 116
Table 13: Cohen's Kappa Results for Reading Difficulty Profiles within the ALP Sample using DTWRP Classifications 117
Table 14: Matrix of Longitudinal Reading Difficulty Profiles within the ALP Sample with Z-Score Classifications 118
Table 15: Cohen's Kappa Results for Reading Difficulty Profiles within the ALP Sample using Z-Score Classifications 120
Table 16: Variance Components Analysis of ALP Z-Scores 121

List of Figures

Figure 1.1: The Simple View of Reading (Gough & Tunmer, 1986) 19
Figure 1.2: The Dual Route Cascaded Model (Coltheart et al., 2001) 22
Figure 1.3: The Triangle Model of Reading (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) 25
Figure 2.1: ALP Year 1 Nonword Reading Raw Score Distribution 79
Figure 2.2: ALP Year 1 Exception Word Reading Raw Score Distribution 79
Figure 2.3: ALP Year 4 Nonword Reading Raw Score Distribution 80
Figure 2.4: ALP Year 4 Exception Word Reading Raw Score Distribution 81
Figure 2.5: Normal Distribution for Age Standardised Stanine Scores 82
Figure 2.6: ALP Year 1 Nonword Reading Stanine Score Distribution 83
Figure 2.7: ALP Year 1 Exception Word Reading Stanine Score Distribution 84
Figure 2.8: ALP Year 4 Nonword Reading Stanine Score Distribution 85
Figure 2.9: ALP Year 4 Exception Word Reading Stanine Score Distribution 85
Figure 2.10: ALP and DTWRP Year 1 Nonword and Exception Reading Regression Slopes 86
Figure 2.11: ALP and DTWRP Year 4 Nonword and Exception Reading Regression Slopes 87
Figure 4.1: Single Whole-Word Reading Performance across Training Sessions 148
Figure 4.2: Post-Test Reading for Trained and Generalisation Real Word and Nonword Items 149
Figure 4.3: Profile Plot of Single Whole-Word Reading Scores by Screener Score 151
Figure 4.4: Single Whole-Word Reading Scores by YARC Phoneme Deletion Score 152
Figure 4.5: Nonword and Real Word Reading across Phoneme Awareness Scores 152
Figure 4.6: Post-Test Mean GPC Reading Total Score by Phoneme Awareness Scores 154
Figure 4.7: Post-Test Real Word and Nonword Reading by Screener Score 155
Figure 4.8: Post-Test Trained and Generalisation Item Reading by GPC Screener Score 156
Figure 4.9: Post-Test Trained and Generalisation Reading scores by LeST Scores 156
Figure 4.10: Post-Test Trained and Generalisation Item Reading Scores by Vocabulary Scores 157
Figure 4.11: Post-Test Trained and Generalisation Item Reading Scores by YARC Phoneme Deletion Scores 157

1 Introduction

In this thesis, I investigate early reading performance in English primary school aged children since the introduction of systematic synthetic phonics teaching, as part of the England primary school curriculum following the introduction of the 2012 Phonics Screening Check. This includes a key focus on examining children's word type reading performances and profiles of reading difficulty to determine whether there is evidence that children who have received systematic, synthetic phonics teaching display different levels of performance and reading difficulty. Additionally, this investigation into word type reading performance and profiles of reading difficulty also determines whether these early readers demonstrate qualitatively different types of difficulty since the introduction of mandatory phonics teaching in England. Another key focus of this thesis is to determine the extent to which children can learn new GPCs (grapheme-to-phoneme correspondences) independently, after receiving phonics teaching, since independence and self-teaching is a key aim of the phonics method. Additionally, the role of literacy-related skills within this GPC "self-teaching" ability are examined.

Phonics teaching has a clear theoretical basis, and an established evidence base of effectiveness with early readers, however, we are lacking detailed analysis of how children's reading has changed as a result of introducing phonics as the mainstream approach to reading instruction. In particular, there is little evidence which confirms the result of systematic, synthetic phonics teaching on the word type reading ability of early readers. Specifically, how reading varies amongst different word types, such as nonwords, exception words and regular words, as a result of phonics instruction. Theoretically, as discussed within the sections below, phonics targets "decoding" ability amongst early readers and strengthens their Non-lexical pathway to reading. Therefore, this should provide an advantage for word types which require the role of decoding, such as nonwords. While there is evidence from national tests such as the Phonics Screening Check to indicate that phonics teaching facilitates nonword reading (Stainthorp, 2022), there is little existing literature comparing word type reading amongst early readers who have received systematic, synthetic phonics instruction. Additionally, there is also little known about the interaction between systematic, synthetic phonics teaching and word types which cannot be read through decoding alone, such as exception words. This gap within existing knowledge is addressed through Study 1a, which explores the longitudinal reading performance of a sample of early readers in England, across two differing word types: nonwords and exception words. This study also compares the performance of this sample with a pre-phonics sample gathered before the introduction of the mandatory Phonics Screening Check in 2012, with the same standardised measures of word type reading.

Another research area of interest in relation to phonics teaching, is the potential impact systematic, synthetic phonics teaching may have on profiles of reading difficulty, amongst early readers. Much of the existing literature regarding reading difficulties amongst early readers, has focused on children with classifications of Dyslexia, rather than readers within mainstream classrooms. Whilst this knowledge is valuable for designing reading interventions and support for children with Dyslexia, the notion of different “subtypes” or “profiles” of reading difficulty can also be applied to mainstream poor readers. These early readers who fall behind the expected reading performance for their age group may require tailored intervention and support, but to facilitate this, the frequency and type of reading difficulties they display must be investigated. While there is evidence to suggest that these “profiles” or “subtypes” exist outside of Dyslexia classifications (Talcott et al., 2013; Wang et al., 2014) there is little existing evidence to determine the impact that systematic, synthetic phonics teaching has had on the profiles of reading difficulty displayed amongst mainstream poor readers. From a theoretical perspective, phonics should reduce word reading difficulties related to decodable words, as it strengthens the Non-lexical pathway to reading from a dual-route reading perspective. Therefore, early readers should display less profiles of reading difficulty in relation to decodable nonwords. In contrast, phonics teaching is not expected to directly train the Lexical pathways to reading from a dual-route perspective, however through Share’s Self-Teaching Hypothesis (1995), it is possible that this Non-lexical training could facilitate the lexicalisation process and thus, the development of the Lexical pathways to reading, indirectly. Alternatively, mainstream poor readers may display reading difficulties with non-decodable words, specifically exception words, due to phonics teaching providing no training along the Lexical pathways to reading. Therefore, it is important to determine the impact that systematic, synthetic phonics teaching has had on profiles of reading difficulty amongst mainstream poor readers. This gap in knowledge is addressed through Study 1b, which investigates the profiles of reading difficulty found within a sample of post-phonics mainstream poor readers, both in regard to their frequency and longitudinal stability. This post-phonics sample was also compared to a pre-phonics sample of mainstream poor readers, to investigate if profiles of reading difficulty differed between the two groups.

One additional consideration is that systematic, synthetic phonics teaching may provide additional benefits to early readers, as part of training the Non-lexical pathway to reading. Mainstream primary schools within England vary on the number of grapheme-to-phoneme correspondences that are explicitly taught to early readers, but little is known regarding how early readers transition from reading with a limited set of taught grapheme-to-phoneme correspondences, to eventually learning novel grapheme-to-phoneme correspondences which are not taught to them. Computational models such as the GPC-LM (Pritchard et al., 2016) and studies with young readers (Apfelbaum et al., 2013) have shown that learning novel grapheme-to-phoneme correspondences from whole words is possible, without explicit phonics teaching. Nevertheless, there is little existing literature which examines the role of phonics teaching on the ability to detect, learn and generalise novel grapheme-to-phoneme correspondences from whole words. Therefore, research is required which investigates if phonics teaching provides early readers with the additional advantage of "self-teaching" novel grapheme-to-phoneme correspondences from whole words, facilitating this transitional period in reading development. This gap in knowledge is addressed through Study 2, which investigates grapheme-to-phoneme "self-teaching" ability within a sample of mainstream early readers who have received initial systematic, synthetic phonics instruction. As part of this research, the role of literacy-related skills which may contribute to early reading development and grapheme-to-phoneme correspondence "self-teaching" are investigated. The role of context within grapheme-to-phoneme correspondence "self-teaching" is also investigated, to provide insight into appropriate teaching strategies and resources for early readers.

These three studies have both practical and theoretical contributions to add to the existing literature. The novel findings from Study 1a and Study 1b provide support for the dual-route reading perspective and the developmental account of reading by Share (1995). Whilst not a developmental account of reading, the results from these studies also reflect the theoretical routes proposed in the Dual Route Cascaded Model of reading (Coltheart et al., 2001). Word type reading results indicate that early readers utilise the Non-lexical pathway to phonologically decode novel words, shown by an initial nonword reading advantage. Over time, Lexical pathways to reading develop and exception word reading improves whilst nonword reading declines, as early readers transition from relying on the Non-lexical pathway to reading, to the Lexical pathway. Similar findings are reflected in Study 1b regarding profiles of reading difficulty amongst early readers. Firstly, profiles of nonword reading difficulty appear to be addressed through early phonics teaching strengthening this Non-lexical pathway, whilst a greater number of children initially demonstrate difficulties with exception word reading. Over time as the Lexical pathway develops through independent reading and text exposure, facilitated through "self-teaching" (Share, 1995), profiles of exception word reading difficulty decline. An overall lack in longitudinal stability within profiles of nonword reading difficulty has practical implications for measures such as the Phonics Screening Check, as this cannot be used as a predictive measure of longitudinal reading abilities or difficulties. Meanwhile the fair to moderate stability of profiles of exception and mixed word reading difficulties indicates that systematic synthetic phonics teaching may not be addressing all types of word reading difficulty within mainstream classrooms. These findings also support the theoretical basis for phonics teaching, through providing empirical support for the underlying mechanisms of phonics (decoding and reading independence) and the impact that this teaching has on word type reading abilities and profiles of difficulty.

The novel findings from Study 2 provide support for the Self-Teaching Hypothesis by Share (1995) but also extends the theoretical hypothesis. Specifically, the empirical evidence from Study 2 indicates that early readers can "self-teach" themselves novel grapheme-to-phoneme correspondences from whole words, following existing systematic, synthetic phonics teaching. Demonstrating that early phonics teaching and this "self-teaching" ability can expand access to independent reading and facilitate this transitional period of reading development. Study 2 also provides wider support for the role of literacy-related skills contributing to this "self-teaching" ability and early reading development, namely: phoneme awareness, letter-sound knowledge and vocabulary. These findings have practical implications for mainstream teaching, as these literacy-related skills which actively contribute to both successful phonological decoding and this "self-teaching" ability should be included as part of classroom teaching practices to facilitate early reading development. Additionally, the role of context was found to not significantly impact this "self-teaching" ability, which contrasts with the existing literature (Landi et al., 2006; Stuart et al., 2000) and has implications for the use of reading resources which include contextual information. Nonetheless, this finding may have been impacted due to the limited contextual design within Study 2 which is discussed further within Chapter 4.

Making causal conclusions are difficult when phonics is now ubiquitous across England, for example, the difficulty in conducting a randomised control trial where the early readers have had no exposure to phonics teaching. The aim of these three studies is to use detailed analysis of patterns of performance and reading difficulty to get one step closer to understanding the way in which phonics teaching works to strengthen Non-lexical reading processes and the consequences this may have for word type reading, longitudinal reading difficulties and grapheme-to-phoneme self-teaching abilities amongst early readers.

Below, I outline the key theoretical perspectives on the process of reading and how this develops. These theories provide a key theoretical rationale for the decision to introduce phonics teaching in early primary school through the National Curriculum.

1.1 Models of Reading

For skilled, fluent, adult readers, reading has largely become an automatic process. As skilled readers, we can quickly derive meaning from written language, which does not require us to be aware of the complex processes that enable skilled reading (Stuart & Stainthorp, 2015). Learning to read is a remarkable process which combines different skills discussed throughout this thesis. For a child novice reader to complete the journey into a skilled adult reader, there is a complex learning process. For a child to become a successful reader, they must master the alphabetic principle, letter-sound correspondences (Ehri, 2005) and eventually recognise whole words and derive their meaning without relying on translating the word into its sounds (Stuart & Stainthorp, 2015). Reading development is also complicated by various routes to reading, such as those proposed by Ehri (2005), who suggested that there are four ways in which written text can be read; through phonological recoding, analogizing, prediction and sight reading. For skilled reading to occur, the novice reader must master all of these routes and their associated skills, which is no simple task.

Theories of reading development provide accounts of how a child develops from a novice reader with no reading experience, through to a fluent adult reader. Before the role that phonics teaching plays in reading development can be examined, the theories which set the broader context for the introduction of phonics teaching must be discussed. Some of these key theories of reading are discussed below.

1.1.1 The Simple View of Reading

In 1986, Gough and Tunmer proposed that successful reading comprehension consists of two factors which form the Simple View of Reading (SVR). The SVR is not explicitly a theory of reading development and is used instead as a framework for reading comprehension, it does describe two components that are crucially important for children to successfully learn to read (Castles et al., 2018). The first factor is the ability to decode text. The definition of this "decoding" is limited in the SVR, however it is often associated with "sounding-out" a word through its letter-sound correspondence rules, which Gough and Tunmer considered "only a primitive form of decoding" (Gough & Tunmer, 1986, p. 7). In Hoover and Gough (1990), it is suggested that this decoding could happen at whole-word level rather than individual letter sounds. As the authors stated, "decoding is simply efficient word recognition", which leads to "the retrieval of semantic information at word level" (Hoover and Gough, 1990, p. 130). Using this information, it is possible that the decoding factor also includes reading by sight, in addition to relying on letter-sound correspondence rules.

The second factor in the framework is linguistic comprehension, which is the ability to understand sentences and spoken language (Gough & Tunmer, 1986). The authors stated that if a child can decode text, but not comprehend the words, they are not reading with comprehension. Meanwhile, a child with good linguistic comprehension cannot read without first decoding the text (Castles et al., 2018). As both factors are needed, the authors present the SVR as multiplicative using the RC = D x LC formula, as neither decoding nor linguistic comprehension alone is sufficient for reading comprehension (Nation, 2019).

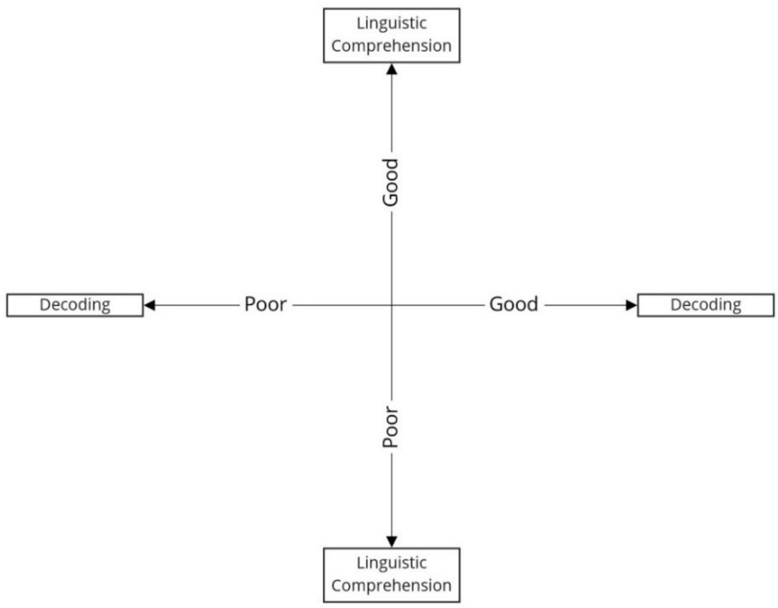
Supporting evidence for the SVR has been found in research such as Hjetland et al. (2019) who found that language comprehension and decoding accounted for 99.7% of the variance in reading comprehension amongst 189 seven-year-old, Norwegian speaking children. Lervåg and Aukrust (2010) found that reading comprehension skills of young first and second language learners in Norwegian were predicted by both vocabulary and decoding skills. While Gustafson et al. (2013) found that the two factors of language comprehension and decoding explained different levels of variance in the reading ability of 36 children with typical reading ability, compared to 36 children with reading difficulties, in fourth grade Swedish schools. The two factors explained 37% of the reading ability variance in typical readers and explained 30% of the reading ability variance in children with reading difficulties (Gustafson et al., 2013). With both factors demonstrated to be significant predictors of reading ability across both groups.

Kendeou et al. (2009) found that in a factor analysis of 221 English speaking children from the US and Canada (113 four-year-olds and 108 six-year-olds), listening comprehension and decoding were found to be distinct components of reading comprehension. Providing further support for the SVR framework.

One of the strengths of the SVR is that the framework reflects how early readers can show variation in their ability to decode text independently from their ability to comprehend spoken language (Nation, 2019), as shown in Figure 1.1 below. Early readers can be placed along two spectrums for decoding and linguistic comprehension ability, which each cause individual variance in reading comprehension (Nation, 2019). Specifically, poor decoding ability provides a barrier to reading comprehension, as children will not have access to the full text, even if their linguistic comprehension ability is good. Conversely, poor linguistic comprehension will limit children's understanding of the text, even if they are able to decode the words accurately.

Figure 1.1

The Simple View of Reading (Gough & Tunmer, 1986)



Note. Reproduced from Castles, Rastle & Nation, 2018.

The SVR is not a model with specified stages of development, lacking information on how these processes develop over time (Castles et al., 2018). Furthermore, the SVR does not clarify whether decoding is only sublexical (letter-to-sound), or also includes lexical (whole-word) knowledge (Chang et al., 2020). As recommended by Castles et al. (2018), to understand reading development, detailed models outlining the underlying cognitive processes should be examined.

1.1.2 Share's Self-Teaching Hypothesis

In contrast with the SVR, Share's Self-Teaching Hypothesis (1995) explicitly describes the process of development, focusing on how children learn to decode text. This account focuses solely on the first factor in the SVR: decoding, and describes this process in detail, focusing on what Share describes as "phonological recoding". During the process of phonological recoding, words are read through letter-by-letter identification where the new reader matches each letter to their corresponding sound (Share, 1995). These mappings between the orthography and phonology are often known as grapheme to phoneme correspondences (GPCs), referred to by Share (1995) as "spelling-sound relationships". Once early readers have acquired a range of different GPCs, they can then use this knowledge when reading through phonological recoding; hereby sounding letters out and blending these together to pronounce a novel word (Share, 1995). Share (1995) hypothesised that through each phonological recoding of a word, the reader developed a "lexicalization" of the recoding process. During this process, GPCs are altered through growing orthographic knowledge, for example, positional letter effects. For example, the "*ea*" sound in the word "*bead*" is different from the "*ea*" sound in the word "*thread*". Combined with an expanding print lexicon, these developments lead to skilled word reading. In this case, the phonological recoding process is a self-teaching mechanism, as it provides the reader with decoding skills which then leads to orthographic knowledge and the ability to acquire a print lexicon without additional support (Share, 1995). This growing orthographic knowledge is then available when reading the previously encountered word in the future, lessening the reliance on phonological decoding (Castles et al., 2018).

Studies such as Cunningham et al. (2002) and Wang et al. (2013) have provided evidence for Share's (1995) hypothesis, through demonstrating that phonological decoding accuracy and orthographic learning and orthographic knowledge were highly correlated with each other, in young readers. This can be interpreted as children who are successful phonological decoders through using their GPC knowledge have better orthographic knowledge than other children, due to their ability to "self-teach" using these skills. Ricketts et al. (2011) also found that target decoding was the strongest predictor of orthographic learning amongst a sample of 88 seven- to eight-year-old children. Whilst Conners et al. (2011) found that amongst 40 seven- to nine-year-old children, the relationship between phonological recoding and word identification was significantly mediated by orthographic knowledge. Relating this back to Share's (1995) hypothesis, children who are more successful at word identification may have greater orthographic knowledge, which they gained through their ability to "self-teach" through phonological recoding.

Share's (1995) hypothesis outlines how early readers begin the journey to becoming independent readers, through acquiring orthographic knowledge through experience with different texts, which are read through phonological recoding. Whilst there is evidence outlined above to support Share's hypothesis, the hypothesis is theoretical and has limited implementation in computational models, such as Pritchard et al. (2018). Nevertheless, it has inspired other models, such as Grainger et al. (2012). Specifically, as early readers "self-teach" themselves through reading, they gradually rely less and less on phonologically decoding words alone as orthographic representations develop (Grainger et al., 2012).

1.1.3 The Dual Route Cascaded Model

In 2001, Coltheart, Rastle, Perry, Langdon and Ziegler described a computational model of reading known as the Dual Route Cascaded Model (DRC). Although the DRC was originally developed to explain the process of skilled, adult reading, the model is widely cited in the reading development literature and the processes are assumed to be similar in child readers. The DRC proposes that there are three routes to word reading through translating orthography to phonology, consisting of interacting layers with multiple units (Coltheart et al., 2001). As shown in Figure 1.2 below, the connections between layers are either excitatory or inhibitory. Excitatory connections result in an activated unit contributing to activation of other units, whereas inhibitory connections result in an activated unit inhibiting activation of other units to rise (Coltheart et al., 2001).

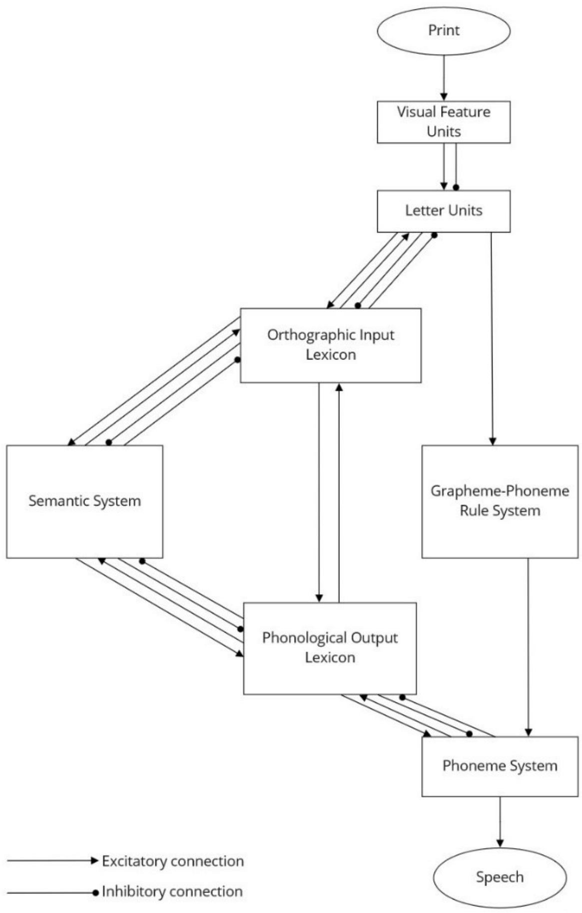
The routes of the DRC are named as follows: the GPC route, the Lexical Nonsemantic route and the Lexical-Semantic route. The GPC route proposes that when reading a printed word, firstly the visual feature units are activated, where each unit represents one of a letter's features (Ziegler et al., 2000). This excitatory connection activates letter units, with each unit representing one letter of the alphabet (Ziegler et al., 2000). This route then converts the letter string through grapheme-phoneme correspondence (GPC) rules into a phoneme string, leading to verbal pronunciation (Coltheart et al., 2001). The DRC learns the correct pronunciation of a word through attempting a pronunciation, for example, through the GPC route, and receiving the word's correct pronunciation as feedback (Coltheart et al., 2001).

The Lexical Nonsemantic route begins with the same visual feature and letter unit activation as the GPC route, however once these letter units have been activated, these letters activate the word's entry in the orthographic lexicon (Coltheart et al., 2001). This orthographic lexicon activation then leads to activation in the phonological output lexicon, where the phonological representation for the word is found, through activating the word's phonemes (Coltheart et al., 2001). Unlike the GPC route, this route does not use individual GPC rules to access the phonology of the word. Instead mapping the orthographic form of known whole words, to their corresponding phonological form (Taylor et al., 2013).

The Lexical Semantic route begins with the same visual feature and letter unit activation as the above two routes. Similar to the Lexical Nonsemantic route, this activation leads to the activation of the word's entry in the orthographic lexicon (Coltheart et al., 2001). As shown in Figure 1.2 below, this then causes activation in the semantic system, leading to phonological lexicon activation, through which the whole word phonological representation is retrieved. However, this semantic pathway is presented with limited detail in Coltheart et al. (2001) as the semantic element of the model had not been implemented.

Figure 1.2

The Dual Route Cascaded Model (Coltheart et al., 2001)



Note. Reproduced from Coltheart, Rastle, Perry, Langdon & Ziegler, 2001.

Support for the DRC has been found through computational modelling in reviews such as Coltheart (2006). Subsequent studies such as Coltheart et al. (2010) and Nickels et al. (2008) have used computational cognitive neuropsychology to examine how lesioned DRC models correspond to patterns of reading difficulty shown by real world participants with dyslexia and semantic dementia with varying results. Additionally, studies such as Peterson et al. (2013) and Ziegler et al. (2008) have also applied the DRC model to real world data of participants with dyslexia subtypes. Additionally, the serial processing procedure of the DRC has been supported by studies such as Mulatti et al. (2007) and Spencer (2007).

A strength of the DRC is that it provides a detailed computational model which demonstrates the different routes to reading. Additionally, the model is organised so that the different routes operate in parallel to each other. The consequences of this parallel organisation are firstly, that a combination of these routes may be used, such that the pronunciation of a word could be jointly determined through both routes (Rayner & Reichle, 2010). Through this parallel route and unit propagation design, the model also accounts for words that are frequently encountered in written text. In this case, activation is faster among units which are frequently encountered, therefore frequently encountered words are pronounced more accurately and rapidly, than novel words which require the slower GPC route (Rayner & Reichle, 2010).

One limitation of the DRC is the limited detail regarding the third Lexical Semantic route, and the limited detail regarding the role that semantics contributes to reading development. A second limitation of the DRC is that the model is a static model of adult reading, rather than a developmental model (Castles et al., 2018). As the DRC is not explicitly a developmental model, the model does not explain how GPC rules are learnt in early readers or how lexical representations of words are formed (McKague et al., 2001).

1.1.4 The Triangle Model

The Triangle Model of Reading is a connectionist, computational model of reading with multiple implementations, from the early model of Seidenberg and McClelland (1989) through to Plaut et al. (1996) and more recent work of Harm and Seidenberg (2004). The Triangle Model uses three sets of simple processing units: the orthography set which is a bank of grapheme units, the phonology set which is a bank of phoneme units and the semantic units set (Powell et al., 2006). As shown in Figure 1.3 below, there are additional hidden layers (represented by blank boxes in Figure 1.3) which mediate the three pathways between the processing units (Plaut et al., 1996). In the Triangle Model, a written word's pronunciation is produced through propagating activation from these orthographic and semantic sets of units along connections to units representing phonological output (Rayner & Reichle, 2010). Unlike the DRC model (Coltheart et al., 2001), the Triangle Model uses a set of input-to-output connections to identify target words, with specific patterns of activity across units representing orthographic input and phonological output (Rayner & Reichle, 2010). When compared to the DRC model (Coltheart et al., 2001), lexical information in the Triangle Model is contained in the connections which mediate orthographic input and phonological output, instead of discrete units in a lexicon proposed by the DRC (Rayner & Reichle, 2010).

Despite the Triangle Model displaying three sets of units, it is considered a dual-route model, in that there are two pathways to reading aloud and identifying the meaning of words, either directly or indirectly (Coltheart, 2005). The Triangle Model as shown in Figure 1.3 below, outlines two distinct routes to the pronunciation of a written word, the first route being the orthography to phonology route, considered the direct route. In contrast, the second route moves from orthography to semantic representations, then to producing the phonology of the word, considered the indirect route (Coltheart, 2005).

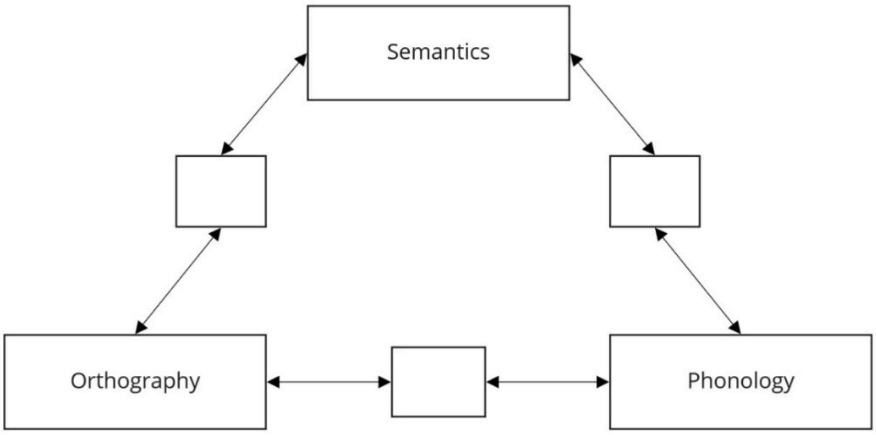
As described by Powell et al. (2006), when a target word is presented to the Triangle Model network, it is presented as a pattern of activity across grapheme units which propagates through the network. In the direct pathway, this activation results in a pattern of activity across the phoneme units which produces the target word pronunciation (Powell et al., 2006). The Triangle Model learns this relationship between orthography and phonology through feedback (Taylor et al., 2013). After attempting a pronunciation, the model receives the correct pronunciation as feedback, using this information to modify the strength of connections between units (Taylor et al., 2013).

Through experience with language, the Triangle Model learns relationships between each of the orthographic, semantic and phonological representations (Chang & Monaghan, 2019). This learning through experience also allows the model to learn context sensitive orthography to phonology activation mappings, as the model encodes how often a letter surrounded by other specific letters is pronounced in a particular way (Taylor et al., 2013).

The Triangle Model also outlines two distinct routes from orthography to semantics as shown in Figure 1.3 below, the first being the direct orthography to semantics route and the second route indirectly mediated through phonology (Coltheart, 2005).

Figure 1.3

The Triangle Model of Reading (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989)



Note. Reproduced from Castles, Rastle & Nation, 2018.

The Triangle Model, whilst simulating adult reading, also allows predictions to be derived about how children learn to read (Castles et al., 2018). Before the model is trained, the resources of the network are allocated to building connections between orthography and phonology in the direct route (Bishop & Snowling, 2004). This is similar to how early readers begin learning to read according to the DRC (Coltheart et al., 2001) and Share's Self-Teaching Hypothesis (1995): through focusing on grapheme to phoneme correspondences and reading through phonological recoding. As training is applied to the Triangle Model, the network relies less on the direct orthography to phonology mappings and begins to utilise mappings from the orthography to semantic to phonology indirect route (Bishop & Snowling, 2004). Once again this is consistent with how reading develops in the DRC (Coltheart et al., 2001) and Share's Self-Teaching Hypothesis (1995): as readers gain orthographic knowledge through experience with text, this allows them to read via whole-word orthographic representations rather than via phonological recoding. In an implementation of the Triangle Model by Harm and Seidenberg (2004), it was found that the direct orthographic to semantic pathway took longer to develop than the indirect orthographic to phonology to semantic route. However, when this direct orthographic to semantic pathway was established, it had a significant speed advantage compared to the indirect route (Harm & Seidenberg, 2004).

The Triangle Model has received support through experimental work with typical adult readers (Woollams et al., 2016), dyslexic readers (Woollams et al., 2007), partially with child readers (Nation & Cocksey, 2009) and has also been successfully applied to computational modelling (Chang & Monaghan, 2019; Monaghan et al., 2017; Smith et al., 2021).

One limitation of the Triangle Model is that some of the implemented models have produced results that contrast with the performance of early readers, potentially due to the way orthography is represented in the model (Nation, 2019). Powell et al. (2006) found that in an implementation of the Plaut et al. (1996) Triangle Model, the network did not simulate reading errors made by early readers with a mean age of 4 years and 10 months. Compared to the early readers, the network was also slower to read nonwords (Powell et al., 2006). Powell et al. (2006) then adapted the network to include grapheme-to-phoneme correspondences during the training, to simulate the explicit correspondences often taught to early readers, before whole words were shown to the network. The network then produced nonword reading performance which closely represented the performance shown by early readers (Powell et al., 2006).

Additionally, the Triangle Model utilises external feedback to strengthen connections between units, however one limitation with this feedback process is that the model is unable to account for the internally generated feedback that young children appear to use when reading (McKague et al., 2001). As discussed by Harm and Seidenberg (2004) the feedback procedure used in their model implementation provided explicit semantic and phonological feedback. However, this may not reflect the experience of early readers who receive more abstract feedback, such as listening to their own pronunciations (Harm & Seidenberg, 2004).

1.1.5 Summary of Reading Theories

The sections above have outlined key theories regarding the process of reading across development, from a novice early reader through to a skilled fluent reader. As stated by Rayner and Reichle (2010) it should be highlighted that these are models and frameworks are limited to outlining the process of reading aloud or reading single words, rather than the all-encompassing process of "reading" which includes fluent reading of whole sentences and reading comprehension.

In the case of Share (1995), the development of reading is explicitly addressed: early readers begin reading through phonological recoding, which with repeated exposures and successful decoding, leads to orthographic knowledge and orthographic representations. Resulting in less phonological recoding with familiar words over time, leading to faster fluent reading (Share, 1995).

The SVR (Gough & Tunmer, 1986) although not a developmental model, has some similarities with Share (1995). Specifically, the SVR highlights the importance of the ability to decode text as part of reading, both at a sublexical and whole-word level. As linguistic comprehension alone without the ability to decode text, is not enough for successful reading comprehension (Gough & Tunmer, 1986). The SVR is also the only framework described here that addresses comprehension processes, with the other models describing the processes of single word reading. In addition, the SVR has a broader definition of decoding when compared to Share (1995) and the DRC (Coltheart et al., 2001). The SVR's decoding could refer to phonic recoding as discussed by Share (1995), or it could refer to either the GPC route or the whole-word Lexical Nonsemantic route in the DRC. In other words, "decoding" as defined by the SVR encompasses any means of reading a single word, which could be based on lexical knowledge and vocabulary at a whole-word level (e.g., recognising a word "by sight") or through phonological processes (e.g., sounding out the graphemes of the word before it is recognised) or through a combination of both processes.

The DRC (Coltheart et al., 2001) breaks down these processes into further detail and proposes two pathways to single word reading. The GPC route, similar to sublexical decoding in the SVR (Gough & Tunmer, 1986) and phonological recoding in Share's (1995) Self-Teaching Hypothesis, occurs through applying GPC rules to the graphemes of the target word, to retrieve the phonological representation of the word. Similarly to Share (1995) until orthographic representations are developed, early readers can use this route to phonologically decode words. In contrast, the Lexical Nonsemantic route utilises existing orthographic representations of the target word in the orthographic lexicon to retrieve the phonological representation of the word. This could be compared to Share's (1995) lexicalisation process, whereby existing orthographic representations directly activate the phonology of the target word, resulting in the reader relying less on GPC decoding.

The Triangle Model of Reading (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) outlines two routes to reading aloud: directly through orthography to phonology, or indirectly from orthography through semantics to phonology. As well as two routes to identifying the meaning of words: directly through orthography to semantics or indirectly from orthography through phonology to semantics. Both the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) and the SVR (Gough & Tunmer, 1986) include representations of oral language skills, reading comprehension and phonological decoding (Chang & Monaghan, 2019). In the Triangle Model, oral language skills are represented in the phonology to semantics pathway, whilst written word comprehension is represented in the orthography to semantics pathway and decoding is represented in the orthography to phonology pathway (Chang & Monaghan, 2019). As discussed above, the Triangle Model shares similarities with Share's (1995) Self-Teaching Hypothesis, through demonstrating how early untrained models focus on building connections in the direct orthography to phonology pathway, akin to phonological recoding. Trained models over time then rely less on this pathway and can utilise the orthographic to semantic to phonology pathway, similarly to how orthographic knowledge in Share's (1995) hypothesis is used to build whole-word orthographic representations.

Whilst the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) differs from the DRC (Coltheart et al., 2001) in design and implementation, in particular the Triangle is a connectionist model and the DRC is a nonconnectionist model; there are similarities between the models in their dual-route approach to reading aloud. Both models have a phonological recoding route associated with early reading and a faster whole-word reading route associated with later reading development, utilising whole-word orthographic representations and weighted connections. The models also differ on their implementation of semantics. The Lexical-Semantic route is not implemented in the Coltheart et al. (2001) DRC and the authors made no statements regarding how the semantic route operates (Taylor et al., 2011). Whilst the Triangle Model states that semantics operates through either direct or indirect connections (Coltheart, 2005). These indirect connections enable semantics to influence the computation of the phonology of a target word from orthography (Taylor et al., 2011).

In summary, all of these reading frameworks and models include the important skill of decoding. This is conceptualised as lexical or sublexical decoding in the SVR (Gough & Tunmer, 1986), phonological recoding in Share's Self-Teaching Hypothesis (1995), the GPC route in the DRC (Coltheart et al., 2001) and the orthography to phonology route in the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989). During this decoding process, which is often attributed to early readers, children use learnt grapheme to phoneme correspondences to sound out the target word into spoken language, which can then lead to accessing semantic information about the word (Castles et al., 2018).

In some of these models, this decoding ability allows the early reader to independently read novel text, which over time provides them with orthographic knowledge, as shown by the lexicalisation process in Share's Self-Teaching Hypothesis (1995). Over time and with experience with text, early readers can begin to rapidly map the orthography of target words to their meanings without requiring phonological decoding (Castles et al., 2018). Castles et al. (2018) refer to this process as, "orthographic learning", which includes the process of acquiring word-specific knowledge to access the semantic information of a target word and also general orthographic knowledge about the writing system. This faster route of whole-word reading through orthographic representations can be seen in the Lexical Nonsemantic route of the DRC (Coltheart et al., 2001) and the direct orthography to semantic route in the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989).

As these influential frameworks and models of reading have demonstrated, there are two key pathways to successful word reading (Lexical and Non-lexical) and both pathways develop as children progress from being novice readers through to building their reading expertise through experience with text. When considering how best to measure reading ability, it is important to consider which route is being assessed, as different word types vary in terms of the extent to which they draw on each route (i.e., requiring more or less access to lexical and/or phonological information).

1.2 Measuring Reading Ability

Following the discussion of prominent models and frameworks of word reading, it is important to consider the types of words early readers encounter in text and how these are processed in the theory of dual-routes to reading. This section will also outline how Share's Self-Teaching Hypothesis (1995) the DRC model (Coltheart et al., 2001) and the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) process these various word types.

1.2.1 Dual-Routes to Reading

The summary of the reading frameworks and models in Section 1.1.5 noted that Share's Self-Teaching Hypothesis (1995), the DRC model (Coltheart et al., 2001) and the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989), encompass two routes to reading. This theoretical proposal of "dual-route" reading models attempt to outline mechanisms for translating print to sound in English and as such require two mechanisms (Harm & Seidenberg, 2004). The first mechanism utilises knowledge of grapheme to phoneme correspondences in the English language whilst the second mechanism utilises knowledge of whole-words (Harm & Seidenberg, 2004). These two mechanisms together allow for processing of a wide range of words, from novel unfamiliar words to frequently encountered familiar words (Castles et al., 2018).

The DRC model, Triangle Model and Share's Self-Teaching Hypothesis as discussed in Sections 1.1-1.1.5 have been applied to early readers to attempt to explain how children move from "novice" readers to skilled adult reading. As such, these general dual-route mechanisms to reading have also been applied to reading development. In the first route to reading, novice readers "sound-out" a target word through grapheme-to-phoneme correspondences, hereafter referred to as GPCs. GPCs, also known as letter-sound relationships, are the correspondences between a grapheme (a letter or letter combination that represents a phoneme in a word) and their sound (Stuart & Stainthorp, 2015). The novice reader phonologically decodes the target word through "sounding-out" these GPCs and blending these sounds together to produce a pronunciation of the target word (Stuart & Stainthorp, 2015). Hereafter, decoding refers to this sublexical GPC decoding and blending unless stated otherwise. If the word is known to the reader in oral form as part of their oral vocabulary, then the meaning (semantics) of the target word can also be accessed (Castles et al., 2018). This Non-lexical, GPC decoding route is often attributed to early readers, as it is a slow, attention demanding process which we do not rely on, as skilled adult readers (Castles et al., 2018).

The second route to reading is associated with faster, skilled adult reading where the orthographic representation of a word links directly to its phonological representation to produce target word pronunciation (Stuart & Stainthorp, 2015). Alternatively, the phonological representation of a target word can be produced indirectly, through the orthographic representation's link to semantic information, then to the phonological representation (Stuart & Stainthorp, 2015). This route does not require decoding and so is sometimes referred to as "reading by sight", "whole-word reading" or "sight reading", as the word is read as a whole rather than decoding into individual GPC sounds. Whole-word reading refers to this process hereafter unless stated otherwise. Over time, oral vocabulary develops which facilitates additional semantic information being stored, along with phonological representations of these new words (Stuart & Stainthorp, 2015). Links are then formed between these recently acquired word meanings and their pronunciations (Stuart & Stainthorp, 2015).

These two routes, hereafter referred to as the Non-lexical route and the Lexical route, operate in tandem in skilled reading, as one route alone is not sufficient for skilled adult reading (Stuart & Stainthorp, 2015). Additionally, the Lexical route is associated with faster processing than the GPC route (Stuart & Stainthorp, 2015). This requirement for both routes is reflected in the parallel processing of the GPC route and the Lexical Nonsemantic route shown in the DRC (Coltheart et al., 2001) and the division of labour shown in the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989). Whilst Share's Self-Teaching Hypothesis (1995) provides a developmental account of how early readers transition from the Non-lexical route to developing orthographic knowledge and representations required for the Lexical route. The combination of these two routes allows skilled adult readers to read novel words for which they have no existing orthographic representation, through decoding, and quickly read aloud and access the meaning for familiar words through the Lexical route (Stuart & Stainthorp, 2015).

These two routes also account for how skilled readers can read a wide variety of word types in a complex language such as English, which will now be discussed.

1.2.2 Word Types

One theoretical proposal of these dual-route reading models is that each of the two routes facilitates reading particular word types, either novel or familiar: regular or irregular (Harm & Seidenberg, 2004). As such, reading ability along each route can be measured with specific word type reading performance.

The first word type to consider is what is commonly referred to as "nonwords", these are pronounceable, pseudo-words without existing semantic information, such as "*glip*" or "*flum*". As these words are novel to the reader and have no associated semantic information, these words cannot be read through the Lexical route, as there are no existing orthographic or semantic representations for these words which could lead to phonological representations (Coltheart, 2005). In order to create a pronunciation for these nonwords, the Non-lexical route must be used to decode these words, blend the GPCs together and create a pronunciation (Stuart & Stainthorp, 2015). As a result, nonword stimuli can be used to assess Non-lexical routes to reading in computational models and as discussed later, the existing GPC knowledge of early readers.

In contrast, the second word type to consider, known as "irregular" or "exception words" utilise the Lexical route to reading. The English orthography has been influenced by multiple languages such as Germanic, Norman-French and Latin-Greek and is associated with a high level of spelling-sound complexity (Share, 2008). The English language is considered a deep orthography, due to its substantial inconsistencies in the relationships between graphemes and phonemes (Castles et al., 2018). One source of this spelling-sound complexity is the six vowel sounds in English (A, E, I, O, U, Y) which can represent 20 vowel phonemes, creating a complex system separate from one-to-one grapheme to phoneme mappings (Share, 2008). However, it is important to note that there are subregularities in English, whereby surrounding context can mitigate particular spelling-sound inconsistencies, for example, the vowel sound in "*wash*" is irregular compared to "*cash*", but this irregularity is also shared in other words beginning with the "*wa*" combination, such as "*wand*" (Castles et al., 2018). These subregularities provide some additional consistency within the English language (Castles et al., 2018).

As such, the English language contains a mixture of words, either described as "regular" or "irregular". Regular words, also known as "transparent words", are words which can be phonologically decoded as they contain consistent grapheme-to-phoneme mappings, such as "*bed*" and "*cat*" (Stuart & Stainthorp, 2015). Whilst irregular words, otherwise known as "exception words", contain irregular spelling-to-sound mappings, such as "*comb*" and "*eye*" (Castles et al., 2018).

These exception words cannot be read accurately through the Non-lexical route, due to their irregular grapheme-to-phoneme mappings which contrast with the major grapheme-to-phoneme mappings of the language (Stuart & Stainthorp, 2015). If read through the Non-lexical route, exception words are "regularized", resulting in pronunciation errors (Coltheart, 2005). This occurs when the irregular grapheme-to-phoneme mappings are read through the Non-lexical route, as though the GPC mappings obey the language rules (Stuart & Stainthorp, 2015). In their example, Stuart and Stainthorp (2015) state that the word "*gauge*" would be regularized in the Non-lexical route to the word "*gorge*" and therefore the incorrect semantic information for the target word would be accessed. Alternatively, exception words such as "*medicine*" when regularized in the Non-lexical route, would not lead to accessing the semantic information of a regular word and instead would not be understood by the reader.

Therefore, for exception words to be read accurately without regularization errors, the Lexical route must be activated for a "sight reading" approach (Stuart & Stainthorp, 2015). During this route, the existing orthographic representation of the exception word activates both the phonological representation of the target word and its associated semantic information, without the requirement of phonological decoding which would lead to regularization errors (Stuart & Stainthorp, 2015). However, the exception word must be familiar to the reader in order to be read through the Lexical route (Kirby et al., 2008).

Moving from the dualism of nonwords and exception words, there are also predictions made by dual-route reading theories which separate familiar words and unfamiliar words. As noted by Share (2008) this familiar or unfamiliar dualism is not only for categories of words we know versus words we do not know, it is more so a within-item transition, as at some point during our reading development, all words were unfamiliar to us and functionally similar to nonwords with no phonological, orthographic or semantic representations. The Lexical route is associated with the ability to process familiar words known to the reader, regardless of whether these are regular or irregular in grapheme-to-phoneme mappings, as the route does not rely on decoding (Kirby et al., 2008). However, the Lexical route cannot read unfamiliar words, as there are no established orthographic representations for the target word (Kirby et al., 2008). Meanwhile, the Non-lexical route is associated with the ability to process unfamiliar words through decoding, providing that the word has regular grapheme-to-phoneme mappings (Kirby et al., 2008). Additionally, the Non-lexical route can read familiar words with regular mappings through decoding, however the Lexical route is considered optimal for familiar words due to its faster holistic recognition process (Kirby et al. 2008).

These word types: nonwords, exception words, familiar and unfamiliar words according to dual-route reading theory, rely on one route more than the other. In summary, the Lexical route is associated with familiar and exception word processing, while the Non-lexical route is associated with unfamiliar and familiar regular word and nonword processing (Kirby et al., 2008).

1.2.3 Predictions about Word Types according to Reading Models

To elaborate on word type reading in dual-route reading models discussed above, this section outlines word type reading predictions made by some of the reading models described above. As the SVR (Gough & Tunmer, 1986) is a framework of reading comprehension rather than a model of reading aloud, it makes no predictions regarding how different word types are read.

Whilst Share's (1995) Self-Teaching Hypothesis proposes that unfamiliar words are first read through phonological recoding in early readers, as a form of Non-lexical route. Due to their unfamiliarity, typically regular GPCs and lack of existing orthographic representations, nonwords could also be read through phonological recoding in Share's (1995) hypothesis.

As the early reader gains experience with various texts through phonological recoding and repeated exposures, the "advancing reader eventually assembles a large stock of instantly familiar words, each recognised as an integrated autonomous unit" (Share, 2008, p. 604). Proposing that familiar words are read through a whole-word Lexical route rather than a slower phonological recoding Non-lexical route.

In Share's Self-Teaching Hypothesis (1995), unfamiliar exception words would be read incorrectly through the phonological recoding Non-lexical route due to regularization errors. However, Share proposed that partial decoding can occur with exception words, where limited phonological information is available through phonological recoding (Share, 1995). This partial decoding leaves the reader with ambiguity as to the pronunciation of the word, however when combined with contextual information, facilitates selection of a target word amongst appropriate candidate pronunciations (Share, 1995). Over time, developing orthographic knowledge and representations for these exception words to be read through a whole-word Lexical route. Nevertheless, this orthographic knowledge is developed firstly through the Non-lexical phonological recoding route combined with contextual information for partial decodings (Share, 1995).

In the DRC model (Coltheart et al., 2001), computations of a target word pronunciation occurs through the Non-lexical and Lexical routes simultaneously, however the different routes in the DRC are associated with varying word type processing. Unfamiliar words in the DRC do not have existing orthographic representations in the orthographic lexicon or semantic information in the semantic system, therefore cannot be read through the Lexical Nonsemantic or Lexical-Semantic route. Instead, unfamiliar words would be read through serial, letter by letter decoding in the GPC route to produce a pronunciation (Rastle & Coltheart, 1999).

Similarly, nonwords would lack orthographic and semantic representations in the DRC and therefore would be read through the GPC route through serial letter decoding (Coltheart, 2005). However, the GPC route is not solely responsible for producing a pronunciation of a target nonword, as nonwords can produce activation in the orthographic lexicon for visually similar words (Coltheart, 2005). For example, the nonword "*flot*" could activate orthographic lexicon representations of similar words such as "*flat*" or "*foot*" (Rastle & Coltheart, 1999). This activation of visually similar words in the orthographic lexicon can then activate the phonological lexicon to help produce a pronunciation for the target nonword, however the phonological lexicon alone would not produce the correct pronunciation for a nonword and thus input from the GPC route is required (Coltheart, 2005). The DRC model (Coltheart et al., 2001) also suggests that the more letters there are in a nonword target, the slower it is read aloud, and that real words are read aloud faster than nonwords (Coltheart, 2005).

Familiar words in the DRC (Coltheart et al., 2001) are accessed through their existing orthographic representations in the orthographic lexicon, leading to activation in the phonological lexicon and a pronunciation from the phoneme system, therefore using the Lexical Nonsemantic route. Despite the fact that the Lexical-Semantic route was not implemented in the Coltheart et al. (2001) model, it could be predicted that a familiar word would activate its corresponding orthographic representation in the orthographic lexicon, causing further activation of the words semantic information in the semantic system (McKague et al., 2001). This would then lead to activation in the phonological lexicon and phoneme systems, thereby reading familiar words through the Lexical-Semantic route (McKague et al., 2001).

Corresponding with dual-route reading theory, if exception words were read through the GPC route in the DRC (Coltheart et al., 2001) alone, there would be regularization errors (Coltheart, 2005). Instead, known exception words are read mostly through the Lexical Nonsemantic route: activating orthographic representations in the orthographic lexicon, causing activation in the phonological lexicon. Exception words are read slower than regular words, due to competition at the phoneme level for the phoneme that has an irregular GPC mapping, as activation from the irregular phoneme is slowed by inhibition from the regularized phoneme (Coltheart et al., 2001).

The Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) does not explicitly distinguish between a Lexical and Non-lexical route across its iterations, as the model pronounces various word types using weighted connections between the three sets of units (orthographic, phonological and semantic) (Zevin & Seidenberg, 2006). When the model is presented with a target word for pronunciation, this pronunciation is generated through the information provided by all existing orthography to phonology to semantic mappings (McKague et al., 2001). Additionally, the Triangle Model does not classify word types as “regular” or “irregular”, but rather along a continuum of “consistency”, that is, the degree of consistency between the orthographic mapping and phonological mapping (Zevin & Seidenberg, 2006).

When presented with familiar words, either the direct orthography to phonology route, or the indirect orthography to semantics to phonology route can produce a pronunciation of a familiar word, as reading aloud is thought to involve both direct and indirect routes to phonology, regardless of the target word’s familiarity (Nation & Cocksey, 2009). As the orthography to phonology pathway is direct, it is considered to be faster and contribute to reading aloud more than the indirect orthography to semantic to phonology pathway (Nation & Cocksey, 2009).

Similarly to the models discussed above, nonwords and unfamiliar words have no associated semantic information in the Triangle Model, and therefore are read directly through the orthography to phonology pathway (Coltheart, 2005). As noted by Nation & Cocksey (2009) above, the indirect orthography to semantic to phonology pathway is also thought to be active during reading aloud of unfamiliar words. This indirect semantic pathway information is then utilised when the unfamiliar word has inconsistent mappings between orthography and phonology, as these words are read less accurately by the direct orthography to phonology route (Nation & Cocksey, 2009).

The orthography to phonology connections are mediated by hidden units, therefore this route of the Triangle Model is able to encode some irregular grapheme-to-phoneme mappings (Taylor et al., 2013). This is supported by simulations such as Plaut et al. (1996) that have shown that the orthography to phonology route could learn to pronounce both regular and exception words. Conversely, exception words are processed more economically through the indirect orthography to semantics to phonology pathway (Bishop & Snowling, 2004). Plaut et al. (1996) in Simulation 4 (S4) created an orthography to phonology route which read exception words with errors after training. When this simulation was given frequency-weighted activation of phonology which was proposed to serve as an approximation of information from the semantic system, a division of labour was found which maximised the models efficiency (Woollams et al., 2007). In this simulation, the orthography to phonology pathway was able to read frequent and consistent mappings, while exception word reading relied more on the orthography to semantic to phonology pathway (Woollams et al., 2007).

In summary, the Self-Teaching Hypothesis (Share, 1995), the DRC model (Coltheart et al., 2001) and the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) are able to process a variety of word types encountered by early readers and skilled adult readers, including familiar words, novel words, nonwords and exception words. How these word types are processed often align with a dual-route theory of reading, with decoding used for novel words and nonwords, while a whole-word sight reading approach is used for familiar words and exception words.

1.3 Important Skills for Reading

Reading is a complex skill, drawing upon a wide range of skills and knowledge. As part of determining how early readers develop from novices to skilled adult readers, the role of a range of literacy-related skills which have been proposed to contribute to reading development, must be considered. As discussed by Ehri (2005) there is a transition from early reading to skilled reading. Whereby early reading is initially slower, while conscious attention is given to the act of phonologically decoding unfamiliar words, potentially hindering the reading comprehension of the target word (Ehri, 2005). As reading develops, there is a transition to faster, efficient word reading, whereby readers recognise words “by sight” and read these as a whole-word from memory, without disrupting the comprehension of the text (Ehri, 2005). This generally corresponds with the two routes to reading outlined within the dual-route reading perspective, but what is also important is how early readers develop this faster, efficient, skilled form of reading and what literacy-related skills are required to facilitate this transition.

Ehri (1987; 2005) outlined a four-phase theory of reading development, which describes how readers move from novice reading to reading words through whole-word, sight reading in greater detail than outlined within the dual-route reading perspective. The first phase known as the "pre-alphabetic" phase, proposes that early readers rely on visual and context cues, as they have little alphabetic knowledge and do not possess GPC mappings, therefore reading words through visual cues such as the two "*o's*" in "*book*" (Ehri, 2005). Readers then transition to the "partial alphabetic" phase, whereby readers have some limited knowledge of letter sounds and names, typically for boundary letters at the start and end of a target word, but due to this lack of alphabetic mastery, these readers cannot yet decode novel words through GPC mappings (Ehri, 2005). Instead, these pre-alphabetic readers can read known words through utilising partial GPCs (Ehri, 2020). Readers then move to the "full alphabetic" phase, when they know the majority of GPCs and utilise this information from memory to phonologically decode novel words (Ehri, 2020). The final phase, known as the "consolidated alphabetic phase" occurs when readers are able to read increasingly more sight words from memory, rather than relying on phonological decoding (Ehri, 2005). As part of this final phase, the GPCs within these known words are memorised as larger units of information, such as syllables, which can then be applied to decoding new words with the same features, such as multisyllabic words, and eventually reading these multisyllabic words from memory (Ehri, 2020).

As part of Ehri's influential phase theory, there is the requirement for two literacy-related skills in particular: phoneme awareness (specifically phonemic segmentation ability) and letter knowledge (including letter shapes, sounds and names), in order to facilitate phonological decoding using GPCs, within this slower, initial route to reading (Ehri, 2020). Additionally, these two skills facilitate orthographic mapping, whereby the creation of GPC mappings to store spellings matched with pronunciations of known words in memory occurs, when words are successfully decoded (Ehri, 2022). However, the role of semantic information within reading development is also important, according to Ehri's phase theory - specifically, the role of vocabulary. According to Ehri (2022), vocabulary contributes to the reading of novel words in two ways. Firstly, when the target word has variable grapheme-to-phoneme spellings, vocabulary knowledge is utilised to select the correct pronunciation from potential known options. Secondly, when decoding a novel word that has not been encountered previously and producing a pronunciation, the reader requires knowledge from their vocabulary to recognise the meaning of the word (Ehri, 2022). If the reader has no existing vocabulary knowledge of the novel word, they can potentially use the context of the surrounding text to deduce the meaning of the word (Ehri, 2022).

Therefore, the transition from a novice to a skilled reader appears to include wider literacy-related skills such as phoneme awareness, letter-knowledge and vocabulary. As these literacy-related skills contribute to this skilled reading development. Evidence for the contribution of these skills to reading development and further theoretical accounts of these skills will now be discussed.

1.3.1 Phoneme Awareness

Phoneme awareness has been linked to success in early reading as one of the possible building blocks of the decoding process used by early readers. Phoneme awareness, sometimes encompassed under phonological awareness, is the explicit awareness of speech units of a language - in this case, the phonemes - which are then deliberately processed and acted upon (Castles & Coltheart, 2004). This ability in early readers to attend to, isolate and then manipulate the phonemes in spoken words has been suggested as a causal influence on word-reading skills (Muter et al., 2004). Byrne (1998), (as cited in Hulme et al., 2012) proposed that reading development depends upon the reader mastering the alphabetic principle; whereby it is understood that written letters (graphemes) represent speech sounds (phonemes). As part of this, Bryne (1998) also proposed that phoneme awareness contributes to the early reader understanding the alphabetic principle (as cited in Hulme et al., 2012).

The phonological representations hypothesis (Snowling & Hulme, 1994; Swan & Goswami, 1997) proposed that phonological awareness skills of early readers, depends upon the early reader possessing accurate underlying phonological representations of words (i.e., their phoneme awareness). Additionally, the segmental organization of these phonological representations is important (Swan & Goswami, 1997). Combined with evidence from phonological awareness tasks which indicates that early reader's segmentation ability is partially determined by their existing underlying phonological representations (Snowling & Hulme, 1994).

As Share (1995) proposed in their Self-Teaching Hypothesis, this knowledge of how graphemes correspond to the phoneme sounds is then utilised in the phonological recoding stage, to sound out words using GPC knowledge and blend the word together for a correct pronunciation. To support Share's (1995) hypothesis, there is research evidence which suggests that children who have a wider knowledge about the "constituent sounds of words", often have successful reading performance (Castles & Coltheart, 2004, p. 79). For example, Hulme et al. (2002) found in a study of five to six-year-old children that measures of phoneme awareness were longitudinal predictors of reading skill, while measures of onset-rime skills were not independent longitudinal predictors. Elbro and Petersen (2004) found that after training 35 kindergarten children in phoneme awareness over a 17-week programme, these trained children outperformed untrained children in both real word and nonword reading longitudinally (in grades 2, 3 and 7). Hatcher et al. (2004) found in a training study of Reception aged children that the level of phoneme awareness developed at the end of the training programme, with a phoneme training element, accounted for a large proportion of variance in real and nonword reading ability, for children at risk of reading failure only. In a meta-analysis of studies concerning phoneme awareness and children's word reading skills by Melby-Lervåg et al. (2012), phoneme awareness was found to be the strongest correlate of individual differences in word reading ability. This effect of phoneme awareness also remained once variations in verbal short-term memory and rime awareness were controlled for (Melby-Lervåg et al., 2012).

In their review, Castles and Coltheart (2004) concluded that if phonological awareness does play a causal role in reading acquisition, then this occurs through the ability to perceive and manipulate phonemes, linking back to the hypothesis by Share (1995). Additionally, the authors also highlight a debate regarding whether phoneme awareness and knowledge of GPCs aids reading acquisition, or this reading acquisition develops GPC knowledge and phoneme awareness (Castles & Coltheart, 2004). Additionally, Melby-Lervåg et al. (2012) highlight the issue of causality between phoneme awareness and reading development in their meta-analytic review. However, they determined that current meta-analyses combined with existing training and longitudinal studies provide support for phoneme awareness likely acting as a causal influence in early reading development (Melby-Lervåg et al., 2012). In a recent study, Cunningham et al. (2020) found a longitudinal, bi-directional relationship between reading and nonword repetition in a sample of 780 early readers in primary school. Interestingly this influence occurred both indirectly via phoneme awareness and directly. They proposed that this relationship may occur due to reading experience creating associations between orthography and phonology through orthographic restructuring; ultimately leading to high quality lexical representations of words (Cunningham et al., 2020).

Additionally, there is some debate about whether improvement in phonological awareness is separate from an improvement in letter knowledge. Knowledge of individual letter names may teach children that words are made up of small letter sized units, which can then be used to decode words (Blaiklock, 2004). Blaiklock (2004) found in their review that predictive correlations between phonological awareness scores and reading ability were often significant, however these became nonsignificant when letter knowledge was controlled for. Conversely, Ball and Blachman (1991) found that in a training study of 90 kindergarten children, phoneme awareness instruction combined with phoneme-letter matching instruction, significantly improved the early reading ability of these children. Meanwhile instruction in letter names and letter sounds alone, did not significantly improve early reading of children in another training group (Ball & Blachman, 1991). Lundberg et al. (1988) were also able to demonstrate that phonemic awareness could be successfully trained amongst pre-school children outside of an alphabetic writing system. Hulme et al. (2005) found a similar effect amongst Czech and English children, who were able to isolate phonemes although they had no knowledge of the corresponding letter for the phonemes. These studies suggest that phoneme awareness skills can be separated from letter knowledge, although these factors may still be closely related, as in Blaiklock (2004).

Considering the findings from the research discussed above, there is some indication that phoneme awareness may be separate to letter knowledge, which when trained and utilised, may lead to improvements in reading development. Therefore, it is important to distinguish between phoneme awareness and letter sound knowledge as contributing skills to reading development.

1.3.2 Letter-sound Knowledge

As discussed above, Byrne (1998) proposed that reading development depends upon the early reader mastering the alphabetic principle, for which phoneme awareness is required (as cited in Hulme et al., 2012). Additionally, letter knowledge is also required as part of understanding the alphabetic principle (Hulme et al., 2012). It is important here to distinguish the ways in which letter-knowledge has been examined. One approach is to examine the letter-sound knowledge of early readers, or to examine the letter-name knowledge of early readers (Hulme & Snowling, 2014). Hulme and Snowling (2013) proposed that learning both letter names and sounds serves as a measure of visual-phonological associative learning that contributes to learning to read, as reading aloud involves creating an association between a printed word and its verbal pronunciation. However, this section will focus on letter-sound knowledge alone, as this is explicitly outlined as part of the alphabetic principle as opposed to letter-naming ability and the two measures are often correlated with each other (Hulme & Snowling, 2014).

In a recent longitudinal study, Clayton et al. (2020) investigated the reading development of 191 children within their first year of UK primary school, along with a range of phonological language skills. They found that phoneme awareness, rapid automatized naming and letter-sound knowledge were all strong independent predictors of reading development (Clayton et al., 2020). Providing supporting evidence for the separate contributions made by phoneme awareness and letter-sound knowledge to reading development and understanding the alphabetic principle. A similar result was found in a study by Foy and Mann (2006), whereby reading performance in a sample of 66 pre-school children was associated with i) phoneme awareness including phoneme judgement ($R = 0.33$, $p < 0.01$) and phoneme manipulation ($R = 0.49$, $p < 0.001$) and ii) letter knowledge including letter sounds ($R = 0.44$, $p < 0.001$) and letter naming ability ($R = 0.31$, $p < 0.001$). Additionally, Foy and Mann (2006) also found that phoneme awareness predicted letter-sound knowledge, suggesting that phoneme awareness facilitated early readers learning of letter-sound relationships through the conceptual knowledge associated with phonological awareness.

Therefore, the issue of causality arises again, while both phoneme awareness and letter-sound knowledge contribute to early reading development, perhaps one skill facilitates the other. This issue of causality is also reflected in the work of Huang et al. (2014) who assessed the letter-sound knowledge of 1197 kindergarteners. A medium effect size was found for phonological awareness on letter-sound knowledge, which may reflect this facilitation effect; whereby increased levels of phonological awareness are associated with letter-sound knowledge (Huang et al., 2014).

Training studies have also sought to determine how letter-sound knowledge may contribute to early reading development. A training study by Bowyer-Crane et al. (2008) placed 152 children with poor oral language skills at school entry (either poor vocabulary or verbal reasoning skills) into one of two interventions over 20 weeks. The "OL" oral language intervention group received instruction on narrative skills, inference, vocabulary and comprehension, whilst the "P + R" phonology with reading intervention group received instruction on book reading skills, phonological awareness and letter-sound knowledge (Bowyer-Crane et al., 2008). Once the interventions were completed after 20 weeks, the "P + R" group performed better than the "OL" group on measures of nonword reading, spelling, phoneme segmentation and blending and letter-sound knowledge (Bowyer-Crane et al., 2008). Whilst the "OL" group performed better than the "P + R" group on measures of oral language including vocabulary and grammatical skills (Bowyer-Crane et al., 2008). It was concluded that the "P + R" intervention increased the participants' decoding ability whilst the "OL" programme improved grammatical and vocabulary skills which contribute to reading comprehension; these gains in various skills between the intervention groups were also maintained 5 months later (Bowyer-Crane et al., 2008). This combination of phoneme awareness and letter-sound knowledge training in the "P + R" group increasing decoding ability is potentially due to the mastery of the alphabetic principle gained by the intervention participants.

In 2012, this Bowyer-Crane et al. (2008) data was re-analysed by Hulme et al., who sought to determine if the "P + R" intervention had no further effect on reading ability once the impact of phoneme awareness and letter-sound knowledge was considered. Their mediation model demonstrated that the "P + R" intervention produced significant improvements in phoneme awareness and letter-sound knowledge along with word-level reading and spelling skills (Hulme et al., 2012). Crucially, the improvements gained in phoneme awareness and letter-sound knowledge fully mediated the longitudinal improvements found in word-level literacy skills 5 months later (Hulme et al., 2012). As participants were randomly assigned to either of the intervention groups in the original Bowyer-Crane et al. (2008) study, these results provide evidence that the longitudinal improvements in literacy skills, phoneme awareness and letter-sound knowledge, are causal effects rather than general reading development, as these results were not found in the "OL" group (Hulme & Snowling, 2013). Overall, this contribution from both phoneme awareness and letter-sound knowledge with a longitudinal outcome provides further support for both factors contributing independently to the mastery of the alphabetic principle in early readers.

The skills of phoneme awareness and letter-sound knowledge contributing to the mastery of the alphabetic principle provides early readers with the framework required for phonological recoding in Share's (1995) Self-Teaching Hypothesis. With evidence such as Carson et al. (2019) demonstrating that a 10-week intervention with 50 four-year-old children, delivering phoneme awareness and letter-sound knowledge, led to better nonword decoding when compared with a control group. This effect was found in typically developing early readers, however this effect of intervention on nonword decoding was not found in a group of 13 children with Spoken Language Difficulties (SLD) (Carson et al., 2019). However, the SLD group did perform better in phoneme awareness and letter-sound knowledge tasks post-intervention, akin to the typically developing group (Carson et al., 2019). The authors concluded that both phoneme awareness and letter-sound knowledge are important for early reading outcomes (Carson et al., 2019).

As the evidence above has demonstrated, there is support for the role of letter-sound knowledge in contributing to early reading development. Namely, through the mastery of the alphabetic principle and facilitating decoding in early readers. Even though there is some evidence to suggest an overlap with phoneme awareness through a potential facilitation effect, it is important to consider the contribution of letter-sound knowledge to early reading development.

1.3.3 Vocabulary

The influence of oral vocabulary has been the subject of research which seeks to identify what role oral vocabulary contributes to reading development. As some research highlights, the nature of the relations between oral language, including oral vocabulary and reading have been debated, with little consensus as to what role oral vocabulary plays (Ouellette, 2006). This is complicated by arguments which state that phoneme awareness and decoding skills create orthographic representations through phonological decoding, therefore building a vocabulary of these lexicalised words over time, whereby your vocabulary is mediated through your phonological processing skills (Ouelette, 2006). In contrast, vocabulary may support word recognition skills and reading development, through the mappings between orthographic, phonological and semantic representations (Muter et al., 2004). As suggested by Stuart and Stainthorp (2015), as early readers encounter a variety of texts and their oral vocabulary develops, more word pronunciations and word meanings are stored in their respective phonological and semantic lexicons. Links are then formed between these newly acquired word meanings and pronunciations, which will be required for Lexical-semantic routes to reading (Stuart & Stainthorp, 2015). Additionally, they propose that it should be easier for an early reader to learn unfamiliar words if the word is already stored in the child's oral vocabulary, as the pronunciation of the word is already linked to its corresponding semantic information (Stuart & Stainthorp, 2015). The early reader then must associate the novel printed word to its existing pronunciation and semantic information (Stuart & Stainthorp, 2015). In an investigation of orthographic learning, Ouelette and Fraser (2009) examined literacy-related skills which contributed to orthographic learning when fourth grade readers were exposed to nonwords with and without semantic information. Results demonstrated that vocabulary contributed to both spelling performance and performance within a recognition task, combined with the wider finding that nonwords presented with semantic information were better identified than nonwords presented with no semantic information (Ouelette & Fraser, 2009). Suggesting that vocabulary contributed to the orthographic learning process separately from decoding and irregular word reading skills which were also measured (Ouelette & Fraser, 2009). Demonstrating that it is a combination of phonological and semantic skills (including vocabulary) which contribute to successful orthographic learning and subsequent visual word recognition (Ouelette & Fraser, 2009).

In the Lexical Restructuring Model (Walley et al., 2003), it is proposed that phonological representations become fine-grained as reading develops, namely through vocabulary growth. As vocabulary develops, phonological representations are altered on an individual item basis, as the items become more familiar and are separated from phonologically similar items (Walley et al., 2003). With this increased sublexical detail of individual items, there is potential growth in phoneme awareness which could facilitate phonological decoding (Ouelette, 2006). In a study by Ouelette (2006) of 60, grade 4 readers, it was discovered that this receptive vocabulary breadth (number of phonological representations) predicted word decoding performance after age was controlled for. Providing support for the proposed associations between growing vocabulary and decoding ability. Partial support for the Lexical Restructuring Model was also found from the longitudinal study of 2790 Dutch children by Verhoeven et al. (2011), who found that early reader's beginning vocabulary predicted word decoding and reading comprehension. It is important to note that the association between early vocabulary and word decoding was weak, with stronger associations between vocabulary and reading comprehension (Verhoeven et al., 2011). Later in their reading development, it was found that word decoding predicted vocabulary development instead, indicating a continuous, reciprocal relationship between reading performance and vocabulary (Verhoeven et al., 2011).

Share's Self-Teaching (1995) hypothesis outlines the role of vocabulary as providing some top-down information when encountering novel words. This is especially useful when the early reader encounters exception words, which cannot be read through the traditional phonological recoding route, due to their exceptions to the GPC and the regularization errors that would follow. Ricketts et al. (2007) also propose that a child with a wider vocabulary can draw on this top-down support for word reading. In this case, if a child has a wider vocabulary of exception words such as "*would*" and "*should*", then they may draw on this exception word knowledge to sound out the word "*could*" correctly, as when using the decoding route to reading aloud, they can compare the incorrect pronunciation they create to their vocabulary and determine that the mispronunciation is not a real word (Ricketts et al., 2007).

Furthermore, there has been empirical evidence which has separated oral vocabulary from phoneme awareness and phonological processing. Ricketts et al. (2007) found that amongst 81 eight to ten-year-old children, vocabulary did not account for any unique variance in decoding ability or text reading accuracy. In contrast, vocabulary did account for variance in reading comprehension and exception word reading (Ricketts et al., 2007). Providing support for the role of vocabulary in Share's (1995) Self-Teaching Hypothesis as top-down support, especially for exception word reading. While Muter et al. (2004) found that in a 2-year longitudinal study of 90 early readers starting at four-years-old, vocabulary did not predict word recognition, whilst letter knowledge and phoneme sensitivity did. In contrast, vocabulary did predict reading comprehension whilst letter knowledge and phoneme recognition did not.

As briefly discussed in Section 1.1.1, vocabulary is not only important for reading aloud, but also reading comprehension. A key demonstration of this is how knowledge of oral language including vocabulary, along with decoding skills interact in the Simple View of Reading (Gough & Tunmer, 1986). Lervåg and Aukrust (2010) conducted a longitudinal study of second-grade children consisting of 198 first and 90 second language learners in Norwegian. They discovered that whilst vocabulary and decoding skills predicted reading comprehension performance, only vocabulary predicted further growth in reading comprehension. Suggesting that once these children had become proficient in phonological decoding, that vocabulary became increasingly important for reading comprehension (Lervåg & Aukrust, 2010). Additionally, the second language learners demonstrated a delay in their reading comprehension skills compared to the first language learners, which was mediated by limitations in their vocabulary (Lervåg & Aukrust, 2010).

Ouelette and Beers (2010) conducted a study with 67 children from grade 1 and 56 children from grade 6 in Canada, measuring reading related skills of phonological awareness, decoding, irregular word reading, listening comprehension, reading comprehension, vocabulary breadth and vocabulary depth. At grade 1, it was demonstrated that oral vocabulary did not predict reading comprehension beyond phonological awareness, decoding, listening comprehension and irregular word recognition (Ouelette & Beers, 2010). In contrast, at grade 6, oral vocabulary predicted reading comprehension even when the aforementioned variables were accounted for (Ouelette & Beers, 2010). Suggesting that in early reading development, decoding skills are crucial for early readers, however over time, the role of decoding is diminished and instead oral vocabulary plays a larger role in reading comprehension (Ouelette & Beers, 2010). This idea of a diminishing role of decoding was also supported by their grade 6 finding, where decoding was no longer a significant predictor of reading comprehension as in grade 1 (Ouelette & Beers, 2010).

The evidence discussed above indicates that oral vocabulary is not limited by a child's phonological processing skills and may contribute to early reading separately from phonological decoding. Although there may be a reciprocal relationship between vocabulary and phonological growth (Ouelette, 2006; Walley et al., 2003) combined with vocabulary and phonological skills contributing to successful orthographic learning (Ouelette & Fraser, 2009) and growing links between orthographic, phonological and semantic representations (Stuart & Stainthorp, 2015). Thus, it is important to consider the role that vocabulary plays in reading development, which could be addressed in three ways. Firstly, vocabulary acts as a top-down source of information for encountering novel words, perhaps through drawing on the learnt phonological information from words in the vocabulary for successful novel word decoding (Ehri, 2022; Ouelette, 2006; Ricketts et al, 2007). Secondly, semantic and phonological information, including vocabulary and phonological decoding, may contribute to the process of orthographic learning (Ouelette & Fraser, 2009) which facilitates the development of orthographic representations and growth of the Lexical route to reading from a dual-route reading perspective. Thirdly, vocabulary contributes to later reading comprehension, once decoding has been mastered and faster Lexical routes to reading are utilised, which rely on both the fully specified phonological and semantic information stored in their respective Lexicons, which are provided in part by oral vocabulary (Ouelette & Beers, 2010).

1.3.4 Home Literacy Environment

The importance of the home literacy environment that a child is exposed to, has been linked to reading performance at an early age. DeBaryshe et al. (2000) stated that a home literacy environment provides opportunities for a child, which may impact their reading development. These include opportunities to become familiar with literacy materials, independently explore literacy behaviours, engage in joint reading activities with others and benefit from teaching strategies that family members use in these joint tasks (DeBaryshe et al., 2000). The idea of a rich home literacy environment is often combined with the role that the parent plays in their child's reading development. For example, parents are thought to promote their child's learning and development through structuring this home literacy environment through preparing routine literacy activities, providing age-appropriate learning materials and supportive engagement (Rodriquez et al., 2009). Studies such as Foy and Mann (2003) found that exposure to reading-related media combined with parental involvement in reading children's literature was directly and indirectly linked to the child's letter and vocabulary knowledge.

The findings from home literacy environment research can be divided into two categories, which often correspond with the DeBaryshe (1995) model regarding maternal belief systems. Firstly, that the parental reading beliefs are associated with the parent-child literacy and language activities (Weigel et al., 2006). Secondly, that these parent-child literacy and language activities are associated with the child's literacy and language skills (Weigel et al., 2006). Therefore, a parent with reading beliefs which encourage shared and independent reading can increase the child's interest in reading and this rich environment combined with an interest in reading can provide children with an advantage in early reading development.

Weigel et al. (2006) found in a study of 85 parents and their pre-school aged children (approximately three-years-old), that parent engagement in literacy and language activities was positively associated with the child's interest in reading and greater print knowledge performance, which remained significant at a one-year follow up. Despite this, more frequent parent-child literacy and language activities were not associated with the child's receptive or expressive language or writing skills (Weigel et al., 2006). Suggesting instead that these verbal language skills were linked to the parents' level of education and income, whilst reading skills could be impacted through the quantity and quality of parent-child literacy related activities (Weigel et al., 2006).

Burgess et al. (2002) examined the home literacy environment of 97 four to five-year-old children and how these environments related to their literacy related skills such as word decoding, phonological sensitivity, letter knowledge and oral language over one year. Their findings indicated that the home literacy environment was statistically significantly related to the assessments of oral language, phonological sensitivity and word decoding ability (Burgess et al., 2002). This suggests that the home literacy environment that a child is exposed to may have early and potentially lasting influence on their reading development (Burgess et al., 2002).

While some research suggests that home literacy environments do not tell us the full story, and we are potentially ignoring heritable parental traits (van Bergen et al., 2016), the research on home literacy environment does provide compelling evidence for how this environment can impact early reading development.

1.4 Reading Difficulties

From all of the elements discussed so far, it can be concluded that reading by an early reader and transitioning into a skilled fluent reader is no easy task, as the reader must develop dual-routes to reading, successfully read various word types and develop literacy-related skills. This section discusses reading difficulties which can occur from a dual-route reading perspective both in early readers and how reading difficulties are reflected in dyslexia. Whilst this thesis considers reading difficulties shown by children in mainstream classrooms, not specifically children with dyslexia; much of the literature and theoretical basis regarding reading difficulties along the dual-route pathways outlined above, is found within dyslexia literature, as highlighted by Wang et al. (2014). Therefore, this introduction to the area will discuss reading difficulties both in terms of research conducted with mainstream poor readers and research using dyslexia classifications.

At this point it is important to specify how mainstream poor readers differ from those with a dyslexia classification. Dyslexia is assessed by an Educational Psychologist and is often characterised as difficulties in phonological awareness, verbal memory and verbal processing speed, with a wider profile than just reading difficulties (Rose, 2009). In contrast, a poor reader is any child who performs poorly in reading, relative to their peers in the mainstream classroom (Rose, 2009). As poor readers are defined compared to the performance of their peers in that year, there is no objective criteria for classifying a poor reader. Due to this lack of criteria, practices for identifying and supporting poor readers vary across classrooms; with some schools referring to standardised tests while others use teacher observations to identify a poor reader. Alternatively, poor readers can be defined more systematically in research studies, compared to a large sample of expected performance for their age group, for example, the use of a standard score below 85 on the Diagnostic Test of Word Reading Processes (DTWRP; Forum for Research into Language and Literacy, 2012).

McArthur et al. (2013) described dyslexia in terms of two main subtypes, the first being the classic “phonological dyslexia” subtype, where the individual has poor phonological recoding abilities, so they struggle to read novel regular words and nonwords. However, these individuals possess intact lexical reading including exception words, as these have been learnt as whole-word orthographic representations (McArthur et al., 2013). The other subtype they describe is “surface dyslexia”, whereby lexical reading is poor, so these readers rely on phonological recoding as a sublexical reading strategy, resulting in difficulties with exception word reading, but adequate nonword reading (McArthur et al., 2013).

From a dual-route reading perspective, phonological dyslexia would occur when the Non-lexical pathway is hindered, resulting in a dependence on existing orthographic knowledge and representations to read words by sight, without the ability to decode target words accurately. In contrast, surface dyslexia would occur when the Lexical pathway is hindered, so orthographic representations cannot be created or accessed, resulting in a reliance on decoding to read target words.

McArthur et al. (2013) in a study of 138 children with different dyslexia profiles sought to understand the underlying deficits across these profiles. McArthur et al. (2013) found that the most common deficit within the phonological subtype was poor GPC knowledge, while surface subtypes had impaired orthographic lexicons, with impaired links to semantic knowledge and the phonological lexicon. Providing supporting evidence for how dual-route reading models would explain difficulties within both dyslexia subtypes. Whereby poor GPC knowledge would hinder decoding in the Non-lexical route and the inability to access or create new orthographic representations would hinder the Lexical route, resulting in default decoding of target words. Griffiths and Snowling (2002) also found that in their sample of 59 dyslexic children, the only predictor of exception word reading was reading experience. Corresponding with predictions made by dual-route models of reading and orthographic learning, such as Share's Self-Teaching Hypothesis (1995), whereby print exposure builds orthographic representations of sight words which cannot be taught via decoding.

These subtypes and their corresponding deficits have been found across different studies (Castles & Coltheart, 1993; Manis et al., 1996; Romani et al., 2008). Despite this, the literature is inconclusive as to whether each of these subtypes are due to a developmental delay which may improve with age, or a deviant form of reading development. Manis et al. (1996) found an overlap between their younger comparison group and surface dyslexia group on nonword and exception word reading; suggesting surface dyslexia is caused by a developmental delay (Manis et al., 1996). Meanwhile the phonological dyslexia group fell below scores for the younger group with nonword reading; suggesting that they may have some sort of deviant development to fall behind these younger children (Manis et al., 1996). Similar results of phonological dyslexia as deviant development and surface dyslexia as a developmental delay were also found by Stanovich et al. (1997) in a study of 68 children and Wolff (2009) in a study of 40 university students. Melby-Lervåg, et al. (2012) also found in their meta-analyses that children with a dyslexia classification performed significantly worse than both age-matched controls and younger, reading-level matched control groups with no reading difficulties, on measures of phonemic awareness.

The suggestion that phonological and surface dyslexia subtypes may have different developmental trajectories is consistent with dual-route theories, whereby independent reading exposure and practice is required to develop orthographic routes of reading. Therefore, if children do not develop these Lexical and Lexical-semantic routes, they may be developmentally delayed compared to their peers and they display reading difficulties consistent with surface dyslexia. Meanwhile, decoding is considered the earlier pathway to reading, leading to established orthographic representations through the ability to "self-teach" reading novel texts through decoding (Share, 1995). If this Non-lexical pathway is not established, children display difficulties consistent with phonological dyslexia.

Alternatively, some researchers have suggested that this matter is not so simple. Ziegler et al. (2008) in a study of 24 dyslexic children found that some dyslexic participants had more than one deficit, so subtyping is not as clear cut as a difficulty along the Non-lexical or Lexical route to reading. Peterson et al. (2013) uncovered multiple deficits in their study, where two types of deficit were described. "Pure" deficits, where the individual has a reading difficulty in either the Non-lexical or Lexical pathway alone and "Relative" deficits, where the individual has deficits in both pathways. Therefore, we must be cautious when grouping individuals to ensure that one difficulty is not emphasised whilst a secondary difficulty is ignored, as this may skew our view of how best to improve that individual's reading ability.

Furthermore, Peterson et al. (2013) also suggested that both phonological and surface dyslexia was associated with patterns of deviant development rather than a developmental delay. When compared to a control group matched at reading level, participants with phonological and surface dyslexia showed a-typical patterns of performance, including selective impairment in their defining component process (phonological coding composite and orthographic coding composite respectively) (Peterson et al., 2013). Additionally, the surface dyslexia group performed significantly worse than the reading level matched control group on the sub-type defining measures of orthographic coding and exception word reading (Peterson et al., 2013). This suggests that as neither dyslexic subtype group performed identically to the younger control group of typically developing readers, that these results reflected a deviancy in typical reading development rather than a developmental delay (Peterson et al., 2013). Therefore, to address the conflicting findings regarding the trajectory of dyslexia subtypes, further research is required to address whether dyslexia subtypes are caused by a deviancy in reading development, or developmental delay.

Similar research regarding reading difficulties has also been conducted with participants without a dyslexia classification. Wang et al. (2014) measured orthographic reading in groups of poor readers, classified as surface or phonological subtypes. The phonological group performed at a lower level on a GPC measure than the surface group, while the surface group displayed impaired orthographic learning. Both findings are consistent with the phonological and surface profiles of reading difficulties according to dual-route models of reading. Wang et al. (2014) conducted a follow-up study with 91 poor readers and found that phonological decoding predicted orthographic learning, while orthographic knowledge predicted timed accuracy in reading. These are both consistent with dual-route models of reading and Share’s (1995) Self-Teaching hypothesis, whereby phonological decoding builds orthographic knowledge and direct access to orthographic representations, which results in quicker reading ability than decoding alone.

In terms of developmental trajectories of poor readers, similar findings to dyslexia research have also been uncovered in studies of children without a dyslexia classification. Talcott et al. (2013) found that children with impaired exception word reading abilities displayed a trajectory of developmental delay, compared to children with impaired nonword reading, who displayed a trajectory of deviant development.

In summary, it seems that many problems can arise along the dual pathways to reading. Difficulties can occur across individual pathways (Lexical versus Non-lexical) and can manifest itself as one overarching deficit or as multiple deficits, making it difficult to tackle the underlying deficit to ensure normal reading develops.

1.4.1 Exception Word Reading Difficulties

In this introduction thus far, models and frameworks of skilled adult reading, which are often applied to reading development have been discussed. Along with discussions of these potential word reading difficulties and how these models, frameworks and the dual-route reading perspective, accounts for these reading difficulties. Nonetheless, the literature is still inconclusive as to why children may struggle to read exception words specifically. As discussed in Section 1.2.2, according to the dual-route reading approach, exception word reading difficulties would reflect difficulties present along the Lexical pathways, as known exception words are read mostly through the Lexical Nonsemantic route. This occurs through activating existing orthographic representations in the orthographic lexicon, causing activation in the phonological lexicon. As the Non-lexical route cannot read these exception words correctly due to regularization errors. While there is little literature specifically targeted at investigating exception word reading difficulties, there is evidence supporting multiple hypotheses which will be discussed.

Johnston et al. (2014) investigated if exception word reading was better predicted by direct reading measures of orthographic processing and nonword reading, than general measures of vocabulary and reading frequency. Johnston et al. (2014) tested 180 children aged six to eight-years-old on a battery of tests and found that these direct measures were better predictors of exception word reading ability than the general reading measures. Additionally, nonword reading ability was found to be a predictor of irregular word reading than orthographic processing. Johnston et al. (2014) suggested that direct training on word recognition would be more beneficial with exception word reading difficulties than interventions focusing on wider factors such as reading frequency.

From a dual-route reading perspective, we would expect phonological decoding skills to impact exception word reading, as successful phonological decoding can be used to create orthographic representations with exposure to text, consistent with reading development outlined in Share's Self-Teaching hypothesis (1995). However, we would also expect reading frequency and vocabulary to impact irregular word reading, as a varied reading environment and vocabulary would hypothetically aid the creation of these orthographic representations.

Nonetheless, as this research was conducted with children who received extensive sight word training as part of their education, it cannot be generalised to all mainstream classrooms. Especially since there is limited guidance for exception word teaching in England and it is acknowledged that exception words taught vary amongst the approved phonics programmes used, we cannot guarantee that all children are receiving the same sight word training, which may have an impact on results. Ideally this research would be applied to mainstream classrooms under current synthetic phonics teaching to investigate potential differences in these groups due to teaching practices.

Steacy et al. (2017) investigated exception word reading in a sample of 170 children with reading difficulties. They found that nonword decoding, vocabulary and orthographic knowledge were significantly related to exception word reading (Steacy et al., 2017). This included children with higher receptive vocabularies and greater orthographic knowledge being more likely to read exception words correctly (Steacy et al., 2017). In contrast to Johnston et al. (2014), Steacy et al. (2017) did find evidence that wider reading factors, in the form of vocabulary, impacted exception word reading. These results were also found in studies such as Ricketts et al. (2007). Ricketts et al. (2016) also found that semantic knowledge, tested through expressive vocabulary, shared a close relationship with exception word reading.

Wider reading factors have also been proven to contribute to exception word reading, such as the Griffiths and Snowling (2002) finding that reading experience predicted exception word reading. Nation (2017) also suggested that experiencing novel words in diverse environments is critical for reading development and may inform children of exceptions to GPCs within exception words.

From the literature discussed, we do not have a clear picture of what could be causing difficulties in exception word reading as there is evidence for multiple factors. Therefore, there are multiple possibilities for the cause of this reading difficulty. The first is that these children have a lack of exposure to text, which hinders their development of orthographic representations and knowledge of exception words, as found by Griffiths and Snowling (2002). The second is that whilst children may have experience with text, they are not forming orthographic representations which they can activate later for quick "by sight" exception word reading. Which was partially confirmed in Johnston et al. (2014) who found that orthographic processing skill contributed to accurate exception word reading. The final possibility is that a poor vocabulary may impact exception word reading, as in Steacy et al. (2017) and Ricketts et al. (2007). Additionally, these children with exception word difficulties and poor vocabularies cannot rely on their top-down information from their vocabulary knowledge to sound these words out correctly, if they are relying on phonological recoding within the Non-lexical route.

Due to the lack of consensus regarding the origin of exception word reading difficulties, further research is required in this area. Although literature regarding reading interventions may provide some insight into how this profile of reading difficulty can be addressed.

1.4.2 Reading Interventions

From the literature discussing reading difficulties within a dual-route reading perspective, there is an understanding of how these difficulties might be addressed through reading interventions, often through a tailored approach to address the pathway responsible for the word type difficulty. For example, children struggling to read nonwords may benefit from interventions targeting GPC knowledge and phoneme awareness to facilitate phonological recoding along the Non-lexical route. Meanwhile children struggling to read exception words may benefit from interventions targeting their vocabulary, exposure to accessible texts, underlying phonological recoding ability and exception word knowledge, in order to facilitate independent reading and development of orthographic and semantic representations along the Lexical routes.

Gustafson et al. (2007) conducted an intervention with 80 children with reading disabilities in Sweden. They proposed, based on a previous study (Gustafson et al., 2000), that interventions should target the child's existing strategy for reading, as phonological training may not be beneficial for all children. In the 2007 study, Gustafson et al. tested children on a battery of tests. Children were classified as having either a phonological or orthographic profile of difficulty and were randomly allocated to an intervention which focused on improving either orthographic or phonological skills.

Post-intervention results revealed a three-way interaction, where phonological training was more effective for children with phonological difficulties, while orthographic training was more effective for children with orthographic difficulties (Gustafson et al., 2007). The authors recommended that interventions be selected on the basis of individual differences in reading, as what works for one child may not work for another, based on their underlying deficit (Gustafson et al., 2007). Gustafson et al. (2007) also suggested that for the phonological group, orthographic instruction would not have helped them develop basic letter sound knowledge, while phonological training for the orthographic group would have been too simple, as this group has already solidified these skills.

The same recommendations have been made by O'Brien et al. (2012) who suggested that different approaches to intervention should be used, depending on the subtype of dyslexia the programme was addressing, for example, with or without phonological deficits. In contrast, studies such as McArthur et al. (2015) recommend both phonics and sight word training for poor readers in general, as their training study had significant effects on nonword and irregular word reading in a group of 41 poor readers. Providing evidence for the efficacy of interventions which recruit phonics training and sight word training as intervention strategies.

Fiorello et al. (2006) proposed a Cognitive Hypothesis Testing (CHT) model which would aim to develop targeted interventions for children with reading difficulties, who do not respond to standard interventions in schools. Using this model, the child's reading difficulty and previous intervention data are used to develop a theory about the underlying problem, which is then investigated with standardised cognitive and intellectual testing (Fiorello et al., 2006). Through uncovering the child's underlying deficit, a tailored intervention can then be created which will target their specific area of difficulty. Fiorello et al. (2006) used case studies to highlight how the model works and conducted discriminant analyses of 128 children with reading disabilities, which found disparate cognitive profiles across readers which potentially could not be addressed with a standard intervention. Providing further support for interventions tailored to the type of reading difficulty shown individually.

Interventions are not currently covered as part of the National Curriculum in England, as part of this, schools are responsible for which programmes are chosen, which skills these address and which children are placed onto an intervention. Therefore, it cannot be concluded that all poor readers are receiving adequate support which is tailored to their area of difficulty. Previously, the UK used the National Strategies Primary Literacy Framework up until 2011 to provide guidance and support for teachers of primary level literacy. Rose (2009) elaborated on this framework, which used a three wave system. Wave one consisted of quality first teaching, with a focus on systematic phonics to develop children's phonological awareness and wider reading skills (Rose, 2009). During wave one, teachers could identify if a child's performance was falling behind their peers (Snowling & Hulme, 2012). Wave two recommended small group or one-to-one interventions for children identified in wave one as requiring support (Rose, 2009). Wave three then recommended intensive support for children who did not respond to intervention in wave two and require a personalised approach, using an intervention tailored to their specific area of difficulty (Rose, 2009).

It could be argued that since the use of phonics has only become more widespread (see Sections 1.5.1 and 1.5.2. below for a further discussion), this three-wave system could still be used to implement tailored interventions in schools. In theory, this could help poor readers struggling with the current phonics curriculum and improve potential areas of phonological deficit. Alternatively, from a dual-route reading model and intervention perspective, it is suggested that phonics does not address all poor readers' area of difficulty. Therefore, schools may need to implement other interventions to address reading difficulties outside of phonological awareness and GPC knowledge.

1.5 How Reading is Taught in the UK

The literature discussed thus far has provided a theoretical basis for why phonics teaching has been introduced into the English National Curriculum. Specifically, phonics teaching is aligned with theories of early reading development, through facilitating word reading through phonological decoding within the Non-lexical route, from a dual-route reading perspective. Early readers who rely on this phonological decoding for unfamiliar words are explicitly taught target GPCs, to enable translation of letters into speech sounds, which can then be blended together for whole-word pronunciation (Shapiro & Solity, 2016). Combined with the phoneme awareness and letter-sound knowledge provided through systematic synthetic phonics teaching, early readers are provided with some of the key skills required for successful reading along the Non-lexical route. With time, these instances of successful novel word decoding and blending provide the early reader with the opportunity to create orthographic representations of the target word (Shapiro & Solity, 2016). Thus, facilitating independent reading and creation of further orthographic representations and knowledge, as early readers gain experience with text as outlined by Share's Self-Teaching Hypothesis (1995). The sections below outline the influence of the 2012 Phonics Screening Check on the English National Curriculum and how this mandatory systematic, synthetic phonics is delivered in primary schools to facilitate the reading development of early readers.

1.5.1 Introduction of the Phonics Screening Check

In 2012, the UK government introduced a National Phonics Screening Check for all children at the end of Year 1. Government guidance states that the Phonics Screening Check (PSC) is designed to assess whether each child has developed phonological decoding skills to an age-appropriate level by the end of Year 1, building upon phonics knowledge introduced in Early Years Foundation Stage (Department for Education, 2013; Standards and Testing Agency, 2011). Explicitly, the Phonics Screening Check was designed to assess early reader's knowledge and application of GPCs taught through synthetic phonics during Reception and Year 1, not as an overall measure of reading ability (Stainthorp, 2020). During these years of education, children are taught 85 GPCs. This knowledge and their ability to phonologically decode words is assessed through each child being asked to read 20 nonwords and 20 real words aloud, with a mixture of items varying in difficulty included across both word types (Darnell et al., 2017). Data is collected for each child in that Year 1 class and is then submitted to the local authority. The threshold score for that year is released on the government website, where administrators compare each child's score with this threshold, which determines if the child has passed the Phonics Screening Check (Standards and Testing Agency, 2019). With data suggesting that the initial threshold to pass the Phonics Screening Check is 32 correctly read items out of 40, which has remained the threshold across all iterations (Stainthorp, 2020).

Conversely, some research has argued that the Phonics Screening Check does not assess all of which it claims to. Darnell et al. (2017) reviewed the content of the Phonics Screening Check from 2012 until 2014, exploring the use of GPCs and how GPCs chosen for the assessment reflect what is being taught through a synthetic phonics approach. This research found that 27 GPCs taught in synthetic phonics were not tested in any of the three Phonics Screening Checks, whilst four additional GPCs which were not specified in the framework were tested in these versions of the Phonics Screening Check. Darnell et al. (2017) also found that children could reach the pass rate (80% for those respective years) if they only had knowledge of high frequency GPCs, suggesting that the Phonics Screening Check is not thorough enough to determine a child's knowledge of taught GPCs and therefore successful phonics instruction. However, since this research focused on the earlier years of the screener, it cannot be determined that the Phonics Screening Check has not improved since this review.

The Phonics Screening Check also has two subsequent aims, the first is to identify which children require further phonics support outside of classroom teaching, who will then have to retake the Phonics Screening Check in the following academic year to determine if progress has been made (Standards and Testing Agency, 2011). A study by Duff et al. (2015) sought to confirm whether the Phonics Screening Check was both valid as a measure of phonic skills and was able to identify children displaying signs of being at risk for a reading difficulty. The results demonstrated that the Phonics Screening Check was a valid measure of early reader's phonic skills, as the Phonics Screening Check correlated strongly with wider measures of phonic skills such as nonword reading (Duff et al., 2015). Whilst displaying weaker correlations for skills not assessed within the Phonics Screening check, such as maths (Duff et al., 2015). Interestingly, the Phonics Screening Check also correlated strongly with wider measures of reading such as single-word reading accuracy (Duff et al., 2015). In terms of the Phonics Screening Check's ability to detect those at risk of a reading difficulty, it was discovered that the measure had a sensitivity of 88% when compared to identification rates through standardised tests (Duff et al., 2015). However, the Phonics Screening Check had a specificity of 82%, resulting in some overestimation of at-risk children, although the authors state that this may be desirable in a screening measure to ensure all potential candidates for a reading difficulty are provided support (Duff et al., 2015).

The second aim of the Phonics Screening Check was to encourage schools to implement phonics programmes for all children early in the Year 1 curriculum, resulting in intensive synthetic phonics teaching (Standards and Testing Agency, 2011). It is important to state that phonics was used in primary school teaching before the Phonics Screening Check was introduced and the government have been providing guidance on systematic synthetic phonics teaching materials since 2010 (Department for Education, 2010). During 2010, the government introduced 11 core criteria for synthetic phonics programmes, including points such as, "demonstrates that phonemes should be blended, in order, from left to right, 'all through the word' for reading" (Department for Education, 2010, p. 2). As following the Rose Review in 2006, it was advised that all English primary schools should include a programme of systematic, synthetic phonics teaching during Key Stage 1 (Stainthorp, 2020). These criteria allowed primary schools in England to choose their own phonics programme, whilst ensuring the proposed programme met the government's criteria for systematic phonics teaching (Stainthorp, 2020).

Although, results from the first year of the Phonics Screening Check revealed that only 58% of children met the required threshold, suggesting that in 2012, phonics teaching was less successful in teaching phonics strategies and GPCs to this cohort of early readers (Stainthorp, 2020). As a result, since the Phonics Screening Check was introduced, there was a further educational shift towards intensive, synthetic phonics teaching early in the curriculum. For example, all phonics programmes used in primary schools across England now have to satisfy 16 core criteria (Department for Education, 2022c). 19 government approved phonics teaching programmes have been outlined by the Department for Education, which have been validated as teaching systematic synthetic phonics effectively (Department for Education, 2022a). Data has shown that since the introduction of the Phonics Screening Check in 2012 and the subsequent educational shift towards synthetic phonics teaching, the percentage of children achieving the threshold to pass the screener has risen from the original 58% and remained at 80%-90% (Stainthorp, 2020). This wider educational shift towards teaching phonics as part of the UK English curriculum is discussed further below.

1.5.2 The increased focus on Phonics Teaching

Following the introduction of the Phonics Screening Check in 2012 and the initial data of only a 58% pass rate, in 2014 the UK government introduced a new national curriculum programme of study for reading, which encompassed requirements for children to be taught phonics explicitly (Stainthorp, 2020). This programme of study for reading consists of two dimensions at Key Stage 1: skilled word reading and comprehension. Within the dimension of skilled word reading, this includes the ability to phonologically decode unfamiliar words and conduct whole-word reading of familiar words (Department for Education, 2014). The combination of these two processes as part of skilled word reading reflects the theoretical approach of dual-route reading models. The programme of study also states that children should understand how letters represent sounds (GPCs) and as a result, phonics should be implemented in the teaching of reading when young readers begin school (Department for Education, 2014).

Whilst within the second dimension of reading comprehension, this includes utilising existing linguistic knowledge, vocabulary and wider knowledge of the world (Department for Education, 2014). As part of this, reading across a wide range of texts is encouraged to develop this world knowledge and vocabulary (Department for Education, 2014). These two dimensions of reading stated in the programme of study for reading therefore reflect the processes outlined in the Simple View of Reading (Gough & Tunmer, 1986) as a framework for developing successful reading skills. In 2022, this framework was also cited in the Department for Education reading framework as guidance for the wider National Curriculum, as early readers can be taught decoding through systematic synthetic phonics (Department for Education, 2022b). Whilst linguistic comprehension can be taught through wider reading related activities, such as discussing stories, talking with peers and learning poems (Department for Education, 2022b).

The 2013 National Curriculum and the 2014 programme of study for reading outline key elements to be taught within Year 1 to early readers, which provides an overview of how systematic synthetic phonics programmes are delivered in primary schools across England. In Year 1, statutory requirements include teaching pupils to "apply phonics knowledge and skills as the route to decode words", "correct sound to graphemes for all 40+ phonemes" and "read accurately by blending sounds in unfamiliar words containing grapheme to phoneme conversions that have been taught" (Department of Education, 2013, p. 10). Overall, recommending that phonics is taught through a systematic, synthetic phonics programme, whereby the "synthetic" strategy places emphasis on serial-order decoding of GPCs in a target word, moving from left to right (Cunningham et al., 2021). This approach results in phonics teaching which firstly teaches early readers GPCs so that a novel word can be segmented into its individual GPCs, which are then blended together from left to right, to read the target word aloud (Department for Education, 2022c).

At Year 1, the majority of the statutory requirements for teaching reading focus on either phonic decoding skills and GPCs, or how phonics is taught, for example, the use of decodable books suitable for early readers. All of this statutory guidance within the National Curriculum, requirements for phonics programmes, approved programmes and advice on teaching methods and decodable materials, demonstrates that within Key Stage 1, there is a large synthetic phonics focus in English in the current curriculum.

The 2022 Department for Education reading framework outlines how systematic, synthetic phonics (SSP) programmes aim to deliver this phonics teaching and ensure that all children learn GPCs to facilitate segmenting words into phonemes and then blending these phonemes into spoken words (Department for Education, 2022b). The framework states that SSP programmes vary on their specific systems, such as the use of prompts, mnemonics, routines and key words (Department for Education, 2022b). While these programmes differ on their taught GPC order, the majority of programmes begin with a simple approach with one grapheme mapped to each of the 44+ phonemes (Department for Education, 2022b). This is then followed by introducing common alternative graphemes and gradually introducing new graphemes as the early readers are able to access more texts (Department for Education, 2022b).

In regard to the "systematic" element of the SSP programmes, the reading framework states that "daily phonics sessions should begin as soon as children start their Reception year", emphasising the focus on phonics during early year's education (Department for Education, 2022b, p. 47). These phonics sessions gradually progress from 10 minutes per day to 1-hour sessions by the end of Reception, to accommodate for consolidating this new GPC knowledge across various activities and opportunities to practice (Department for Education, 2022b).

From a dual-route reading perspective, this systematic, synthetic phonics instruction would strengthen the GPCs required for successful phonological decoding of novel words within the Non-lexical route. As predicted by Share's Self-Teaching Hypothesis (1995), over time this ability to phonologically recode text would facilitate access to a variety of texts and development of orthographic representations. Additional orthographic knowledge could be gained through certain GPCs taught through systematic synthetic phonics, for example, the GPC "*ou*" in "*would*" may also provide some orthographic knowledge for positional letter effects. This combination of orthographic knowledge and orthographic representations can then facilitate whole-word reading through the Lexical routes to reading in dual-route reading models. If the early reader also has an adequate vocabulary, once a word is either decoded or read as a whole-word through "reading by sight", then the semantic meaning for that word can be accessed (Stuart & Stainthorp, 2015).

Some research has challenged the ability of synthetic phonics instruction to provide readers with the best route for reading success. Torgerson et al. (2006) conducted a systematic review of randomised control trials utilising phonics instruction and reading outcomes. They found that while systematic phonics instruction positively increased reading accuracy, there was no statistically significant difference between the use of synthetic versus analytic phonics (Torgerson et al., 2006). Additionally, reading accuracy between normally developing children and children at risk of reading failure was not statistically different after receiving synthetic phonics instruction, suggesting that synthetic phonics is not better than other phonics approaches (Torgerson et al., 2006). Although the review did conclude that systematic phonics teaching was more effective for enabling progress for children at risk of reading difficulties and typically-developing children, when compared to no phonics teaching (Torgerson et al., 2006). Overall, the review concluded that systematic phonics programmes with a broad literacy curriculum had a positive effect on reading accuracy (Torgerson et al., 2006).

Wyse and Goswami (2008) reviewed the studies of Torgesen et al. (2001) and Berninger et al. (2003). They found that while these studies support the use of systematic phonics instruction as an effective teaching method, they did not find evidence supporting synthetic phonics as the most effective form of phonics delivery. In a more recent systematic tertiary review of phonics based systematic reviews and meta-analyses by Torgerson et al. (2019), an overall positive effect of phonics teaching was found. Nevertheless, there was limited evidence to conclude which form of phonics teaching was best (synthetic versus analytic) (Torgerson et al., 2019). With the authors suggesting that there are two potential sources of bias in the phonics systematic reviews and meta-analyses published, through design and publication bias which may exaggerate the benefit of phonics teaching (Torgerson et al., 2019).

An earlier study by Johnston and Watson (2004) compared the effectiveness of three phonics programmes by dividing 300 children in early primary school into three matched groups. The first group were taught through a synthetic phonics approach, the second group were taught through an analytic phonics method and the final group were taught through an analytic phonics method combined with systematic phonemic awareness teaching (Johnston & Watson, 2004). Once these phonics programmes had been administered, the post-test measures discovered that the synthetic phonics group had reading and spelling ages 7 months ahead of their chronological age and outperformed their peers from the analytic phonics groups in both reading and spelling (Johnston & Watson, 2004). Additionally, 15 months after the phonics programmes had ended, the reading and spelling age advantages for the synthetic phonics group were still present, when compared to the analytic phonics groups who had since received synthetic phonics teaching (Johnston & Watson, 2004).

The longitudinal impact of phonics teaching was also investigated by Vadasy and Sanders (2013) who found that when a phonics-based intervention was delivered to early readers in the first grade, there were longitudinal advantages for word reading, reading comprehension and spelling which remained into grades 2 and 3. In an extension of their earlier study (Johnston & Watson, 2004), Johnston et al. (2012) found that the children who had been taught within the synthetic phonics group had advantages in word reading, spelling and reading comprehension longitudinally, compared to an English sample of readers taught through analytic phonics.

Walker et al. (2015) published a report after evaluating the Phonics Screening Check from 2012 to 2015. This review concluded that while reading scores had improved over the four years of the screening check, this trend was apparent before the Phonics Screening Check was introduced. Therefore, they could not separate a growing trend in achievement and the Phonic Screening Check's influence on teaching practice. However, they did find some evidence of phonic attainment improving over these years. Therefore, either phonics is helping with phonic decoding skills but not wider literacy skills, such as orthographic whole-word representations and semantics, or the methodological problems in the review (many phonics programmes covered which may not be comparable and a lack of control group without the Phonics Screening Check or phonics teaching), failed to show the true impact of the screener (Walker et al., 2015).

Referring to the notion that phonics cannot encompass all factors required for early reading, another consideration is that wider independent reading and experience with text would be required to develop whole-word orthographic representations, as outlined in dual-route reading models such as the DRC (Coltheart et al., 2001) and in Share's Self-Teaching Hypothesis (1995). As part of the Key Stage 1 Curriculum, there is some guidance which relates to developing orthographic routes, through teaching "common exception words, noting unusual correspondences between spelling and sound when these occur" (Department for Education, 2013, p. 10). In contrast to the specified 40+ phonemes which must be addressed, the national curriculum does not provide guidance for which exception words should be taught in Key Stage 1. The national curriculum states that the "number, order and choice of exception words taught will vary according to the phonics programme being used"

(Department for Education, 2013, p. 10). Nevertheless, the reading framework does provide some examples of common exception words which enable early readers to access age-appropriate texts, such as "*to*", "*said*", "*they*" and "*some*" (Department for Education, 2022).

Shapiro and Solity (2016) compared two existing synthetic phonics programmes available in the UK during the Reception year of primary school: "Letters and Sounds" (L&S) which delivers multiple GPCs with no exception words taught "by sight" and "Early Reading Research" (ERR) which delivers the most consistent GPCs and high frequency words "by sight", including a mixture of exception and regular words. Regarding exception words encountered through the L&S programme, children are taught to recognise components of the target word which can be phonologically decoded and sound out as much of the word as possible (Shapiro & Solity, 2016). Children were assessed on measures of phonological awareness and reading ability at four time points (start of Reception, end of Reception, end of Year 1 and end of Year 2). With results demonstrating that both phonics programmes were effective for the majority of participants, indicating that the inclusion of exception words and regular words taught "by-sight" in the ERR programme, did not interfere with the effectiveness of synthetic phonics (Shapiro & Solity, 2016). Additionally, there was a small exception word reading advantage for the ERR participants who had poor phonological awareness at the start of Reception, with the researchers suggesting this group benefitted from learning "by-sight" rather than relying on their poor phonological awareness (Shapiro & Solity, 2016). Demonstrating that there is little risk to including exception word teaching as part of the phonics curriculum and potentially a benefit for those early readers who have difficulty along the Non-lexical pathway. In their review, Torgerson et al. (2019) also suggested there is insufficient evidence to justify a "phonics only" teaching programme at present.

While early readers across England are being taught the same phonemes, there is a lack of focus on exception word teaching, and there are no systematic guidelines for teaching this word type. Children may therefore have differing levels of development in their orthographic representations, largely dependent on their level of independent reading, text exposure and which common exception words they are taught. This varied exception word teaching may lead to difficulties with exception words which cannot be phonologically decoded. Therefore, it is of interest to determine how word type reading performances across decodable nonwords and non-decodable exception words have been impacted as a result of systematic, synthetic phonics teaching. In addition to determining if systematic synthetic phonics has had a positive or negative effect on emerging profiles of reading difficulty within early primary school readers.

1.6 Overview of Thesis

This thesis aims to address three outstanding areas of inquiry regarding the impact of systematic, synthetic phonics teaching as part of the English National Curriculum, for early readers and their reading development. The introduction presented above has discussed key areas in relation to this topic, including the theoretical basis which formed the rationale for introducing phonics teaching, from the perspective of dual-route reading models, frameworks and reading development literature. In combination with describing how phonics teaching has been implemented within English primary schools, following the introduction of the 2012 Phonics Screening Check. The first area of inquiry concerns the impact of phonics teaching on the word type reading ability of early readers. Specifically, whether phonics is addressing the Non-lexical route as expected, through facilitating successful phonological decoding and whether the lack of consistent exception word teaching has implications for word type reading performance. Additionally, examining word type reading longitudinally to determine whether word type reading performance over time reflects reading development, according to the dual-route perspective of reading, following initial synthetic phonics instruction. Chapter 2 presents a study of children who have received systematic, synthetic phonics teaching and investigates word type reading performance longitudinally, in combination with word type reading comparisons with a pre-phonics sample.

Secondly, the introduction presented above discussed potential reading difficulties that children can display across word types from a dual-route reading perspective. While this topic is mostly informed through literature with a focus on dyslexia classifications and interventions, the question remains: what reading difficulties do poor readers without a dyslexia classification display, especially since the introduction of mandatory phonics teaching? The second area of inquiry concerns the impact of systematic, synthetic phonics teaching on emerging profiles of reading difficulty shown by early readers. Specifically, whether as a result of phonics teaching, nonword profiles of difficulty have been addressed and are therefore less common, while there is potential for increased exception word profiles of difficulty, due to the aforementioned lack of consistent exception word teaching. Additionally, the longitudinal stability of these emerging profiles of reading difficulties are examined, to determine if reading difficulties within early readers, post-phonics teaching, resemble long-term difficulties which require tailored interventions or resemble short-term difficulties which are addressed with educational input and further reading development. Chapter 3 presents a study of children who have received systematic, synthetic phonics teaching and investigates emerging profiles of reading difficulty and their longitudinal stability both within-sample and when compared with a pre-phonics sample.

Thirdly, the introduction presented above has discussed important skills for reading, including phoneme awareness, letter-sound knowledge, vocabulary and home literacy environment. While synthetic phonics teaching places a particular focus on phoneme awareness and letter-sound knowledge to facilitate phonological decoding, the question remains: how much influence do these wider skills for reading contribute to the GPC learning of early readers receiving synthetic phonics teaching? Additionally, the introduction presented above has outlined the main principles of phonics and how phonics is proposed to facilitate the independent reading of early readers, through enabling successful phonological decoding which facilitates orthographic learning and reading “by-sight”. While there is support for the effectiveness of phonics teaching overall, there is little research investigating the ability to “self-teach” and learn GPCs independently. The third area of inquiry concerns the fundamental principles of synthetic phonics teaching, specifically, whether as a result of phonics teaching early readers are able to detect, learn and generalise novel GPCs and develop independent reading. Additionally, the influence of other important skills for reading are examined, to determine if wider skills outside of phonics teaching such as vocabulary and home-literacy environment, contribute to GPC learning. Lastly, the role of context is examined to investigate the influence of context when learning novel GPCs, which can inform how phonics teaching is delivered to early readers. This final empirical work reported in Chapter 4, presents a study of early readers within Reception and Year 1, who have received systematic, synthetic phonics teaching and investigates their ability to detect, learn and generalise GPCs to novel items without explicit instruction.

2 Study 1a: Reading Performance of Early Readers Post-Phonics Introduction

Abstract

This study investigates word type reading performance across nonwords and exception words within a sample of primary school children who have received systematic, synthetic phonics instruction during the first two years of education. This post-phonics sample was also compared to a sample of primary school children who were classified as pre-phonics, as this standardisation sample was gathered before the introduction of the 2012 Phonics Screening Check. Within this post-phonics instruction sample, the standardised assessments of word type reading revealed that in Year 1 there was an advantage for nonword reading which by Year 4, became an advantage for exception word reading. When both samples were compared, the post-phonics sample had better nonword and exception word reading than the pre-phonics sample at both time points. It is suggested that in the early years of primary school education, phonics instruction was successful, leading to an early nonword reading advantage through a trained Non-lexical pathway. Over time, this movement to an exception word advantage reflects the development of orthographic representations and knowledge, as part of Lexical routes to reading within a dual-route framework of reading development.

2.1 Introduction and Rationale

2.1.1 The Year 1 Phonics Screening Check and Phonics Teaching in England

As discussed in Sections 1.5-1.5.2, there has been an increased focus on systematic, synthetic, phonics teaching as part of the National Curriculum for English, in part due to the introduction of the Year 1 Phonics Screening Check (PSC) in 2012. It is important to note that phonics teaching is not a novel concept and has been in use across primary schools in England before the PSC was introduced, following recommendations by the Rose Review (2006) and government criteria published in 2010 which allowed primary schools to select their own phonics programme. However, since the introduction of the PSC, an increased emphasis on the role of phonics teaching within primary schools in England has been demonstrated through changes to the 2013 National Curriculum and the updated guidance for the English programme of study for reading in 2014. Furthermore, newly published guidance from the UK government stresses the importance of systematic, phonics teaching programmes for teaching word reading and spelling successfully, such as the 2022 Department for Education Reading Framework.

In 2015, the UK government commissioned a review of the impact of the PSC (by the National Foundation for Education Research) which determined that following its implementation, the majority of primary schools included adjusted their phonics teaching, including changes to classroom practice or use of assessment (Walker et al., 2015). Specifically, literacy coordinators reported increasing the pace of phonics teaching, increasing the time devoted to phonics teaching and increasing assessment of progress with phonics in the Reception and Year 1 age groups, during the 2013-2014 academic year (Walker et al., 2015). Interestingly, respondents reported that the increased use of assessment was internal rather than external, with teachers addressing reading difficulties through the results of internal assessments rather than referring to the wider PSC data (Walker et al., 2015). This review concluded that while many of the primary schools were already delivering phonics teaching, the introduction of the PSC did have an impact on an increased focus on nonwords within phonics teaching, as well as the changes regarding phonics delivery and assessment discussed above (Walker et al., 2015).

However, there is limited information available regarding the impact this increased focus on systematic, synthetic phonics teaching has had on the longitudinal reading performance of early readers, who are specifically targeted through this phonics teaching. While there is evidence from the published results of the PSC each year, this information reflects the reading ability of these early readers at the time of issue, namely examining GPC knowledge and phonological decoding ability and is not expected to predict reading performance longitudinally. With some research suggesting that the PSC is not an adequate measure of the whole range of 85 GPCs that early readers are meant to learn during the early years of education (Darnell et al., 2017). Therefore, perhaps the validity of the PSC is limited as a measurement of GPC knowledge at the time of issue. Additionally, Gilchrist and Snowling (2018) suggest that the predictive validity of the PSC requires longitudinal data, in order to determine the longitudinal literacy outcomes for all early readers who both pass and fail the assessment.

Conversely, there are research findings which support the validity of the PSC, such as Duff et al. (2015) who found that the PSC strongly correlated with nonword reading and single-word reading accuracy. Double et al. (2019) compared the longitudinal reading comprehension outcomes for children who initially failed the PSC and then passed when the measure was retaken (fail-pass), compared to children who failed the PSC at both time points (fail-fail) and children who passed the PSC the first time (pass). Their results demonstrated four years after the first PSC, the fail-fail group performed worse than the fail-pass group, who were also outperformed by the pass group on a measure of reading comprehension (Double et al., 2019). Interestingly, these results indicated that the fail-pass group were able to partially catch up with their peers in the pass group, due to improving their phonological decoding skills before the PSC was retaken (Double et al., 2019). Therefore, there may be limited predictive validity within the PSC, with results indicating that PSC performance is related to nonword reading, single-word reading and reading comprehension.

Nevertheless, the results of the PSC and its limited predictive validity have not addressed the wider longitudinal impact of the increased focus on systematic, synthetic phonics teaching. Moreso, this impact of increased synthetic phonics teaching has not been examined across the different word type reading performances of early readers, both within Year 1 when the PSC is issued and longitudinally.

2.1.2 Predictions about how Systematic, Synthetic Phonics Teaching Influences Reading of Different Word Types

Systematic, synthetic phonics teaching focuses on decoding words using GPC knowledge, so words with regular GPCs are prioritised. Children will also be exposed to nonwords (pseudo-words) as part of preparations for the PSC, which contains 20 regular items and 20 nonword items as a measure of phonological decoding. However, as discussed in Section 1.2.2, the English language also contains many exception words which cannot be accurately read through phonological decoding alone, as they do not conform to typical letter-sound mappings. Currently, there are no explicit government guidelines on which exception words are to be taught within English phonics programmes and thus the number and choice of exception words which are directly taught to early readers through synthetic phonics differs according to the programme delivered.

As discussed in Sections 1.2-1.2.3, from a dual-route reading perspective, different word types are read successfully through different routes to reading, namely nonwords which require phonological decoding as they have no existing orthographic, semantic or phonological representations, are read through the Non-lexical route. Specifically, within synthetic phonics, these nonwords are read through applying learnt GPC knowledge within the programme to segment the target word into its constituent sounds, which are blended together to produce the sound of the whole word. Conversely, exception words which cannot be read through phonological decoding along the Non-lexical route due to regularization errors, are theoretically accessed through “whole-word” reading along the Lexical route. These exception words are not taught through this phonological decoding process within synthetic phonics, but rather included in phonics programmes as “sight words” or “tricky words” to be read “by sight” once early readers are familiar with the word through repeated exposures. Reading along this Lexical route is either direct from the orthographic representation of the familiar exception word which activates the phonological representation of the word, or indirectly through the semantic representation of the exception word.

From a dual-route reading perspective, systematic, synthetic phonics teaching should strengthen the Non-lexical route to reading, through improving both GPC knowledge and phoneme awareness, which are required for successful phonological decoding (Bowyer-Crane et al., 2008; Hulme at al., 2012; Share, 1995). Additionally, systematic, synthetic phonics teaching should improve phonological decoding ability, through teaching early readers to apply their GPC knowledge and synthesize the sounds of a target word for pronunciation. Therefore, reading performance amongst early readers for words which require phonological decoding along the Non-lexical route, such as nonwords, should be improved following systematic, synthetic phonics instruction.

In contrast, this systematic, synthetic phonics teaching is not expected to improve the Lexical routes to reading directly, as the phonological decoding ability taught through phonics is a synthetic approach rather than a whole-word "sight reading" approach required for the Lexical routes to reading. Therefore, reading performance amongst early readers for words which require the Lexical route, such as exception words which cannot be phonologically decoded accurately due to their irregularities within letter-sound mappings, is not expected to improve following systematic, synthetic phonics instruction. One possibility is that this phonics teaching does not strengthen the Lexical route to reading directly but instead indirectly, through teaching early readers to phonologically decode novel words through applying their growing GPC knowledge. According to Share's Self-Teaching Hypothesis (1995), this phonological "recoding" ability acts as a "self-teaching" mechanism, whereby orthographic learning occurs through the "lexicalization" of the recoding process. Specifically, orthographic knowledge is gained through instances of successful phonological recoding, ultimately creating orthographic representations of decodable words without the need for additional support.

Nevertheless, the longitudinal reading performance between nonwords and exception words, following systematic, synthetic phonics instruction has not been directly compared. According to the dual-route perspective of reading discussed above, it would be expected that nonword reading performance would improve following systematic, synthetic phonics instruction, due to improvement in GPC knowledge, phoneme awareness and phonological decoding. Exception word reading is not expected to improve directly as a result of systematic, synthetic phonics teaching and two outcomes are possible. The first outcome is that exception word reading performance is poor, due to a lack of consistent exception word teaching across phonics programmes and this word type not benefitting from improvement along the Non-lexical route. The second outcome is that exception word reading performance is improved indirectly, through the improvements in phonological decoding and the Non-lexical route acting as a "self-teaching" mechanism, which leads to orthographic learning and developing adequate whole-word "sight reading" for exception word items, as in the developmental hypothesis of Share (1995). Nevertheless, these predictions are complicated by natural reading development which occurs throughout primary school.

2.1.3 Changes in Performance for Different Word Types as Reading Develops

Section 1.2.1 outlined that according to the dual-route perspective of reading, early readers begin with a reliance on the Non-lexical route, whereby words are read aloud through the phonological decoding process in a laboured, slower form of reading than that of a skilled adult reader (Castles et al., 2018). As reading develops, the reader is then expected to transition from reading solely along the slower Non-lexical route through phonological recoding to reading "by sight" along the Lexical route as a faster, less labour intensive form of reading (Stuart & Stainthorp, 2015). To facilitate this Lexical reading route, orthographic learning must occur to provide the reader with orthographic knowledge, including word-specific knowledge such as semantic information and general orthographic knowledge about the writing system such as positional letter effects (Castles et al., 2018; Share, 1995). This growing experience with text through independent reading (facilitated through phonological decoding), then provides the reader with the opportunity to develop orthographic representations of familiar words and gradually rely less on phonological decoding (Grainger et al., 2012). Ultimately leading to development of the Lexical route, whereby the orthographic representation of a target word links directly to its phonological representation for pronunciation, or indirectly through a semantic representation of the target word, resulting in what is considered "whole word reading" or "sight reading", as there is no requirement for sublexical phonological decoding (Stuart & Stainthorp, 2015). It is important to note that both reading pathways are expected to operate in tandem according to the DRC (Coltheart et al, 2001) or through a division of labour approach according to the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989), however as discussed in Section 1.2.2, these reading pathways can be assessed through reading performance of different word types (e.g., nonwords and exception words).

Therefore, it is of interest to determine how word type reading performance changes over time as reading develops. Especially following initial systematic, synthetic phonics instruction which as discussed above may provide an initial advantage for nonword reading due to training along the Non-lexical route. According to the dual-route predictions of reading development, it is possible that early readers who have received synthetic phonics instruction may have a nonword reading advantage initially as they are relying on the Non-lexical route. Meanwhile, these early readers may have poor exception word reading performance as their Lexical route to reading has not yet developed, as they are still refining their phonological decoding skills and Non-lexical reading and have not yet made orthographic representations for whole-word reading. Over time as reading develops, this may change as the Lexical route develops and therefore there may be an advantage when reading exception words which rely on the Lexical route. However, the Non-lexical route is not expected to worsen over time according to the dual-route perspective, as it is often utilised for decoding novel words without existing orthographic representations. Therefore, nonword reading performance is not expected to decline longitudinally. Ultimately, it is of interest to determine if word type reading performance longitudinally corresponds with the dual-route perspective of how the two routes to reading develop.

In summary, there are two main questions that arise from the dual-route reading perspective and the impact of this increased focus on systematic, synthetic phonics teaching, since the introduction of the PSC in 2012. Firstly, does the longitudinal word type reading performance of early readers who have had synthetic phonics instruction, reflect reading development along the two routes to reading from a dual-route perspective? Secondly, does systematic, synthetic phonics teaching within the first two years of primary school, have an impact on word type reading performance. That is, is there a nonword reading advantage and an exception word reading disadvantage due to phonics teaching?

2.1.4 Research Questions

This study sought to address two gaps in the existing literature through firstly investigating how word type reading changes longitudinally from Year 1 to Year 4, as part of early readers' reading development throughout primary school. Secondly, this study investigated the impact of this increased focus on systematic, synthetic phonics teaching, through examining word type reading performance between a pre-phonics and post-phonics sample, to determine if synthetic phonics teaching had outcomes for word type reading that correspond with predictions from the dual-route reading perspective.

In terms of longitudinal reading performance, it was hypothesised that following the development of the two routes to reading outlined in the dual-route reading perspective, initial reading performance at Year 1 would be better for nonword items than exception word items amongst early primary school readers. This hypothesis was informed by the view of reading development from a dual-route perspective (i.e., early readers are relying on the Non-lexical route and phonological decoding, resulting in successful nonword item reading). Meanwhile, these early readers have not yet developed a Lexical route to reading and may be reading exception words incorrectly through phonological decoding, resulting in poor initial exception word reading.

Additionally, it was hypothesised that as reading develops, at Year 4 exception word reading performance would improve, to reflect development along the Lexical route to reading and the creation of orthographic, phonological and semantic representations of exception word items. Nonword reading was predicted to remain stable longitudinally, as the Non-lexical route is not expected to worsen over time and is required for decoding unfamiliar words without orthographic representations according to the dual-route reading perspective.

Regarding the potential impact of the increased focus on systematic, synthetic phonics teaching on the word type reading performances of a post-phonics sample, it was hypothesised that nonword reading performance would be better within the post-phonics sample compared to a pre-phonics sample. If systematic, synthetic phonics teaching is effective in teaching phoneme awareness, GPC knowledge and phonological decoding, it was hypothesised that this would facilitate nonword reading along the Non-lexical route to provide a nonword reading advantage to the post-phonics sample. Additionally, nonwords are explicitly taught within synthetic phonics teaching, both in preparation for the Year 1 PSC, but also as a method of teaching phonological decoding. Therefore, resulting in a nonword reading advantage for the post-phonics sample.

Alternatively, no specific hypotheses were made regarding the exception word reading performance between the pre-phonics and post-phonics sample and was instead viewed as an exploratory investigation. According to the dual-route perspective, systematic, synthetic phonics teaching was not predicted to improve exception word teaching directly, as whole-word exception word training varies across phonics programmes and is not addressed systematically due to a lack of mandatory requirements. Therefore, one possibility was that exception word reading performance would be similar across the two samples, as synthetic phonics had not improved exception word reading in the post-phonics sample.

Conversely, exception word reading performance may have been improved following systematic, synthetic phonics teaching, through an indirect improvement in the Lexical route to reading. Specifically, phonological decoding taught through phonics teaching may act as a "self-teaching" mechanism, whereby early readers can independently read a variety of texts and increase their exposure to print (Share, 1995). Resulting in growing orthographic knowledge and the creation of orthographic representations required to develop the Lexical route to reading, which facilitates exception word reading. In combination with the limited amount of exception words taught through synthetic phonics, one possibility was that exception word reading performance may be worse in the pre-phonics sample, compared to the post-phonics sample, through a lack of indirect Lexical route improvement.

The two research questions addressed in this study are as follows:

Question 1. Does word type reading performance change as reading develops, following systematic, synthetic phonics instruction?

Question 2. Does word type reading performance differ between a pre-phonics and post-phonics sample? Following the introduction of systematic synthetic phonics teaching to the UK English curriculum and the introduction of the Phonics Screening Check during Year 1.

2.2 Method

This chapter discusses two samples of secondary data which have been analysed to investigate how reading performance varies across word types, years of education and between pre-phonics and post-phonics samples. The Aston Literacy Project (ALP) conducted by Aston University, began in 2011 with the aims of developing current understanding of skills which impact upon literacy development, including auditory and phonological skills. In combination with identifying patterns of difficulty shown by poor readers, to enable early identification of children who are at risk of reading difficulties. The ALP is longitudinal in nature, having recruited 16 Birmingham primary schools and testing participants from Reception to Year 6. Before the PhD programme, I contributed to data collection on the ALP in years 2016 and 2017. Therefore, prior to 2016 and following 2017, this chapter will be discussing secondary data from the wider ALP. The period of data collection for the ALP coincided with the first year of the mandatory Phonics Screening Check introduced by the UK Government, with the Year 1 sample having recently completed their Phonics Screening Check in 2012. Therefore, the ALP sample is considered a post-phonics introduction sample, as the participating children in Year 1 had already been taught GPCs and decoding through synthetic phonics both in Reception and Year 1, in preparation for the Year 1 Phonics Screening Check.

The Diagnostic Test of Word Reading Processes (DTWRP) is a package of word reading assessments designed by The Forum for Research in Literacy and Language and GL Assessment, designed for use with six to twelve-year-olds. In 2009 and 2011, the DTWRP standardisation sample was collected, to gather data which would form the age adjusted stanine scores for the assessment package. Unlike the ALP, the DTWRP sample was not longitudinal, instead sampling different participants from each age group, from Reception to Year 7. As the data collection for the DTWRP standardisation sample began three years before the introduction of the Phonics Screening Check and mandatory synthetic phonics teaching in the UK English curriculum, this sample is considered a pre-phonics sample. As discussed earlier, synthetic phonics teaching existed before the Phonics Screening Check, with the UK Government providing limited guidance on phonics programmes from 2010. However, it is unlikely that consistent synthetic phonics teaching was implemented across the four regional centres that were sampled in the DTWRP, before this was introduced into the English curriculum. Hereafter, this chapter will refer to the two samples as the ALP sample and DTWRP sample, with a particular focus on the data from Years 1 and 4 of each sample.

2.2.1 Participants

2.2.1.1 ALP Sample

In the first year of the ALP, parents of the Reception class children in each school were contacted to receive voluntary parental consent for their children to be tested, with opt-out forms used in all schools except one (when the school leadership team requested an opt-in strategy). In the later years of data collection, parents were contacted and given an opt-out form if they wished to withdraw their child from the study. This resulted in a representative sample where all children were tested unless their parents withdrew them from participating, or they were not present during data collection for reasons such as illness. Following the Reception data collection in the first year, 94% of these children participated in Year 1, while in Year 4, 80% of these children participated, including data from 9 new participants who did not take part in Year 1.

This method of recruiting participants resulted in: 783 participants in Reception, with a mean age of 63 months (5.25 years), 727 participants in Year 1, with a mean age of 75 months (6.25 years) and 579 participants in Year 4 with a mean age of 111 months (9.25 years) as some children had been withdrawn or moved to a non-participating school by Year 4. Table 1 below demonstrates this sample attrition from Year 1 to Year 4. Across the ALP sample, there were 414 males and 459 females. At school entry (Reception), 10% of the sample had English as an additional language (EAL), 18% of pupils received free school meals and 3 children had a statement of SEN. Data from this same sample have been reported elsewhere (e.g., Cunningham et al., 2021) but the analyses reported in this thesis are original and stand alone.

Table 1

ALP Participants across Year 1 and Year 4

	Year 4	**Participated**	**Missed**	**Total**
Year 1	**Participated**	570	157	727
	Missed	9	0	9
	Total	579	157	736

2.2.1.2 DTWRP Sample

The DTWRP standardisation sample recruited 34 schools from four regional centres across England with varying socioeconomic and geographical backgrounds. In total, 1125 participants were voluntarily recruited from Reception to Year 7 as follows: 143 participants in Reception, with a mean age of 64 months (5.33 years), 143 participants in Year 1, with a mean age of 75 months (6.25 years) and 143 participants in Year 4, with a mean age of 111 months (9.25 years). Across the DTWRP sample which includes children from all primary school years, there were 554 males and 571 females. 12% of the sample had English as an additional language (EAL), 17% of the sample were on their school's Special Educational Needs (SEN) register and 26% of pupils received free school meals (FSM). The original dataset was requested in order to gain more accurate demographic information for the relevant year groups (Reception, Year 1 and Year 4), but the complete dataset was not available. These figures therefore cannot be directly compared to those reported for the ALP sample above. Although EAL and FSM status are likely to be fairly consistent across the year groups, but SEN will increase significantly during primary school as children are identified (it is rare to enter school with SEN which may explain why the ALP sample had many fewer children with a statement).

Nevertheless, the ALP and DTWRP samples can be considered broadly equivalent in terms of EAL, but the DTWRP sample had more children in receipt of FSM.

2.2.1.3 Power Calculations

This study utilises two retrospective, secondary data samples in the form of pre- and post-phonics groups. A priori power analyses were conducted to ensure that the analyses within the study would have an adequate sample size to investigate the research questions, through detecting effects, if present. To address the first research question, a priori power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007, 2009) to determine the minimum sample size required to achieve 0.8 power for detecting a medium effect (Cohen's dz = 0.5), at a significance criterion of $\alpha = 0.05$ for Wilcoxon signed-rank tests, which produced a result of N = 47.

To address the second research question, a priori power analysis was also conducted using G*Power version 3.1.9.7 (Faul et al., 2007, 2009) to determine the minimum sample size required to achieve 0.8 power for detecting a medium effect (Cohen's f = 0.25) at a significance criterion of $\alpha = 0.05$ for one-way ANCOVAs, which produced a result of N = 128.

Therefore, the larger required sample size was selected as a basis for Study 1a sample size. Thus, the obtained sample sizes of N = 727 from the Year 1 ALP sample (N = 570 in Year 4) and N = 143 participants from the DTWRP sample at both Year 1 and Year 4 is adequate to test the study research questions.

2.2.2 Design

The design for this study varies between the two research questions and the two samples of secondary data. The first research question asks if the ability to read different word types changes over time as part of reading development and utilises the longitudinal data of the ALP sample. The analyses for this research question are within-sample, focusing on the word type reading performance of the ALP sample from Year 1 to Year 4 and are a mixture of descriptive and correlational approaches.

The second research question asks if a pre-phonics and post-phonics sample differ on their word type reading ability, utilising and comparing the DTWRP sample (pre-phonics) and ALP sample (post-phonics). The analyses for this research question are between-samples, comparing the samples through a quasi-experimental design.

2.2.3 Measures

2.2.3.1 DTWRP Assessment Package

The Diagnostic Test of Word Reading Processes (DTWRP) is an assessment package designed by The Forum for Research in Literacy and Language and GL Assessment, designed for use with six to twelve-year-olds (DTWRP; Forum for Research into Language and Literacy, 2012). The DTWRP assesses word reading processes across three different word types: nonwords, exception words and regular words (Stuart & Stainthorp, 2015). The assessment package consists of three subtests (nonwords, exception words and regular words) each assessing reading performance of a separate word type across 30 individual items, whereby a child is asked to read each item aloud and an accuracy score for each item is recorded (Stuart & Stainthorp, 2015). The DTWRP requires 30 minutes to administer and a stop rule is applied on each subtest, after five consecutive errors made by a participant, the test is stopped and the next test is presented.

The DTWRP nonword assessment, similar to the Phonics Screening Check, can only be completed accurately if the items are read through phonological decoding processes as these items do not have orthographic or semantic representations (Stuart & Stainthorp, 2015). Whilst the DTWRP exception word assessment can only be completed accurately if the items are read through lexical processes, (e.g., utilising existing orthographic representations for the target words, as these words cannot be read accurately through phonological decoding alone) (Stuart & Stainthorp, 2015). The DTWRP is based on the DRC model of reading (Coltheart et al., 2001) and aims to both assess word reading processes and identify children who show relative weaknesses across the three word types (Stuart & Stainthorp, 2015).

Two scores are given for each subtest of the DTWRP, first the raw accuracy score for each set of 30 items, ranging from 0-30. Secondly, an age adjusted stanine score on a scale of 0-9, gathered from the DTWRP standardisation sample collected in 2009 and 2011. These stanine scores can help identify children who have not made the expected word reading progress over the academic year, as stanine scores are comparable across years and are thought to be stable if the child is making the expected progress for their age (Stuart & Stainthorp, 2015). Additionally, these stanine scores are used to allocate reading difficulty profiles to children, based on the child's profile of strengths and weaknesses in phonological decoding and lexical word reading processes (Stuart & Stainthorp, 2015). Specifically, children whose scores fall below a stanine score of 4 on one word measure (e.g., nonwords) may indicate a difficulty with that word type when considered with a stanine score of 4 or above with the other word types (e.g., exception words). These profiles can then be used to design targeted interventions and teaching plans for that specific child through addressing their word reading weaknesses (Stuart & Stainthorp, 2015).

2.2.3.2 ALP Assessments

During the ALP data collection, the DTWRP assessment package was administered with participants in both Year 1 and Year 4, including all three subtests of nonword reading, exception word reading and regular word reading. The ALP data collection also included non-standardised tests, two of which will be discussed here as they form part of the analyses conducted in this chapter. The first measure of Orthographic Choice was taken from Olson et al. (1994) and asked participants to view a pair of words on a laptop screen and select which word was a "real" word which was spelt correctly from the pair, for example, "*salmon*" and "*sammon*". The second measure of Phonological Choice was adapted from Olson et al. (1994) and asked participants to view a pair of words on a laptop screen and select which word was a "real" word which sounded correct from the pair, for example, "*sharf*" and "*skore*". Olson et al.'s (1994) original phonological task used three words, whereas this task was simplified to give just two choices per trial, following the same format as the Orthographic Choice task. The full list of items used within the Phonological and Orthographic Choice tasks can be found in Tables 11 and 12, Appendix 7. Additional measures such as digit span, British Ability Scales (BAS) regular word reading and PhAB phoneme deletion were also administered, however these will not be discussed further as they did not form part of the analyses conducted in this chapter. Further information regarding tasks administered during ALP data collection which are not discussed here can be found in Cunningham et al. (2015).

2.2.3.3 Validity Check for Nonword versus Exception word measures on the DTWRP (ALP Data)

This validity check analysis concerns correlations between the DTWRP subtests of nonword and exception word reading and additional assessments that were administered in the ALP sample at Year 1 and Year 4, which did not belong to the DTWRP assessment package. As discussed in Section 2.2.3.2, assessments of Phonological Choice and Orthographic Choice were given to children in the ALP sample to assess the same abilities assessed through the DTWRP subtests. In the case of Phonological Choice, children were assessed as to whether they could phonologically recode two words, to then determine which sounded like a real word. The Orthographic Choice measure assessed if a child could choose between two words which when phonologically decoded sound the same, but only one represented the spelling of a real word.

It was hypothesized that these choice tests tap into the same phonological and orthographic processes as those measured in the DTWRP nonword and exception word reading. Specifically, nonword reading and phonological choice tasks both utilise the Non-lexical, phonological recoding route. In contrast, exception word reading and orthographic choice utilise the Lexical, sight-reading pathway, as Non-lexical phonological recoding alone would not be sufficient during these tests. If this hypothesis regarding the underlying processes of each task is correct, then these measures should be positively correlated. Therefore, this analysis also provides a validity check of the DTWRP subtests, through determining that each measure assesses the underlying processes associated with reading each word type.

This correlational analysis determines the strength of the relationships between the two proposed measures of orthographic and phonological processes (DTWRP assessments and choice tests) to investigate if the DTWRP assessment package was a valid measure of these reading processes. As the Orthographic and Phonological Choice tests do not create a stanine score, raw scores between each measure were used. A non-parametric test was chosen as the ALP nonword and exception word raw scores were abnormally distributed.

Phonological Choice at Year 1 was excluded from this analysis, as many children performed at floor level in this test within this age group. Spearman's Rho correlation coefficients were used to assess the relationship between the DTWRP and ALP Phonological and Orthographic Choice tests. A correlation matrix is provided in Table 2 below. As hypothesised, Phonological Choice and DTWRP nonword reading were moderately, positively correlated at Year 4 ($r_s(481) = 0.54$, $p < 0.01$). Whilst Orthographic Choice and DTWRP exception word reading were strongly positively correlated at Year 1 ($r_s(722) = 0.75$, $p < 0.01$) and moderately positively correlated at Year 4 ($r_s(481) = 0.63$, $p < 0.01$).

Table 2

Spearman's Rho Correlation Matrix for DTWRP and ALP Choice Assessments

Measure	1	2	3	4	5	6
1. Orthographic Choice Year 1	-					
2. DTWRP Exception Year 1	.75	-				
3. Orthographic Choice Year 4	.22	.30	-			
4. DTWRP Exception Year 4	.21	.25	.63	-		
5. Phonological Choice Year 4	.19	.23	.46	.49	-	
6. DTWRP Nonword Year 4	.32	.43	.56	.61	.54	-

Note. All correlations achieved a significance of p<0.01.

2.2.4 Procedure

2.2.4.1 ALP Procedure

During the ALP data collection, each participant worked individually with a researcher in a quiet testing area away from their main classroom. The researcher first explained the aims of the tasks through reading a script found on the child consent form. The child was asked to write their first name on the consent form to indicate that they were happy to participate. The child was also presented with a reward card on which their progress would be recorded through congratulatory stickers. If the child did not wish to participate, they were taken back to their classroom and the next participant was brought to the testing area. Children completed the ALP assessments over three sessions, with each session occurring on separate days during the school week. For the purposes of this study, tasks administered in Session 1 and Session 2 will be discussed as no data was included from Session 3 in this study.

The first task administered upon completion of the consent form in Session 1 was the DTWRP nonword assessment. During this task, the DTWRP nonword printed booklet was displayed in front of the participant. The researcher first read a script to outline the task to the child including instructions and an example nonword to be read aloud by the child as a practice item. The researcher then pointed to each individual nonword item and the participant was prompted to read this word aloud, moving through each item until all 30 nonwords had been attempted. Following the DTWRP stop rule, if five consecutive errors were made by the participant, the task was stopped. Researchers provided no corrective feedback however provided encouragement as the participant read items aloud. The researcher recorded an accuracy score for each nonword read aloud (1 for correct, 0 for incorrect). Once this task was completed, the researcher congratulated the child and provided them with a sticker for their reward chart.

The second task administered was the DTWRP exception word assessment. Following a similar procedure to the task above, the DTWRP exception word printed booklet was displayed in front of the participant. The researcher then read another script to outline the task to the child and began the task by pointing to the first word for the child to read aloud. As before, the researcher pointed to each word individually until all 30 items had been attempted, providing no corrective feedback. Following the same DTWRP stop rule, the task was ended if the participant made five consecutive errors. A score of 1 or 0 was recorded for each item (1 for correct, 0 for incorrect). The child was congratulated for their effort and given a second reward chart sticker.

The third task administered was the DTWRP regular word assessment, completing the three DTWRP subtests. As this does not form part of the analyses conducted in this chapter, this will not be discussed further.

The following day, the child completed the tasks of Session 2, there were three tasks administered on this day, however only two tasks form part of the analyses conducted in this chapter and therefore the digit span task will not be discussed. The first task presented in Session 2 was the Orthographic Choice task, which began with the researcher reading a script with instructions to the participant, followed by an example item. During this task, the child viewed a laptop screen which displayed a pair of words which when phonologically decoded, produce the same sound, for example, "*room*" and "*rume*". The child was given a stylus pen and asked to press the "real" word beginning with 4 example items. For a correct answer during the example items, a smiley face icon was displayed. For an incorrect answer during the example items, a sad face icon was displayed. The researcher then prompted the participant to think carefully about the words whilst not providing corrective feedback. The child then completed 3 blocks (levels 1-3) of 10 items following the example items, with the E-Prime software recording correct and incorrect responses during the task. Alternatively, the task had an automatic stop rule when the child scored 2 correct words or fewer during levels 1 and 2. Following the completion of the Orthographic Choice task, a reward sticker was given for the participants reward chart.

The second task during Session 2 was the Phonological Choice task, which following the same format as above, began with the researcher reading a script and task instructions. Similarly to the Orthographic Choice task, the child viewed a pair of words on the laptop screen, both of which were nonwords, but when phonologically recoded, only one sounds like a "real" word, for example, "*nite*" and "*kile*". The child was asked to press the "real" word with their stylus pen and ignore the other nonword. The task followed the same format as above, with 4 example items followed by 3 blocks of 10 items (levels 1-3) with no corrective feedback given. As before, E-Prime recorded participant responses and an automatic stop rule of 2 or fewer correct answers during levels 1 and 2 was applied. Following the completion of the Phonological Choice task, a reward sticker was given to the participant.

2.2.4.2 DTWRP Sample Procedure

The DTWRP standardisation sample is secondary data of which I was not involved with the data collection. Therefore, there is limited information to report regarding the procedure of the standardisation sample data collection. In accordance with the DTWRP manual which provides guidance for administering the tasks, the three DTWRP subtests were administered with a similar procedure as the ALP sample (DTWRP; Forum for Research into Language and Literacy, 2012). This includes participants working individually with a researcher, reading aloud from the same printed word stimuli booklets, researcher recording of raw accuracy scores and an automatic stop rule of 5 consecutive errors.

2.3 Results

2.3.1 Data Analysis Strategy and Exploration of Data

The analyses for this study vary according to the research question addressed. The first research question asked if word type reading performance differed between nonwords and exception words in a sample of children taught through intensive synthetic phonics (ALP).

Distributions for both raw and stanine word reading scores were analysed through histograms and descriptive statistics. This analysis presented direct distribution comparisons between word types and across year groups, to determine if word type reading performance varies by word type (nonwords versus exception words) and across year groups as the children moved through the primary school curriculum. Stanine score distributions were also compared to normal distributions for stanine scores to determine if the ALP sample differed from the normal distribution. This analysis also informed the use of parametric or non-parametric tests.

Raw scores were significantly different from a normal distribution, requiring non-parametric tests. In contrast, stanine scores were not significantly different from a normal distribution, therefore analyses with this data utilised parametric tests. Wilcoxon signed-rank tests using raw scores were conducted to determine if nonword and exception word reading scores at Year 1 and Year 4 were significantly different within the ALP sample at each time point, as a non-parametric equivalent to a within-subjects t-test. This determined if word reading performance within the ALP sample was more successful with one word type than the other, for example, nonword reading is more accurate than exception word reading within either time point, as a result of their early systematic, synthetic phonics teaching.

The second research question asked if word type reading performance differed between a pre-phonics (DTWRP) and post-phonics sample (ALP), as a result of intensive synthetic phonics teaching. In order to compare the DTWRP standardisation sample and the ALP sample on their word type reading performance, a series of one-way ANCOVAs, with age as a covariate to statistically control for the potentially compounding variable of relative age, were conducted using raw scores for nonword and exception word reading. These ANCOVAs were conducted using Year 1 and Year 4 data from each sample. Descriptive statistics were also given for each samples nonword and exception word reading performance across Years 1 and 4. Due to the ALP data displaying an abnormal distribution and requiring non-parametric analyses, other potential analysis options were considered before the parametric ANCOVA analyses were conducted. The first option of a Mann-Whitney U test would have allowed for comparison of median group differences between the ALP and DTWRP samples, but not statistically accounted for the potentially compounding variable of relative age. The second option of a one-way ANOVA also presented the same issue, firstly requiring parametric data and secondly failing to control for the covariate of age and as a result, potentially reducing the error variance. The third option of a Multiple Regression would have investigated predictions rather than group differences, whereas this analysis aimed to compare retrospective reading performance across existing groups. The Multiple Regression option also required parametric data, displaying the same limitations as the ANCOVA and ANOVA options, but was also less suitable for the aim of the analysis. The fourth option considered was a MANCOVA, which would have required multiple dependant variables. In this analysis, the variables of nonword and exception word reading would have been combined, therefore providing less information regarding word type reading within the

ALP and DTWRP samples. It would not have identified whether there was a statistically significant mean difference between the two groups from each of the word types, therefore, providing less information than the ANCOVA. While the ANCOVA analysis did not satisfy all of the required assumptions, specifically, the use of parametric data and the lack of homogeneity of variances in Year 4, this is acknowledged and therefore the results of this analysis should be interpreted with caution.

At this point it is important to discuss when raw scores (0-30) were used versus stanine scores (0-9). Both types of score are gained for each subtest of the DTWRP assessment, however the analyses presented in this chapter differ on which score type was used. Raw scores are ideal for comparing between measures within the DTWRP assessment and non-standardised assessments for which there are no stanine scores. As these raw scores are based on word reading accuracy, they also provide more detail than standardised stanine scores. As discussed later in Section 3.3.1, raw scores also allow for norming within a sample independent to the DTWRP standardisation sample.

In contrast, stanine scores are age adjusted, which allows us to view the position of an individual child amongst the wider sample in terms of their word reading performance scoring amongst the average, below average or above average. Stanine scores are also directly comparable across the three DTWRP subtests, whereas raw scores are not expected to be exactly the same across time and are not as comparable across different word type measures. As discussed later in Section 3.2.3, stanine scores also allow us to allocate a profile of reading difficulty to a child based on their performance across word types in the DTWRP subtests.

2.3.2 Question 1: ALP Word Type Reading Performance (Raw Scores)

The following results display the ALP raw score distributions across the measures of nonword and exception word reading, to determine if word reading accuracy varied by word type at Year 1. Figures 2.1 and 2.2 below display nonword and exception word reading raw scores at Year 1.

At Year 1, the distribution for nonword reading had a moderate positive skew (0.55) with the majority of children gaining total scores below 15 out 30 displaying average to below average nonword reading. Specifically, the nonword reading distribution peaked at scores of 8-12. However, 126 children (17.33%) achieved high total scores of 20 and above, with 2 children achieving the maximum total score of 30.

In contrast, exception word reading at Year 1 displayed more variance than nonword reading (47.42 and 46 respectively) with a wide mixture of children performing above and below average. Specifically, the distribution peaked at scores of 6, 11 and 18 demonstrating varied exception word reading performance. In contrast to nonword reading, no participants achieved the maximum total score of 30 when reading exception words and fewer children (114, 15.68%) gained a total score of 20 and above. Additionally, more children scored a total of 0 on exception word reading than nonword reading (32 and 12 respectively).

Nonword reading at Year 1 also had a higher mean score (Mean = 12.07, SD = 6.77) compared to exception word reading at Year 1 (Mean = 11.53, SD = 6.89). A Wilcoxon signed-rank test demonstrated that this difference in raw scores was statistically significant ($Z = -3.79$, $p < 0.001$).

Figure 2.1

ALP Year 1 Nonword Reading Raw Score Distribution

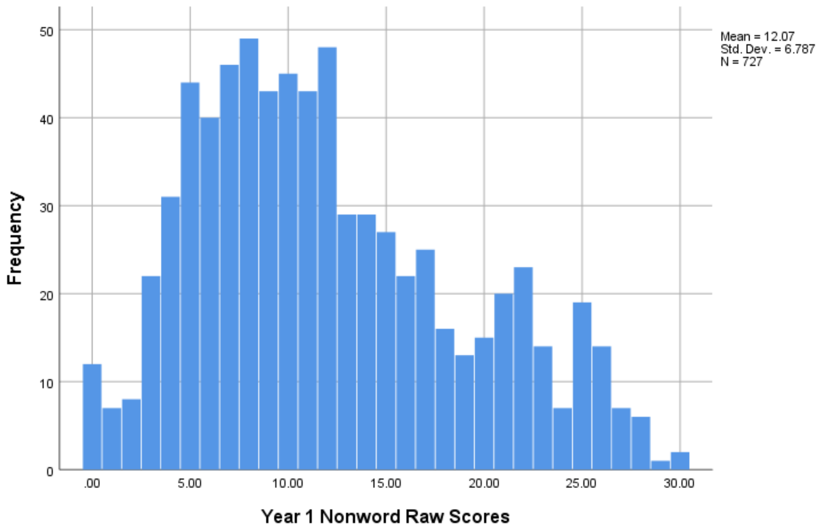
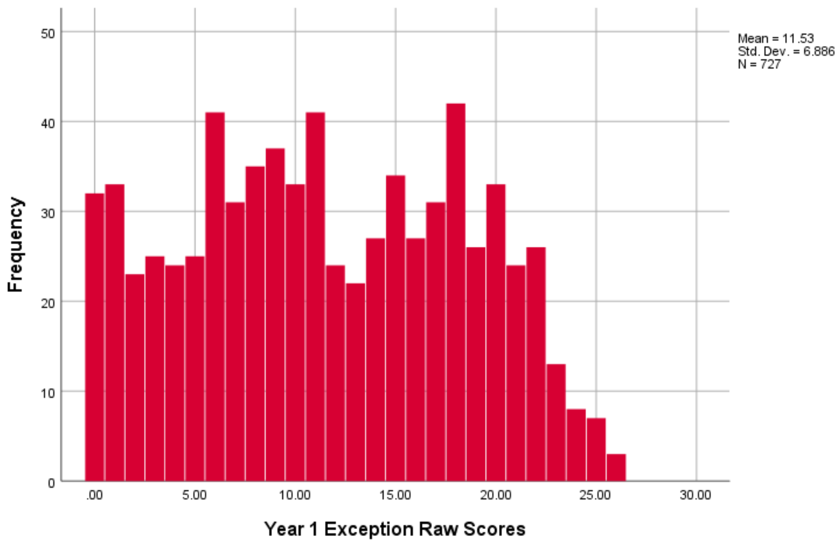


Figure 2.2

ALP Year 1 Exception Word Reading Raw Score Distribution



The histograms in Figures 2.3 and 2.4 below display nonword and exception word reading raw scores at Year 4. At Year 4, both nonword and exception word reading scores were highly negatively skewed, strikingly with exception words (-1.03 and -1.78 respectively). These distributions indicate that at Year 4, most children were performing above average when reading both word types. When reading nonwords, 423 children (73.06%) displayed high total scores of 20 and above and when reading exception words, 520 children (89.81%) displayed high total scores of 20 and above.

However, the nonword reading distribution contained more variance than the exception word reading distribution (34.65 and 15.61 respectively) demonstrating a wider variation in nonword reading performance compared to exception word reading. While both distributions peaked around scores of 25-27, there was more variance in children scoring below a total score of 20 at Year 4 when reading nonwords.

By Year 4, it is clear that word reading performance had improved for both nonword and exception word reading. At Year 4, in contrast to Year 1, exception word reading achieved a higher mean score (Mean = 23.84, SD = 3.95) than nonword reading (Mean = 22.22, SD = 5.89). A Wilcoxon signed-rank test demonstrated that this difference in raw scores was statistically significant ($Z = -7.30$, $p < 0.001$). This therefore shows an interaction pattern where an initial advantage for nonword reading changes to an advantage for exception word reading at Year 4. Due to the non-normal distributions of the data, it was not possible to test whether this was a significant interaction through parametric analyses.

Figure 2.3

ALP Year 4 Nonword Reading Raw Score Distribution

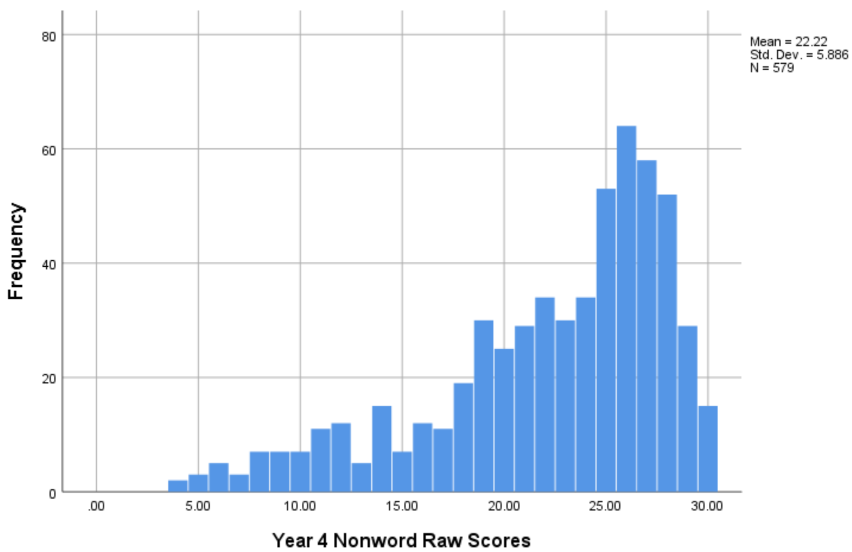
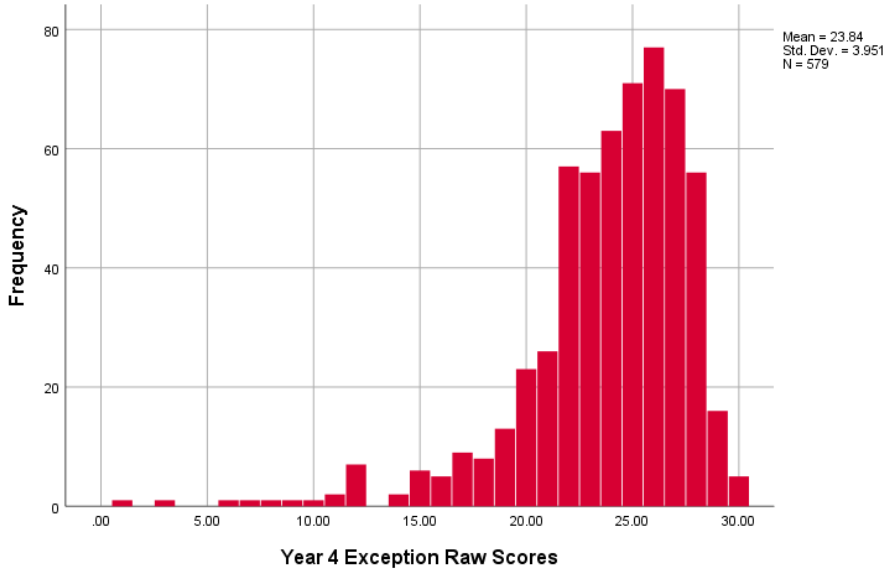


Figure 2.4

ALP Year 4 Exception Word Reading Raw Score Distribution



2.3.3 Question 1: ALP Word Type Reading Performance (Stanine Scores)

As discussed earlier, raw scores are not expected to be comparable across year groups, as performance is expected to change over time. Alternatively, standardised stanine scores are available from the DTWRP assessment package, which are comparable across year groups. Each stanine score histogram has a normal distribution curve overlaid (Mean = 5, SD = 2) to demonstrate how closely the data fits to a normal distribution. An example of this normal distribution can be seen in Figure 2.5 below. Followed by Table 3 which describes the typical percentage of children we would expect to achieve each stanine score under this normal distribution, as outlined by GL Assessment (GL Education Support, n.d.). Figures 2.6 and 2.7 below display nonword and exception word stanine scores at Year 1.

Figure 2.5

Normal Distribution for Age Standardised Stanine Scores

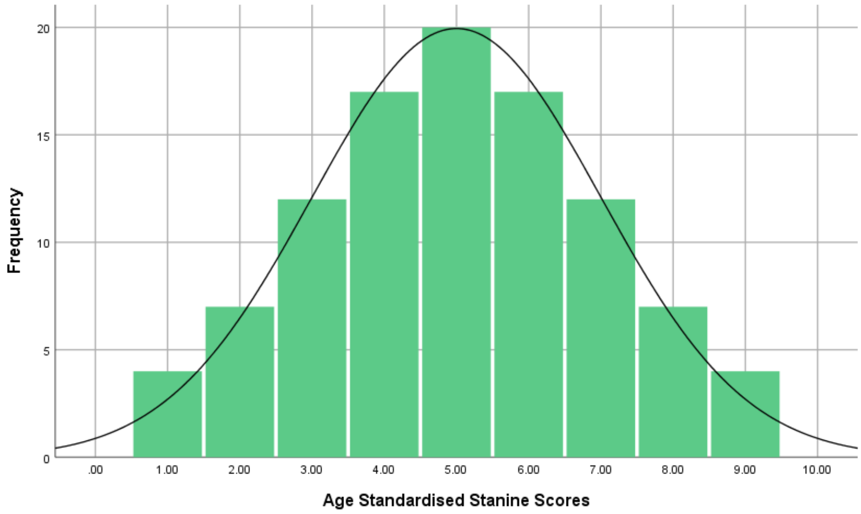


Table 3

Percentage of Stanine Scores achieved within a Normal Distribution

Stanine Score Achieved (1-9)	Expected Percentage of Sample
1	4%
2	7%
3	12%
4	17%
5	20%
6	17%
7	12%
8	7%
9	4%

At Year 1, the stanine score distribution for nonword reading was similar to a normal distribution, with the majority of children scoring between the stanines of 5-7 (402, 55.3%). There was also a similar pattern shown for exception word reading at Year 1 with more children than expected gaining high scores of 8-9 although there was a broader spread of exception word reading stanine scores compared to nonword reading (variance of 4.98 for exception words versus 3.52 for nonwords).

The distribution of exception word reading stanine scores at Year 1 reflected more children achieving a minimum stanine score of 1 (23, 3.16%) compared to nonword reading (5, 0.69%). Combined with more children scoring a stanine of 4 or below when reading exception words (222, 30.54%) compared to nonwords (166, 22.83%) which according to the DTWRP manual reflects a word type reading deficit (DTWRP; Forum for Research into Language and Literacy, 2012). Interestingly, both word types at Year 1 displayed a higher proportion of children achieving a maximum stanine score of 9 compared to the normal distribution curve (87, 11.97% for nonword reading and 103, 14.17% for exception word reading). Profiles of reading difficulty will be explored fully in the next chapter.

Figure 2.6

ALP Year 1 Nonword Reading Stanine Score Distribution

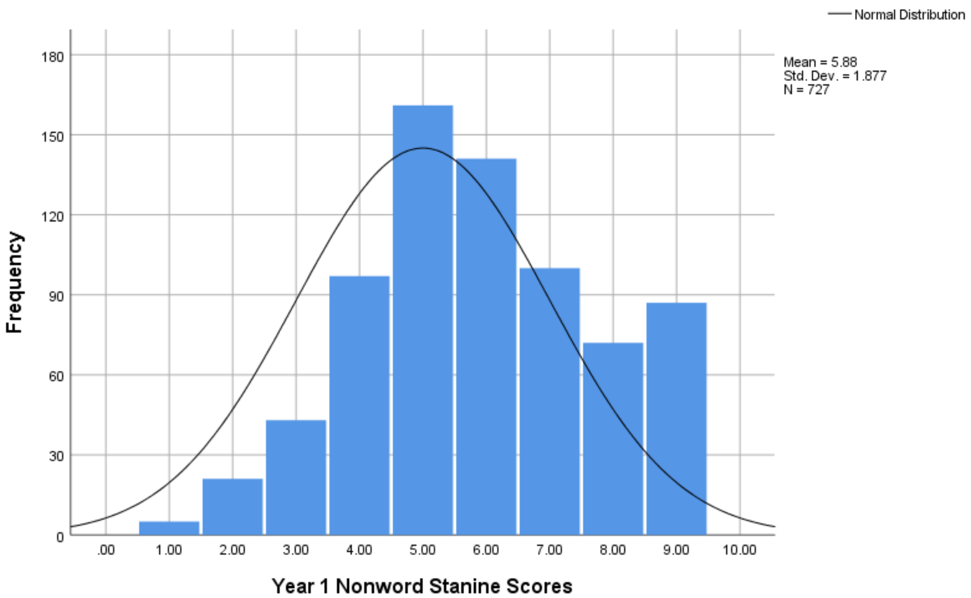
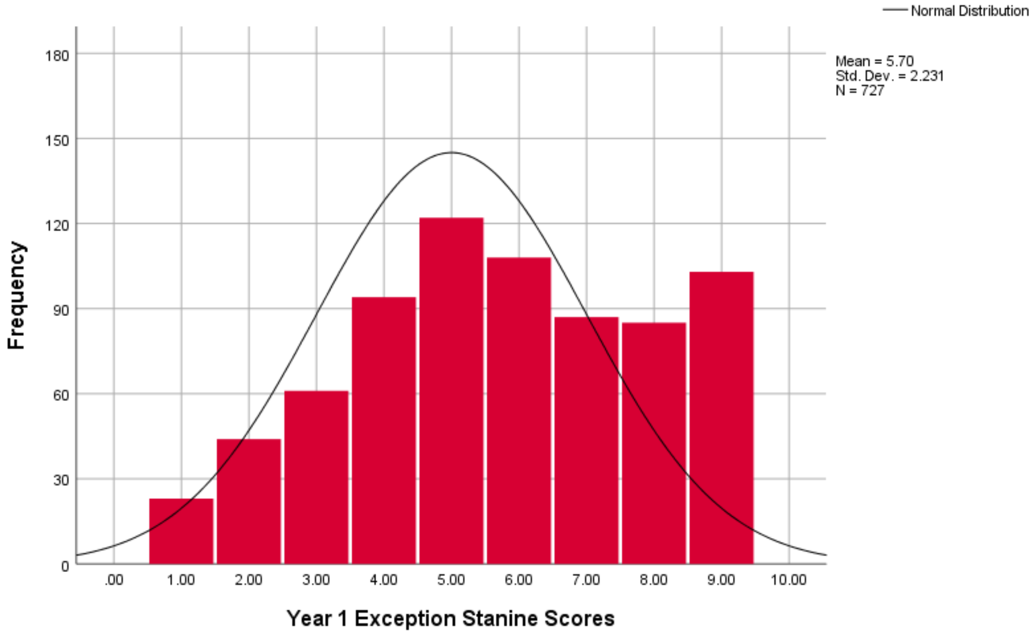


Figure 2.7

ALP Year 1 Exception Word Reading Stanine Score Distribution



Figures 2.8 and 2.9 below display nonword and exception word stanine scores at Year 4. At Year 4, both distributions for nonword and exception word reading were similar to the normal distribution curve, with the majority of children scoring between the stanines of 5-6. Both distributions also had similar levels of variance (2.95 for nonword reading and 2.91 for exception word reading) although for nonword reading, there was a slightly larger negative skew (-.24) compared to exception word reading (-.22) with a higher proportion of children scoring a stanine of 7. Additionally, more children achieved a maximum stanine score of 9 at Year 4 during nonword reading when compared to exception word reading (44, 7.6% and 21, 3.63% respectively).

Figure 2.8

ALP Year 4 Nonword Reading Stanine Score Distribution

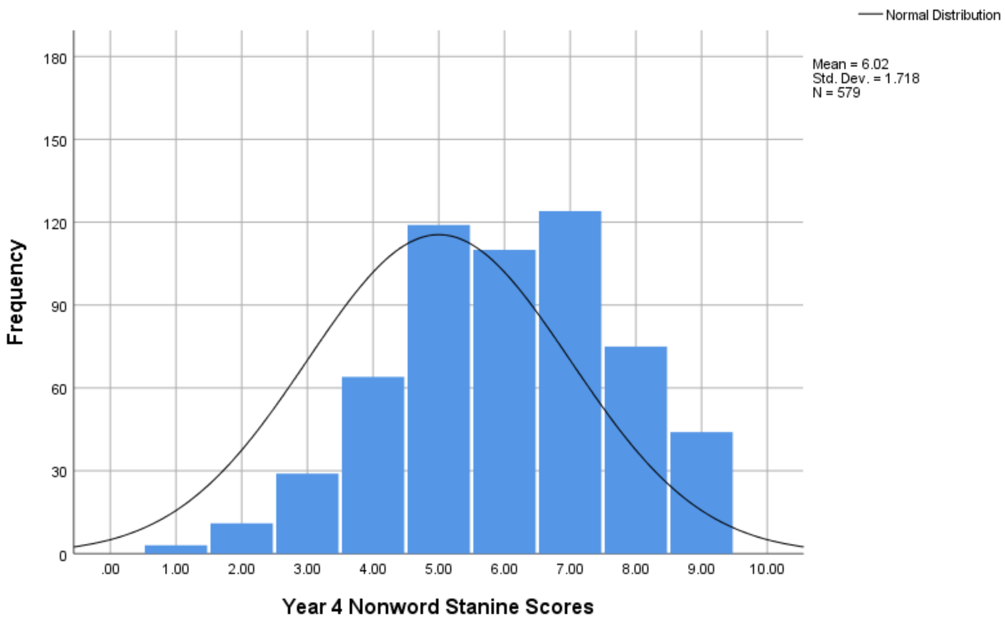
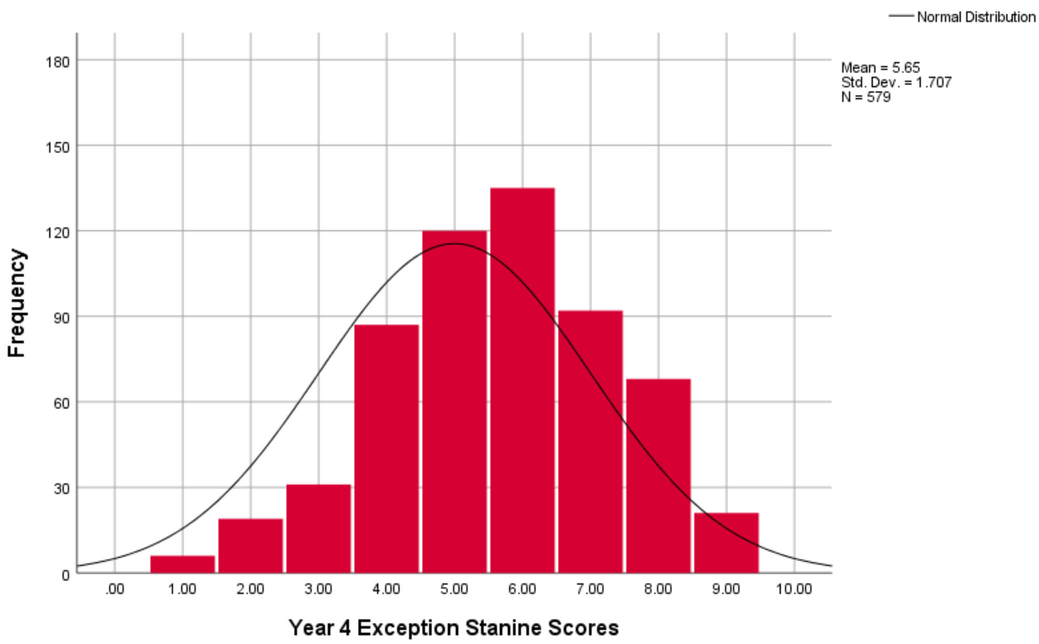


Figure 2.9

ALP Year 4 Exception Word Reading Stanine Score Distribution



2.3.4 *Question 2: Word Reading Performance between Pre and Post Phonics Samples (ALP and DTWRP Sample Data)*

In order to address the second research question and directly compare the word type reading performance of the pre-phonics (DTWRP) and post-phonics (ALP) samples, a series of one-way ANCOVAs using raw scores for each year group were conducted, followed by comparing the descriptive statistics for each sample during Year 1 and Year 4 across word types. As the raw data for the ALP sample was not normally distributed, up until this point non-parametric analyses have been used, however without a non-parametric equivalent of the ANCOVA, parametric ANCOVAs were conducted. Alternative statistical analyses were considered as part of the data analysis process and have been outlined previously as part of Section 2.3.1 within the data analysis strategy. This analysis was selected as it allowed for the statistical control of a third variable - relative age - which could have confounded reading performance within each time point. This was informed by research conducted by the Institute for Fiscal Studies (IFS) in 2007 and 2013, which stated that relative age within a year group provides an attainment advantage amongst primary school aged children (Crawford et al., 2007; 2013). Specifically, the IFS reported that summer born children are at risk of poor academic performance compared to their peers with older relative ages.

As part of this analysis, the regression slopes for the ALP and DTWRP samples were compared through scatterplots as shown in Figures 2.10 and 2.11 below. These figures indicated that the regression lines for the two samples were similar across Years 1 and 4.

Figure 2.10

ALP and DTWRP Year 1 Nonword and Exception Reading Regression Slopes

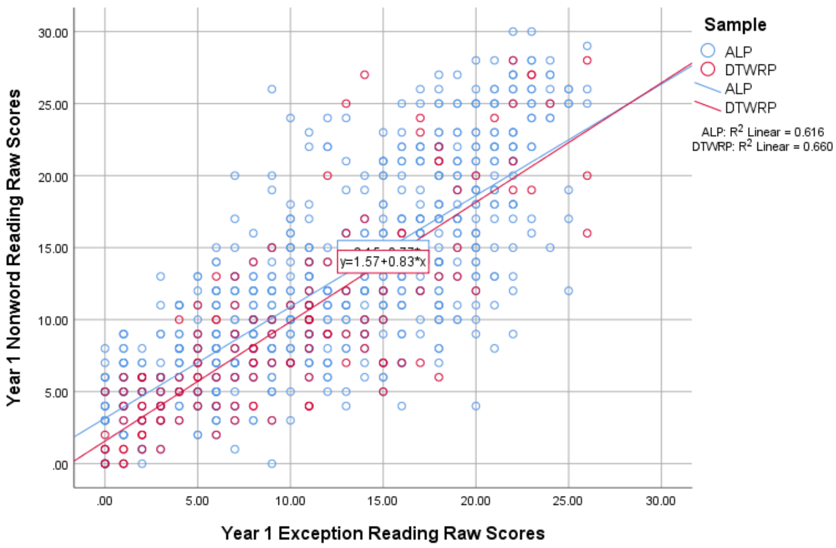
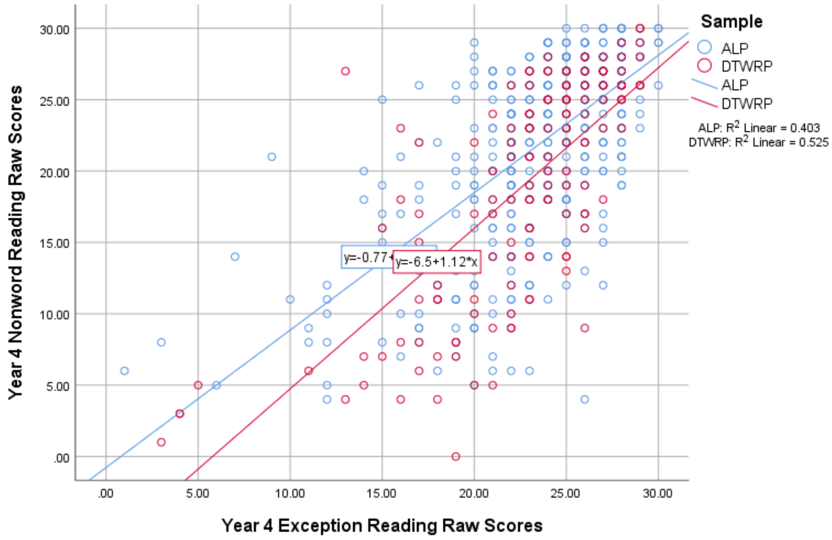


Figure 2.11

ALP and DTWRP Year 4 Nonword and Exception Reading Regression Slopes



Following this, Levene's test of homogeneity of variances was conducted to establish if both samples had similar levels of variance across Years 1 and 4. At Year 1, Levene's test indicated equal variances on nonword reading (F = 0.21, p = 0.88) and exception word reading (F = 0.46, p = 0.50). , In contrast, at Year 4, Levene's test indicated unequal variances on nonword reading (F = 32.5, p < 0.01) and exception word reading (F = 8.99, p < 0.01). Therefore, Year 4 results should be interpreted with caution as not all assumptions of the ANCOVA were met. Effect sizes are interpreted with the guidance provided by Cohen (1988).

In Year 1, the DTWRP sample consisted of 143 participants compared to 727 participants in the ALP sample. In the DTWRP sample, the mean age was 75 months, SD = 3.93, while the ALP sample also had a mean age of 75 months, SD = 3.6. Overall, the Year 1 datasets had a mean age of 75 months, SD = 3.65. The Year 1, one-way ANCOVA included raw scores for nonword and exception word reading at Year 1 from the ALP and DTWRP samples with age included as a covariate.

At Year 1, there was a significant difference in nonword reading performance (F(1,867) = 15.81, p < 0.01) between the DTWRP and ALP samples, with a small effect size of $\eta^2 = 0.018$. The covariate of age at Year 1 was also significant (F(1,867) = 39.76, p < 0.01) with a small effect size of $\eta^2 = 0.044$.

Year 1 exception word reading performance was also significantly different between the two samples (F(1,867) = 8.54, p < 0.01) with a small effect size of $\eta^2 = 0.010$. The covariate of age was also significant in exception word reading at Year 1 ($F(1,867) = 52.43$, $p < 0.01$) with a medium effect size of $\eta^2 = 0.057$.

Across both word types at Year 1, the ALP sample displayed a reading advantage compared to the DTWRP sample, with higher mean scores as follows. When reading nonwords, the ALP sample had a higher mean raw score (Mean = 12.10, SD = 6.79) versus the DTWRP sample (Mean = 9.71, SD = 6.95). When reading exception words, the ALP sample also had a higher mean raw score (Mean = 11.53, SD = 6.89) versus the DTWRP sample (Mean = 9.83, SD = 6.81).

In Year 4, the DTWRP sample consisted of 143 participants compared to 579 participants in the ALP sample. In the DTWRP sample the mean age was 111 months, SD = 3.38, while the ALP sample also had a mean age of 111 months, SD = 3.48. Overall, the Year 4 datasets had a mean age of 111 months, SD = 3.46. The Year 4, one-way ANCOVA included raw scores for nonword and exception word reading at Year 4 from the ALP and DTWRP samples, with age included as a covariate.

At Year 4, there was a significant difference in nonword reading performance ($F(1,719) = 33.92$, $p < 0.01$) between the DTWRP and ALP samples, with a small effect size of $\eta^2 = 0.045$. The covariate of age at Year 4 was not significant ($F(1,719) = 2.76$, $p = 0.097$) with an effect size of $\eta^2 = 0.004$, falling below the threshold for a small partial Eta squared effect size.

Year 4 exception word reading performance was also significantly different between the two samples ($F(1,719) = 11.96$, $p < 0.01$) with a small effect size of $\eta^2 = 0.016$. Age was not significant in exception word reading at Year 4 ($F(1,719) = 3.36$, $p = 0.067$) with an effect size of $\eta^2 = 0.005$, falling below the threshold for a small partial Eta squared effect size.

Similarly to Year 1, at Year 4 the ALP sample displayed a reading advantage compared to the DTWRP sample, with higher mean scores as follows. When reading nonwords, the ALP sample had a higher mean raw score (Mean = 22.22 SD = 5.89) versus the DTWRP sample (Mean = 18.78, SD = 7.78). When reading exception words, the ALP sample also had a higher mean raw score (Mean = 23.84, SD = 3.95) versus the DTWRP sample (Mean = 22.48, SD = 5.01).

2.4 Discussion

The findings from this study are interpreted in the context of intensive synthetic phonics teaching and that the ALP sample formed part of the first cohort to undertake the Phonics Screening Check at Year 1. The research questions for Study 1a (Chapter 2) were as follows:

Question 1. Does word type reading performance change as reading develops, following systematic, synthetic phonics instruction?

Question 2. Does word type reading performance differ between a pre-phonics and post-phonics sample?

Before discussing these research questions, it is important to note that the initial correlation analyses showed an expected pattern of correlations between standardised and non-standardised tests administered during the ALP data collection. Across both year groups, the Orthographic Choice and DTWRP exception word subtest had stronger positive correlations compared to the Phonological Choice and DTWRP nonword subtest at Year 4, although this was also a moderate, positive correlation. As these standardised and non-standardised measures were moderate to strongly correlated, this suggests that they were measuring similar underlying reading skills across the different tests. In this case, Non-Lexical processes including phonological decoding were assessed during nonword reading and Phonological Choice, while Lexical processes including existing orthographic representations were assessed during exception word reading and Orthographic Choice. Therefore, supporting both the use of the DTWRP as an assessment measure and this study's focus on the results of the DTWRP subtests.

To address the first research question, word type reading performance within the ALP sample was examined, using the two DTWRP subtests of nonword and exception word reading. ALP raw scores demonstrated a slight advantage for nonword reading at Year 1, which was significant despite using subtests from the DTWRP which were matched for difficulty. Since nonwords provide the best test of phonic skills, this advantage suggests intensive synthetic phonics teaching was having the intended effect that early readers were successfully using the Non-lexical, phonological decoding route to read nonword stimuli. Additionally, these children may have been practiced at reading nonwords through this intensive synthetic phonics teaching and preparations ahead of the Year 1 Phonics Screening Check. Stanine scores at Year 1 for both the DTWRP subtests of nonword and exception word reading demonstrated that the ALP sample were scoring slightly better than expected on both tests. Therefore, this nonword reading advantage was not at the cost of exception word reading, as the exception word reading ability of the ALP sample was not worse than expected for their age group. This finding for exception word reading performance may be caused by the limited amount of "sight word" teaching of exception words delivered through the synthetic phonics programmes. This may have corresponded with exception word items within the DTWRP, leading to successful exception word reading. However, this is only a potential explanation and could not be confirmed through a lack of information regarding the exception word items taught in each primary school.

At Year 4, interestingly this pattern of results reversed. Raw scores were higher for exception word reading compared to nonword reading, although stanine scores once again demonstrated that children in the ALP sample were ahead of age-related expectations on both word types. According to

Share's Self-Teaching Hypothesis (1995) and dual-route models of reading such as the DRC (Coltheart et al., 2001) which have been applied to reading development, it would be expected that at Year 4, children would be relying less on the Non-lexical phonological decoding pathway. Instead, orthographic representations and the Lexical pathways have developed to read exception words, "by sight". Additionally, these children would have been without phonics in the curriculum for 2+ years, so the overall teaching emphasis would have been less focused on phonological decoding. This is potentially reflected in this declining advantage for nonword reading by Year 4 but an increase in exception word reading ability. Interestingly, very few children showed nonword reading difficulties at Year 4, so this lack of focus on phonics appears justified. Reading difficulties within the ALP sample will be discussed further in the next chapter.

To address the second research question, the ALP sample was compared to the DTWRP standardisation sample, as this allowed for a direct comparison of two samples who were collected post-2012 and pre-2012: before the introduction of the 2012 Year 1 Phonics Screening Check and therefore mandatory phonics teaching. ANCOVAs directly compared the ALP sample with the performance of the DTWRP standardisation sample, which demonstrated that the ALP sample had a reading performance advantage across both Years 1 and 4 and both word types compared to the DTWRP sample. Additionally, the covariate of age was no longer significant at Year 4, while it was significant at Year 1. This indicates that by Year 4, initial age-related advantages of being months older than their peers is no longer a significant factor in reading performance across nonwords and exception words. Instead, group membership between the samples is significant at Year 4 with large effect sizes, again reflecting the advantage that the ALP sample had in reading performance across both word types.

These results support the findings of the ALP stanine distributions, suggesting an overall reading performance advantage for post-phonics-screener children who have undertaken intensive synthetic phonics as part of the primary school curriculum. Additionally, these results reflect predictions from a dual-route reading perspective regarding word type reading performance, with a nonword advantage found within the post-phonics sample at both time points compared to the pre-phonics sample. Indicating that this increased focus on systematic, synthetic phonics teaching has provided early readers with training along the Non-lexical route to successfully phonologically decode nonword items, which remains longitudinally. Interestingly, a similar result was discovered for exception word reading performance between the pre- and post-phonics samples, with an exception word reading advantage found in the post-phonics sample which remained longitudinally. Therefore, this may indicate that this increased focus on synthetic phonics teaching for early readers, indirectly improves development along the Lexical routes to reading. As independence within reading is gained through the improved phonological decoding skills which acts as a "self-teaching" mechanism according to Share (1995) and allows early readers to develop their exposure to print and create orthographic representations for reading along the Lexical route.

It is important to consider the limitations of this study, to inform the conclusions that can be made and identify issues which future research may address. Firstly, it is important to note that while these two large secondary data samples provide retrospective quasi-experimental groups which undertook the same standardised measures of word type reading, the samples were not perfectly matched. This was due to their differing sample sizes, locations sampled within the UK, differing percentages of FSM and the lack of a complete DTWRP sample, as part of the raw DTWRP data was not available for analysis. Additionally, the ALP sample allowed for longitudinal analyses of the post-phonics sample, whilst the DTWRP sample was cross-sectional and thus no longitudinal reading performance within this pre-phonics sample could be analysed. In future research, it may be of interest to compare the longitudinal reading performances of pre- and post-phonics samples in a randomised control trial within a country without existing phonics education, as this is beyond the scope of this study. There was also attrition of 157 participants (21.2%) within the longitudinal ALP sample which may have impacted results related to longitudinal reading performance.

Additionally, due to the use of secondary, retrospective data from the two selected samples of the ALP and DTWRP, it is difficult to infer causation from these results. While this study provides a step towards determining the role that systematic, synthetic phonics teaching has had on early reading development and performance, extraneous variables may have contributed to the findings, such as the home literacy environment of each group. Overall, the ALP sample has provided valuable information regarding the longitudinal reading performance of a post-phonics sample, which has produced a pattern of reading performance development across different word types within the same sample. Whilst the DTWRP sample has provided a retrospective sample of word type reading within a pre-phonics sample before the introduction of the 2012 phonics screening check, to allow for comparisons between age matched pre- and post-phonics samples. This combination of cross-sectional and longitudinal data from these respective samples supports an overall pattern within the data, suggesting that reading performance at differing time points reflects development along the dual pathways to reading, with an initial nonword reading advantage at Year 1 along the Non-lexical pathway, followed by an exception word advantage at Year 4 along the Lexical pathways. There may also be a wider impact of systematic, synthetic phonics teaching on the reading performance of the post-phonics sample, who appeared to have an overall reading advantage across both word types, when compared to the pre-phonics sample. Nevertheless, we cannot determine that systematic, synthetic phonics teaching is the only explanation for this advantage, due to the lack of control of extraneous variables.

Regarding limitations within the results, it is acknowledged that all of the assumptions for the ANCOVAs were not met, specifically the use of parametric data and the homogeneity of variances at

Year 4, therefore, the ANCOVA results should be interpreted with caution. These results indicate a pattern within the data between the two groups, which is supported by the descriptive statistics provided within Section 2.3.4. Overall, the results from this study indicate a pattern of word type reading performance differences between a pre- and post-phonics sample, with evidence that early readers with systematic, synthetic phonics teaching have a word type reading performance advantage across both nonwords and exception words, compared to pre-phonics early readers. This is combined with evidence of word type reading performance changing longitudinally within the post-phonics sample to reflect word type reading advantages that are consistent with the dual-route pathways to reading.

2.4.1 Conclusions and next steps

In conclusion, the results of this study demonstrate that children's relative ability to read each word type changes as reading develops, in a way that is consistent with dual-route theories of reading. Initially, beginner readers appear to prioritize the Non-lexical phonological decoding pathway, as indicated by their relatively better performance when reading nonwords. This also suggests that intensive synthetic phonics teaching can successfully teach early readers to decode and read aloud novel words. Whilst over time, this advantage for word types reverses, with an advantage for exception word reading over nonword reading, as developing readers rely less on the Non-lexical pathway and instead utilise growing orthographic representations and Lexical pathways to reading.

Another key finding was that pre- and post-phonics samples showed a different pattern of performance. Specifically, the post-phonics ALP sample demonstrated an overall reading advantage compared to the pre-phonics DTWRP sample. This advantage was found across both word types and at both time points of Year 1 and Year 4. It is important to note that the samples could not be directly compared as the DTWRP sample had a higher percentage of children in receipt of FSM and therefore may have sampled a slightly lower SES group. Unfortunately, the full DTWRP dataset was not available for a direct comparison, taking into account SES-related covariates. Nevertheless, the analyses reported above were able to account for the relative age of children within a year group, which is a key predictor of reading success. When age was included in the ANCOVAs, initial age-related advantages compared to peers in each sample became non-significant by Year 4 whereas the group membership of being in a pre or post phonics sample remained significant. This advantage for the ALP sample does potentially indicate that intensive synthetic phonics teaching is succeeding in teaching early readers how to read and developing important skills such as phoneme awareness and GPC knowledge. Suggesting that the introduction of mandatory intensive synthetic phonics teaching strategies have positive consequences for the reading abilities of primary school children.

Since this chapter has focused on all participating children from the ALP and DTWRP samples, the next step for this data is to focus on children with reading difficulties in early primary school. Whilst the majority of poor reader research focuses on dyslexic children, everyday mainstream teachers are faced with poor readers in their classroom who will not receive an assessment for dyslexia. Therefore, it is crucial that we examine the profiles of reading difficulty shown by all poor readers in mainstream classrooms. From this, the longitudinal stability of these reading difficulties can be examined, along with potential effects of phonics teaching on these profiles. Ultimately resulting in research which investigates how best to support these mainstream poor readers.

3 Study 1b: Reading Difficulty Profiles in Early Readers Post-Phonics Introduction

Abstract

This study investigates profiles of word reading difficulty within the post-phonics sample discussed in Chapter 2. The profiles of word reading difficulty found within this sample were also compared to the same pre-phonics sample from Chapter 2 in order to investigate the potential impact of phonics teaching on emerging profiles of reading difficulty amongst early readers. The longitudinal stability of these profiles within the post-phonics sample was also examined. Within the post-phonics sample, profiles of reading difficulty were relatively low, when compared to the pre-phonics sample and normed on the post-phonics sample data. Strikingly, the post-phonics sample rarely displayed nonword reading profiles of difficulty, corroborating the findings in Chapter 2, with regards to phonics instruction strengthening nonword reading and the Non-lexical pathway amongst early readers. The results also reflect a lack of stability in profiles of nonword reading difficulty within the post-phonics sample, while profiles of exception word and mixed difficulty displayed fair to moderate stability throughout primary school. It is suggested while systematic synthetic phonics instruction during early primary school has potentially reduced early nonword reading difficulties, exception word and mixed profiles of difficulty may not be addressed through this phonics instruction and may reflect a wider issue with the detection and support available for overall poor readers. It is important to note that there is no evidence to suggest that phonics instruction has resulted in an increase in exception word reading difficulties. In fact, the pattern suggests a reduction in reading difficulties overall.

3.1 Introduction and Rationale

3.1.1 Reading Difficulties along Dual-Routes to Reading

Sections 1.4-1.4.1 discussed reading difficulties from a dual-route reading perspective, with much of the existing literature and knowledge regarding profiles of reading difficulty examining reading difficulties within dyslexia subtypes (Castles & Coltheart, 1993; Manis et al., 1996; McArthur et al., 2013; Romani et al., 2008; Stanovich et al., 1997; Wolff, 2009; Ziegler et al., 2008). Whilst the focus of this thesis is to examine word reading and word reading difficulties within early readers, without dyslexia classifications as a wider representation of early readers within mainstream classrooms, dyslexia literature makes important contributions to the topic of childhood reading difficulties. Additionally, there is limited information available regarding the reading difficulties displayed by early readers without a dyslexia classification, which this study seeks to address.

Nevertheless, theoretical accounts of reading difficulties derived from research into subtypes of deficit within dyslexia and wider related skills such as orthographic learning and phonological processing, can be applied to early readers without a dyslexia classification such as Wang et al. (2014), when they are consistent with a reading framework such as the dual-route reading perspective which accounts for typical reading development. As such, it is important to revisit how reading difficulties across different word types may manifest, according to a dual-route reading approach. According to the dual-route reading perspective, nonwords are read exclusively through the Non-lexical route, through the process of phonological decoding. Therefore, if an early reader has a nonword type reading difficulty, this may indicate a problem within the Non-lexical pathway, which could include phonological processing, a lack of GPC knowledge to apply to phonological decoding or a lack of phoneme awareness to conduct phonological decoding for example. According to Share's Self-Teaching Hypothesis (1995), these phonological decoding difficulties may then have wider implications for reading, whereby early readers cannot decode unfamiliar words, ultimately leading to a lack of independent access to text and difficulties developing orthographic knowledge. This reading difficulty along the Non-lexical pathway manifests similarly to phonological dyslexia, which includes word type reading difficulties for nonwords and novel regular words; both of which require phonological decoding (McArthur et al., 2013).

In contrast, according to the dual-route reading perspective, exception words are read mainly through the Lexical pathway as whole-words, as to read them through phonological decoding alone would result in regularization errors (Coltheart, 2005). Therefore, if an early reader displays an exception word type reading difficulty, it may indicate a problem within the Lexical pathway, such as a difficulty with orthographic processing, or a lack of independent reading to create the orthographic representations necessary for whole-word reading along the Lexical pathway. Therefore, early readers with an exception word reading difficulty can manifest similarly to surface dyslexia, if these readers are able to read nonwords through phonological decoding, but incorrectly read exception words through the same sublexical approach (McArthur et al., 2013).

As previously discussed in Section 1.4.1, the origin of exception word reading difficulties is inconclusive, with multiple theoretical explanations including: a lack of print exposure which hinders orthographic knowledge growth, difficulties within orthographic processing and the creation of orthographic representations and poor vocabulary knowledge which results in a lack of semantic representations and difficulties within the indirect Lexical-semantic route. Equally, this word type reading difficulty may have a developmental explanation, as the Lexical route to reading is not expected to develop until after the Non-lexical route has been established, providing early readers with phonological decoding skills to begin independent reading.

There is also a third type of reading difficulty to consider, which will be referred to as a "mixed" profile of reading difficulty. Within this profile of reading difficulty, readers demonstrate difficulties when reading both nonword and exception word items and instead display poor reading performance overall, rather than within one specific word type. This mixed form of reading difficulty may therefore indicate difficulties along both the Non-lexical and Lexical pathways to reading. As discussed in dyslexia literature with the concept of "pure" and "relative" difficulties (Peterson et al., 2013), the nonword and exception profiles of reading difficulty may reflect "pure" difficulties along one pathway to reading, whilst this mixed profile display "relative" difficulties along both pathways.

According to the dual-route perspective of reading, one possibility is that there are potential difficulties within the Non-lexical pathway and phonological decoding in this mixed profile of reading difficulty, which hinders reading of nonword items and also hinders development of the Lexical pathway to reading. Alternatively, this difficulty along the Non-lexical pathway could also be combined with one of the potential explanations for exception word reading difficulties to display this "relative" difficulty in both pathways.

What is not currently known is whether current systematic, synthetic phonics teaching is effective in addressing all three profiles of reading difficulty, amongst early readers within mainstream classrooms. Evidence from the dyslexia literature which discusses interventions for phonological and surface dyslexic subgroups, which as discussed above, share similarities with nonword and exception word reading difficulties, often suggests a tailored intervention approach which synthetic phonics alone may not be able to provide (Gustafson et al., 2007; Fiorello et al., 2006; O'Brien et al., 2012; Rose, 2009). Some research indicates that overall poor readers such as those in the mixed profile of reading difficulty, benefit from a combination of phonics and sight word training, to address difficulties along both the Non-lexical and Lexical pathways (McArthur et al., 2015). While at this time it cannot be determined whether mainstream synthetic phonics teaching is able to address all reading difficulties, there is research to suggest that supplementing a phonics programme with exception word teaching is not detrimental to reading outcomes (Shapiro & Solity, 2016).

3.1.2 Reading Difficulties and Phonics Teaching

Since the introduction of the Phonics Screening Check in 2012 and the following increased focus on systematic, synthetic phonics teaching in England, types of emerging reading difficulties amongst early readers in primary school have not been examined. Therefore, it has not been determined whether this increased synthetic phonics teaching has impacted profiles of reading difficulty displayed in mainstream classrooms of early readers. Furthermore, the stability of these mainstream classroom profiles of reading difficulty have not been examined. This is of specific research interest in the case of nonword profiles of reading difficulty, due to the impact that synthetic phonics teaching may have on addressing this reading difficulty over time without additional intervention. Additionally, examining the longitudinal stability of exception word profiles of reading difficulty is of research interest, to examine whether this word type reading difficulty is also improved over time due to a “delay” in developing the Lexical route to reading, or if this reading difficulty requires intervention.

In Section 2.1.2 within the previous study, predictions regarding word type reading following systematic, synthetic phonics instruction were discussed. Similarly to the previous study, the discussion of emerging reading difficulties amongst early readers, following synthetic phonics teaching, are also viewed from a dual-route reading perspective and therefore, share similar predictions. That is, if systematic, synthetic phonics teaching is successful, then the Non-lexical pathway of early readers will be trained, resulting in successful phonological decoding and nonword reading and fewer profiles of nonword reading difficulty.

However, one concern that has been raised regarding phonics teaching, is that children are trained to rely on phonological decoding instead of reading for meaning (Rosen, 2021). Whereby, early readers may display less profiles of nonword reading difficulty after receiving synthetic phonics teaching, but this is not expected to improve the Lexical route to reading exception words directly (however improvement could occur indirectly as improved phonological decoding ability facilitates orthographic learning and representations for the Lexical route). In fact, the Phonics Screening Check only assesses children’s phonic-decoding because the nonwords and regular words used in the test are explicitly selected such that children can read them using the GPCs they have already been taught. If this concern is valid, then current synthetic phonics teaching combined with inconsistent exception word teaching within synthetic phonics, may not be adequately addressing exception word reading difficulties and as a result, there may be an increased amount of this profile of reading difficulty. Additionally, if the origin of exception word reading difficulties is related to difficulties outside of phonological decoding, such as within orthographic processing or a lack of exposure to text, these difficulties are not expected to be addressed through synthetic phonics teaching. Ultimately, further research is required to determine the underlying difficulties associated with exception word reading difficulties.

In terms of mixed profiles of difficulty, predictions from a dual-route perspective of reading are undetermined. One possibility is that systematic, synthetic phonics teaching improves the Non-lexical pathway to reading as discussed above, which results in improved nonword reading for early readers with a mixed profile of reading difficulty. As a result, this “relative” difficulty across both pathways to reading instead changes into a “pure” difficulty, closer representing a profile of exception word reading difficulty, as synthetic phonics is not expected to directly improve this Lexical pathway. Alternatively, this profile of difficulty may represent overall “poor readers” who may not improve their nonword reading difficulty due to synthetic phonics instruction alone and instead may require intensive teaching or tailored intervention.

Therefore, according to predictions from the dual-route reading perspective, the increased focus on systematic, synthetic phonics teaching may have positive outcomes for profiles of reading difficulty amongst early readers. However, this improvement may be limited and a lack of direct exception word teaching which is inconsistent across mainstream classrooms delivering phonics teaching, may increase this profile of reading difficulty or not improve this profile of difficulty beyond what is currently expected for early readers.

The next step is to determine the stability of these profiles of reading difficulty, as little is known regarding the stability of these word reading difficulties outside of the dyslexia literature. Additionally, there is a lack of information regarding the stability of these word reading difficulties following systematic, synthetic phonics instruction, which may have implications for how poor readers are supported in mainstream classrooms.

3.1.3 Stability of Word Reading Difficulties

The second aim of this study was to determine if these profiles of reading difficulty are stable throughout primary school. Little is known regarding the stability of these word reading difficulties within mainstream classrooms, with much of the research discussing stability from within dyslexia subtypes. Findings from research with dyslexic participants indicates that nonword reading difficulties (phonological dyslexia) represent a deviant development with more stability than exception word reading difficulties (surface dyslexia) which improve with time, as this instead represents a developmental delay with limited stability (Peterson et al., 2014).

Peterson et al. (2014) investigated the longitudinal stability of different subtypes of dyslexia which they identified in an earlier study (Peterson et al., 2013). In the 2013 study, Peterson et al. identified 437 children aged eight to thirteen-years-old, who were categorised into a phonological, surface or mixed subtype of dyslexia. In this study, “pure” (either lexical or sublexical difficulties alone) and “relative” (difficulties in both pathways, but one form of reading is worse than the other) subtypes were used. Participants were assessed on a battery of tests, including phonological and orthographic coding, general intelligence, literacy and reading related skills (e.g., phoneme deletion). 72 of these dyslexic children returned for the follow up study five years later (Peterson et al., 2014) and the same subtyping criteria was used to reclassify them. Cohen’s Kappa was used to determine the longitudinal stability of these profiles, with results indicating “fair” stability, with stronger stability for phonological dyslexia than surface dyslexia (Peterson et al., 2014). When analysing the cognitive profiles of each subtype, the phonological subtype had poorer phonological awareness than surface and mixed subtypes, which was consistent across both studies (Peterson et al., 2014).

These results can be interpreted to show that phonological dyslexia is a more stable deficit, potentially linked with a developmental deviance, as shown by a large deficit in phonological awareness longitudinally. Whereas surface dyslexia was not as stable, which may indicate that this subtype is linked to a developmental delay, where across these five years, some children moved out of the surface subtype due to improvement in their orthographic knowledge and representations, as explained by the dual-route reading perspective, as the Lexical route develops with reading experience. Although the results from Peterson et al. (2014) further our understanding of dyslexic subtype stability, they did not elaborate on how pure subtypes compared to relative subtypes. They attributed this to attrition rates, whereby the final sample of pure subtypes was small, which made it difficult to compare to relative subtypes (Peterson et al., 2014).

However, as discussed earlier in this introduction, profiles of nonword, exception and mixed word reading difficulties share some similarities with phonological, surface and mixed dyslexia subtypes, namely on word type performance and potential difficulties along the two pathways to reading, from a dual-route reading perspective. Therefore, these stability findings may also apply to early readers within mainstream classrooms who do not have dyslexia classifications, which has implications for how these profiles of reading difficulty should be addressed to support poor readers. In terms of the developmental trajectories of poor readers, similar findings to dyslexia research have also been uncovered amongst children without a dyslexia classification. Talcott et al. (2013) found that children with impaired exception word reading abilities displayed a trajectory of developmental delay, compared to children with impaired nonword reading, who displayed a trajectory of deviant development.

Additionally, research conducted with poor readers without dyslexia classifications indicate that phonological difficulties demonstrate greater longitudinal stability than orthographic difficulties (Spector, 2005; Steacy et al., 2014). With mixed profiles of reading difficulty amongst poor readers (sometimes referred to as a “double-deficit” subtype which demonstrates difficulties in both phonological tasks and orthographic rapid naming tasks), displaying greater longitudinal stability (Steacy et al., 2014). In their longitudinal study from kindergarten to second grade, Steacy et al. (2014) found that reading difficulties were more stable during the first grade, with all groups demonstrating a high probability of remaining in that same group. In second grade, stability remained fair, although the group with orthographic difficulties (rapid naming deficits) had a 32% probability of moving into a profile of no difficulty as their Lexical route to reading developed (Steacy et al., 2014). Spector (2005) also found that approximately 80% of participants who were classified into a reading difficulty profile in the autumn, also had a difficulty profile in the spring of first grade, with only 30% of participants with an orthographic difficulty classified with this reading difficulty at both time points. Nevertheless, these results must be interpreted with caution, as stability was only measured across the span of a five months, so perhaps this is not enough to measure deviant development. Rather, as these were beginning readers in the first grade, perhaps this explores the skills of emergent readers. In this case, those children who cannot improve their phonological awareness skills show a stable difficulty, as these skills are required for phonological recoding, the first approach in learning to read. Those children who are experiencing orthographic rapid naming difficulties may already have the tools of phonological recoding but require exposure to develop orthographic representations. Therefore, their difficulty is prone to improvement without teaching instruction, while phonological recoding cannot occur without learning GPCs.

Additionally, some children move from one subtype of reading difficulty to another, specifically in the work of Steacy et al. (2014) whereby a proportion of the sample moved from a single difficulty subtype to the mixed "double-deficit" subtype, as they failed to improve their phonological awareness and rapid automatic naming skills to the level of their peers, which was classified as a second reading difficulty (Steacy et al., 2014). While this may complicate the process of improving a reading difficulty, it does demonstrate that the two reading pathways (Non-lexical and Lexical) are connected to each other within reading development, as proposed by the dual-route reading perspective, and that a difficulty within one pathway may impact the alternative pathway and have consequences for reading performance. For example, if an early reader has poor phonological awareness so that they cannot phonologically decode words, this may cause a difficultly in creating orthographic representations with distinct letter order information.

As this thesis investigates emerging reading difficulties within a wider sample of young children who have received systematic, synthetic phonics instruction and who are not selected on the basis of a dyslexia assessment, it is currently unknown whether these difficulties will be stable over time due to the potential impact of increased phonics teaching. This is an important question because if early reading difficulties are stable, this would indicate a potentially long-term problem that must be proactively addressed to remove this barrier to a child's educational potential. Stable difficulties therefore motivate early intervention through tailored interventions and teaching approaches. Furthermore, this may indicate that current synthetic phonics teaching is not successful if there are stable nonword reading difficulties. Alternatively, this may indicate that phonics teaching may not address all types of reading difficulty, specifically those along the Lexical route to reading if there are stable exception word reading difficulties.

In contrast, if early profiles of reading difficulty are not stable, and children who have reading difficulties are likely to recover, then this would mean that although the Phonics Screening Check may be useful for indicating who needs immediate support during that particular period, it would not be useful for informing the longer-term strategy for addressing children at risk of dyslexia or other learning difficulties. Additionally, if nonword reading difficulties are reduced and have little longitudinal stability, then synthetic phonics teaching may have positive outcomes for emerging reading difficulties along the Non-lexical route.

Therefore, further research is required into mainstream reading difficulties in a wide range of poor readers. Specifically, are different profiles of reading difficulty apparent in poor readers, mapping onto a similar conceptualisation of dyslexic subtypes, for example, do poor readers show relative difficulty on different word types (e.g., nonwords versus exception words; as these word types recruit different reading pathways, [Non-lexical versus Lexical]). It is these mainstream profiles of relative word reading difficulty following systematic, synthetic phonics instruction which is the interest of this study and where there is a significant gap in the literature. Through investigating these mainstream profiles of reading difficulty, both the stability of these profiles throughout primary school and how these early readers are supported within schools can then be addressed.

In summary, there are three main questions that arise from the dual-route reading perspective regarding emerging profiles of reading difficulty amongst early readers and the potential impact of the increased focus on systematic, synthetic phonics teaching. Firstly, what profiles of reading difficulty emerge within a sample of early readers after receiving systematic, synthetic phonics instruction? Secondly, do the profiles of reading difficulty found amongst a sample of post-phonics early readers, differ from a sample of pre-phonics early readers, due to the impact of systematic, synthetic phonics teaching? Thirdly, what is the stability of these profiles of reading difficulty amongst early readers? In order to determine whether these profiles of reading difficulty require tailored intervention or whether they represent delays in reading development which improve over time. This study sought to address these gaps in the literature, through investigating profiles of reading difficulty in a longitudinal study of children receiving phonics intensive teaching and their stability over different years of primary education.

3.1.4 Research Questions

Due to the shift towards systematic, synthetic phonics instruction within mainstream classrooms since the Phonics Screening Check was introduced in England in 2012, it was hypothesised that there would be fewer profiles of nonword reading difficulty within a sample of early readers at Year 1 who had received synthetic phonics teaching. According to the dual-route reading perspective, this would occur as intensive phonics teaching approaches delivered within the first two years of primary school would adequately train this Non-lexical pathway of reading.

Alternatively, the increased focus on systematic, synthetic phonics instruction may be teaching early readers to rely on phonological decoding along the Non-lexical pathway instead of reading exception words through a whole-word approach. When combined with the inconsistent manner in which exception words are currently taught in mainstream English classrooms due to a lack of government criteria, it was hypothesised that there would be more profiles of exception word reading difficulty within a sample of early readers who had received synthetic phonics teaching, as this orthographic route to reading was not being adequately addressed through phonics teaching.

No predictions were made regarding the profile of mixed reading difficulties and instead was approached as an exploratory investigation. Theoretically from a dual-route perspective, these profiles of difficulty could both decrease due to improvements in nonword reading along the Non-Lexical route due to phonics instruction, or increase due to a lack of exception word teaching and direct improvement along the Lexical route following phonics instruction. Equally as indicated by Steacy et al. (2014) this profile of reading difficulty may indicate that these are overall poor readers who have difficulties along both routes to reading, or readers who already had a relative word type reading difficulty, who then fall behind their peers when reading the alternative word type, which manifests as a mixed profile of reading difficulty.

Regarding the stability of these profiles of reading difficulty, informed by the findings from both literature with dyslexic readers and mainstream poor readers, it was hypothesised that profiles of reading difficulty would display varying longitudinal stability. Specifically, it was hypothesised that profiles of nonword reading difficulties would demonstrate longitudinal stability, as indicated by the Talcott et al. (2013), this type of reading difficulty may indicate a deviant developmental trajectory. Equally, if systematic, synthetic phonics is not improving nonword reading difficulties so that profiles of reading difficulty occur amongst early readers, then this may indicate difficulties within the Non-lexical pathway which are not being addressed through phonics teaching and persist over time. It was also hypothesised that profiles of exception word reading difficulties would lack longitudinal stability as indicated by Talcott et al. (2013), as this reading difficulty may be due to a developmental delay which improves throughout primary school as the Lexical route to reading develops.

The three research questions addressed in this study are as follows:

Question 1: What is the prevalence of each word type reading difficulty profiles within the ALP post-phonics sample?

Question 2: How does the prevalence of each word type reading difficulty profile in the ALP sample compare to a pre-phonics sample?

Question 3: How stable are these profiles of early reading difficulty?

3.2 Method

This chapter further discusses two samples of secondary data which were previously investigated in Chapter 2: the pre-phonics DTWRP sample and the post-phonics ALP sample. Firstly, reading difficulties within the ALP post-phonics sample were examined, to determine the prevalence of word type reading difficulties following systematic, synthetic phonics teaching. This utilised two approaches, the first utilised a standardised classification system and the second utilised a classification system which was standardised on the performance of the post-phonics sample to examine reading difficulty profiles with greater detail. Secondly, reading difficulty profiles within the pre- and post-phonics samples were compared using a standardised classification system, to investigate how profiles of reading difficulty amongst early readers within mainstream classrooms (without dyslexia classifications), had changed since the increased focus on systematic, synthetic phonics teaching, following the introduction of the 2012 Phonics Screening Check. Additionally, reading difficulty stability within the ALP post-phonics sample was investigated, to determine the longitudinal stability of these reading difficulties and whether these word reading difficulties improve with time and reading development (representing a reading development delay), or whether these profiles required tailored intervention or teaching approaches due to persisting difficulties.

3.2.1 Participants

This chapter utilised the ALP and DTWRP standardisation samples which were previously outlined in Sections 2.2.1.1 and 2.2.1.2. Specifically, the ALP sample at the time points of Year 1 (727 participants) and Year 4 were analysed (579 participants). Meanwhile the DTWRP standardisation sample included participants from the Reception to Year 7 age groups, for a total of 1125 participants.

3.2.2 Power Calculations

In contrast to Chapter 2, as this study focused on the frequency and types of reading difficulty profiles within the pre- and post-phonics ALP and DTWRP samples, no participants were excluded from the analyses within this chapter. As these samples consisted of secondary, retrospective data, there was also no possibility of recruiting additional participants. Therefore, the full samples described above were utilised and a priori power calculations were not conducted. The majority of this empirical chapter focuses on descriptive statistics and percentages, rather than inferential tests, often due to the small sample sizes within each profile of reading difficulty.

3.2.3 Design

Similarly to Chapter 2, the design for this study varied between the three research questions and the two samples of secondary data analysed (the ALP sample and the DTWRP standardisation sample). In regards to the first research question which investigated word type reading difficulties within the ALP sample, the analyses for this research question were within-sample. Focusing on the word type reading difficulties present within the sample through descriptive statistics, both when discussing a standardised classification system and when a within-sample classification system was used to allocate profiles of difficulty. Further descriptive statistics of non-standardised checks were also utilised as a validity check for the standardised measure, examining task performance of participants with a profile of reading difficulty within the sample. In combination with a quasi-experimental analysis of reading difficulty profile risk within the sample.

The second research question investigated if these profiles of reading difficulty vary between a pre-phonics sample and a post-phonics sample, through a between-samples analysis. This comparison of descriptive statistics between the two samples allows for a direct comparison of reading difficulty profile frequencies, when utilising profiles of reading difficulty allocated through the same classification system provided by the DTWRP assessment package.

The third research question investigated the longitudinal stability of these reading difficulty profiles from Year 1 to Year 4 in the ALP sample, using within-sample analyses. These analyses presented descriptive statistics to report frequencies of reading difficulty profiles longitudinally and movements between profile groups within the sample. In combination with quasi-experimental analyses of longitudinal profile agreement and analyses of variance within the sample.

3.2.4 Measures

As discussed in Chapter 2, the ALP sample completed the three subtests of the DTWRP assessment package in both Year 1 and Year 4, which included tests of nonword reading, exception word reading and regular word reading over a total of 90 items. While raw scores out of 30 for each test are utilised as an accuracy measure, the stanine scores for each test can be used to allocate reading difficulty profiles to primary school children. As outlined in the DTWRP manual, profiles of reading difficulty are allocated to children who fall below a stanine score of 4 on one word measure (e.g., nonwords; indicating a difficulty when reading this word type) when considered with a stanine score of 4 or above on the alternative reading measures (DTWRP; Forum for Research into Language and Literacy, 2012). Alternatively, if a child has a stanine score below 4 on the nonword reading measure and a stanine score below 4 on the exception word reading measure, they would be classified as showing a mixed reading difficulty profile, indicating difficulties when reading both word types (DTWRP; Forum for Research into Language and Literacy, 2012).

In total, there are three profiles of reading difficulty, the first is a phonological profile, which is associated with nonword reading difficulties and will hereafter be referred to as “Nonword” profiles. The second profile is a Lexical-Semantic profile, which is associated with exception word reading difficulties and will hereafter be referred to as “Exception” word profiles. Lastly, there are mixed profiles which are associated with difficulties when reading both nonwords and exception words, which will hereafter be referred to as “Mixed” profiles. As discussed in Chapter 2, reading performance on the DTWRP subtests and various word types reflects underlying reading processes associated with each word type. This is also applied to the DTWRP profiles of reading difficulty, whereby a Nonword profile may reflect difficulties in developing phonological processes such as phonological decoding along the Non-lexical route to reading (DTWRP; Forum for Research into Language and Literacy, 2012). Alternatively, an Exception profile may reflect difficulties in developing lexical-semantic processes, such as orthographic representations along the Lexical routes to reading (DTWRP; Forum for Research into Language and Literacy, 2012). Whilst a Mixed profile may reflect difficulties in both sets of phonological and orthographic processes, along both routes to reading in a dual-route model (DTWRP; Forum for Research into Language and Literacy, 2012).

3.2.4.1 Validity Check for Nonword and Exception Profiles within the ALP Sample determined through DTWRP classifications

As discussed earlier in section 2.2.3.3, the non-standardised measures of Orthographic and Phonological Choice administered within ALP at Years 1 and 4, have been used as a validity check for the DTWRP assessment package. This validity check was also applied to children within the ALP who were classed as having a DTWRP profile of reading difficulty. As hypothesised in section 2.2.3.3, the Orthographic Choice measure assessed the same underlying orthographic processes and Lexical reading routes as those measured by the DTWRP exception word reading measure. While the Phonological Choice measure assessed the same underlying phonological processes and Non-lexical reading route as those measured by the DTWRP nonword reading measure. Similarly, to Chapter 2, this validity check was conducted to establish that the DTWRP subtests and the Orthographic and Phonological Choice tasks were assessing the same underlying reading processes. Specifically examining whether the expected difficulties in these underlying processes would be assessed through both measures, that is, if a child has a difficulty when reading nonwords and therefore is assigned a Nonword profile of difficulty within the DTWRP classifications, there may be underlying difficulties in utilising the Non-lexical route to reading and phonological decoding. This would be reflected in both poor DTWRP nonword reading scores and poor Phonological Choice scores.

If this is true, then it would be expected that children classified as having an Exception profile of difficulty through the DTWRP would score lower on the Orthographic Choice than Phonological Choice measure, as their underlying orthographic processes and Lexical routes to reading may be impaired. Whilst the opposite would be found for children classified as having a Nonword profile of difficulty through the DTWRP, who would score lower on the Phonological Choice than Orthographic Choice measure, as their underlying phonological processes and Non-lexical route to reading may be impaired. Alternatively, children with Mixed profiles of difficulty may perform at similar levels for both tasks, as they are hypothesised as sharing a mixture of difficulties across these underlying orthographic and phonological processes and dual-routes to reading.

Due to a small sample size within each group, this validity check consisted of a descriptive analysis of choice scores because there was not sufficient power for inferential tests, also this validity check is only reported for Year 4 children due to floor performance at Year 1. As part of this descriptive statistic comparison, A-Prime scores were used which were generated from raw scores for the choice tests. A-Prime scores were selected as a variant of D-Prime, which adjusts for bias in participant responses such as false alarms or chance performance when there are only two options for participants to choose from. A-Prime provides an advantage compared to D-Prime, as a non-parametric index of sensitivity to the phonological and orthographic choice information. The median score from each choice test was also used to represent the spread of the data from each reading profile group within the ALP sample, as data was non-parametrically distributed on both choice tests. Median scores for children classified with each profile of difficulty, across both choice measures at Year 4 were examined. These scores are displayed in Table 6 below.

Table 4

Difficulty Profile Median Scores on Orthographic and Phonological Choice at Year 4 (ALP Data)

	Orthographic Choice (Median)	**Phonological Choice (Median)**
Nonword Profiles	0.69	0.81
Exception Profiles	0.62	0.70
Mixed Profiles	0.62	0.58
No Profile	0.86	0.87

When examining the scores of children with DTWRP reading difficulty profiles, the median raw choice task scores displayed above reflected the difficulties associated with the profile of difficulty for Exception and Mixed profiles only. Children with an Exception profile performed worse in the Orthographic Choice test (0.62) than the Phonological Choice test (0.7) whilst children with Mixed profiles demonstrated similar scores across both measures, although there was a slight advantage for Orthographic Choice (0.62) than Phonological Choice (0.58). The Mixed profile group did perform at the same level or worse than the other profile groups on both measures, displaying overall difficulties with both word types.

Interestingly, children with a Nonword profile of difficulty unexpectedly performed worse in the Orthographic Choice measure (0.69) than the Phonological Choice measure (0.81). However, it is important to note that this Nonword profile group was a smaller sample of 8 children, when compared to the other profile groups. Therefore, this choice test difference may not be reliable. The wider pattern of results for the Nonword profile group demonstrate that this group outperformed the other two difficulty profile groups across both the Orthographic and Phonological Choice tasks. As expected, children with no DTWRP profile of difficulty outperformed all three difficulty profile groups across both tasks, with similar scores for both measures.

3.2.5 Procedure

The procedure for this study has been outlined previously in Section 2.2.4, as this study utilises the same secondary data which was analysed in Chapter 2. Building upon the existing procedure, profiles of reading difficulty were allocated using stanine scores for the nonword and exception word reading subtests of the DTWRP as outlined in Section 3.2.3 above. Profiles of reading difficulty for the DTWRP sample were allocated as part of their standardisation data collection. Whilst profiles of reading difficulty for the ALP sample were allocated as part of the analysis for this thesis chapter, following the guidance provided in the DTWRP manual (DTWRP; Forum for Research into Language and Literacy, 2012).

3.3 Results

3.3.1 Data Analysis Strategy and Exploration of Data

As in Section 2.3.1, the analyses for this chapter are separated by the research question addressed. The first research question asked: what is the prevalence of each word type reading difficulty profiles within the ALP post-phonics sample? This question was addressed through two separate profile classification systems. Firstly, profiles of reading difficulty were allocated through the guidelines provided by the DTWRP assessment package. The frequencies and percentages of these DTWRP reading difficulty profiles at Year 1 and Year 4 within the ALP sample were then reported. McNemar tests were conducted for both the Year 1 and Year 4 data, to establish if there was a statistically significant risk of displaying a Nonword or Exception word profile of reading difficulty at either time point within the ALP sample. The McNemar test, which is also known as a paired chi-square test, is a non-parametric test for paired nominal data. This test investigates the homogeneity of two dichotomous variables, the first dichotomous dependent variable being displaying a reading difficulty (yes or no) across the DTWRP assessment package and the second independent variable being the two subtests of the DTWRP (nonword reading or exception word reading). This test establishes if there is a statistically significant difference between two related groups, in this case, those who are classified with an Exception profile and a Nonword profile of reading difficulty, due to their reading performance for each word type – nonwords and exception words, within the DTWRP assessment package.

To address the second research question, which asked if profiles of reading difficulty differed between a pre-phonics sample and a post-phonics sample, the profiles of difficulty allocated through the DTWRP to the ALP sample were compared with the reported percentages of the same profiles found within the DTWRP standardisation sample. This allowed for a direct comparison between a pre-phonics and post-phonics sample, utilising the same profile classification criteria to establish if profiles of difficulty had changed since the introduction of phonics teaching to the UK curriculum.

In an alternative approach to the first research question, a second reading difficulty profile classification system was used. This alternative approach sought to provide a more sensitive estimate of the relative reading difficulties within the ALP sample, in combination with acting as a validity check for the overall pattern of profiles found using DTWRP classification criteria. In this approach, raw scores from the DTWRP nonword and exception word subtests were converted into z-scores, to allow for a within-sample classification system. Informed by the research of Peterson et al. (2013; 2014) profiles of reading difficulty were allocated to participants whose z-scores fell below -1.5 standard deviation of the group mean. If z-scores fell below -1.5 standard deviation for both measures, the Mixed profile was allocated, while Nonword and Exception profiles were allocated if the respective measure fell below -1.5 standard deviation, while the alternative measure fell within 1 standard deviation of the group mean. Once profiles of reading difficulty were allocated to the ALP sample using this within-sample classification, frequencies and percentages of these profiles were reported.

To address the third research question, which asked how stable these profiles of reading difficulty were, analyses were separated by the profile classification system used. Firstly, utilising the DTWRP profile classification system, participant profiles across Year 1 and Year 4 of the ALP sample were reported, focusing on participants who were classified with the same profile longitudinally and those who moved from one profile to another. This was then followed by a series of Cohen's Kappa analyses, to assess profile classification agreement longitudinally, as informed by the dyslexia stability analyses of Peterson et al. (2014). Cohen's Kappa was used as a measure of agreement, typically this is between two independent raters, but within this analysis, there was only one rater (the DTWRP assessment guidelines) at two time points (Year 1 and Year 4). Therefore, time was considered the independent rater and this analysis established agreement over time when using the DTWRP guidelines to allocate profiles of reading difficulty at Year 1 and Year 4 within the ALP sample. This was a similar approach to longitudinal profile stability analyses within the literature such as Peterson et al. (2014). Thus, investigating the longitudinal stability of these profiles of reading difficulty throughout primary school, when allocated with the same classification system.

Finally, the third research question was addressed through utilising the ALP within-sample profile classification system. Following a similar format as the analyses above, participant profiles across Year 1 and Year 4 of the ALP sample were reported, including participants with stable longitudinal profiles and reporting movements between profile groups. A series of Cohen's Kappa analyses were then conducted, utilising the same format as the analyses above to ensure consistency between the stability analyses of both classification systems. These separate Cohen's Kappa analyses sought to determine the agreement between profiles of reading difficulty over time, when classified with the within-sample classification system. This would allow for comparisons of agreement between the DTWRP and within-sample classification systems, to determine if participants from the ALP sample displayed longitudinal profiles of reading difficulty, within either classification system. The final analysis was a Variance Components Analysis conducted with nonword and exception word reading z-scores, to establish which factors contributed variance to the stability of the word type reading performance within the ALP sample, which may have had an impact on the stability of the reading difficulty profiles. As part of this analysis, estimates of the contribution of these various factors (participant, time, and word type) were produced, to establish if instability was largely due to participant level individual differences within the ALP sample, or if the factor of time contributed to the longitudinal instability. This analysis was also interested in the interactions between these factors, to determine if reading performance and the rank order of participants changed over time, therefore causing instability within word type reading longitudinally. Overall, having implications for the longitudinal stability of profiles of reading difficulty within the ALP sample and providing a possible explanation for the lack of stability within these profiles.

3.3.2 Question 1: Profiles of Reading Difficulty within the ALP Sample determined through DTWRP classifications

This section examines the proportion of children from the ALP sample in Year 1 and Year 4 who were classified into profiles of reading difficulty according to the DTWRP assessment criteria (a stanine score below 4 on either the nonword or exception word subtests). This information is displayed in Tables 4 and 5 below.

In Year 1, the vast majority of children from the ALP sample did not display a reading difficulty profile (79.51%). Of those who displayed a difficulty profile, the largest proportion of the sample displayed an Exception word profile (9.6%) or a Mixed difficulty profile (9%). In contrast, very few children displayed a Nonword profile (1.8%). These results indicate that at Year 1, very few children displayed *specific* difficulties with nonwords in isolation, although it is important to note that the children with Mixed profiles would have had difficulties with nonwords reading, as part of their overall reading difficulty. Additionally, a higher proportion of children demonstrated a specific exception word reading difficulty, than a specific nonword difficulty.

Table 5

Reading Difficulty Profiles in the ALP Year 1 Sample using DTWRP Classifications

Year 1 Profiles	Frequencies in ALP at Year 1	Percentages in ALP at Year 1	Confidence Intervals
Nonword Difficulties	13	1.8%	0.8-2.9
Exception Difficulties	70	9.6%	7.6-11.7
Mixed Difficulties	66	9.1%	7.0-11.3
No Difficulties	578	79.5%	-

At Year 4, this pattern was partially replicated, with an even greater proportion of children from the ALP sample not displaying any reading difficulty profile (87.2%). The Nonword profile was very rare (1.4%) which had further reduced by 0.4% since Year 1. Additionally, the Exception profile now contained a smaller proportion of the sample (4.4%) than at Year 1, which had reduced by 5.2%. The Mixed profile now contained the largest percentage of the sample, out of the three difficulty profiles (7%), although this profile had also reduced by 2%. Overall, fewer children were classified with DTWRP reading difficulty profiles at Year 4 than at Year 1. The overall pattern in these DTWRP classification profiles demonstrated very few children displaying a nonword reading difficulty at either time point. A larger proportion of children either displayed a mixed difficulty with both word types, or an exception word reading difficulty in isolation, although the incidence of a specific exception word reading difficulty was reduced by Year 4.

Table 6

Reading Difficulty Profiles in the ALP Year 4 Sample using DTWRP Classifications

Year 4 Profiles	Frequencies in ALP at Year 4	Percentages in ALP at Year 4	Confidence Intervals
Nonword Difficulties	8	1.4%	0.5-2.4
Exception Difficulties	25	4.4%	2.8-6.0
Mixed Difficulties	40	7%	5.0-9.2
No Difficulties	497	87.2%	-

3.3.2.1 Question 1: Comparison of the Risk of Nonword versus Exception Profiles within the ALP Sample

The results so far have demonstrated that children within the ALP sample were classified through the DTWRP as having an Exception profile of difficulty, more than they were classified as having a Nonword profile of difficulty. A McNemar test was conducted with the ALP Year 1 and Year 4 data to indicate if these children were more likely to show difficulties with exception word reading than nonword reading in the DTWRP. Assumptions for this test were met in regard to the 2-by-2 contingency table design and random sampling from a target population of post-phonics early readers. The dichotomous dependent variable was a yes or no classification as having a reading difficulty and the independent variable was displaying either a nonword based reading difficulty or an exception word based reading difficulty. Additionally, the two groups of the dependent variable were mutually exclusive – there was no overlap between the profiles, as participants fell into one of four profile types, based on their reading difficulties with nonword reading and exception word reading: Mixed, Exception, Nonword and None. The input for the McNemar tests are displayed in Tables 7 and 8 below.

Table 7

ALP Year 1 McNemar Input

Year 1 Input	**Nonword Difficulties**	+	-
Exception Difficulties	+	Mixed (66)	Exception (70)
	-	Nonword (13)	None (578)

Table 8

ALP Year 4 McNemar Input

Year 4 Input	**Nonword Difficulties**	+	-
Exception Difficulties	+	Mixed (40)	Exception (25)
	-	Nonword (8)	None (497)

When interpreting this McNemar test, a significant p-value < 0.05 indicates that there is a statistically significant difference in the proportion of participants with a Nonword and Exception profile of reading difficulty. At Year 1, the McNemar results were as follows: McNemar $\chi 2$ (1, n = 727) = 37.78, p < .001. Therefore, there was a significantly unequal risk of profiles of reading difficulty between the DTWRP nonword and exception word measures, with participants from the ALP sample significantly more likely to show an exception word reading difficulty than a nonword reading difficulty, as supported by the descriptive statistics.

This result was also found at Year 4: McNemar $\chi 2$ (1, n = 570) = 7.76, p < 0.01. Where children from the ALP sample at Year 4 were significantly more likely to display an exception word reading difficulty than a nonword reading difficulty. However, these results are less extreme than at Year 1. These McNemar results support the existing findings above, that within the ALP sample, these children displayed more exception word reading difficulties than nonword reading difficulties longitudinally.

3.3.3 Question 2: Profiles of Reading Difficulty within the ALP sample compared to the DTWRP sample using DTWRP classifications.

The results above have discussed the profiles of reading difficulty found within the ALP sample, when using the DTWRP profile classification criteria. In contrast, this section will compare these ALP profile findings with those discovered by the DTWRP through their standardisation sample gathered in 2009 and 2011. According to the DTWRP manual, the standardisation sample were classified into each of the three reading difficulty profiles, with 5% of children displaying a Nonword reading difficulty profile, 5% of children displaying an Exception reading difficulty profile and 17% of children displaying a Mixed word reading difficulty profile (DTWRP; Forum for Research into Language and Literacy, 2012).

Using the same DTWRP stanine score classification criteria for profiles of reading difficulty, the percentages of each profile in the DTWRP pre-phonics sample can be compared to the percentages of profiles found in the ALP post-phonics sample. In order to determine if the prevalence of these profiles of reading difficulty vary post synthetic phonics introduction. This comparison can be found in Table 9 below.

Within the ALP sample, the percentage of children classified as having a Nonword reading difficulty profile at both Year 1 and Year 4 is much smaller than the 5% found in the DTWRP sample. In contrast, this was not the case for Exception reading difficulty profiles. At Year 1, the ALP sample displayed more children classified as having an Exception profile than found in the DTWRP, with almost double the DTWRP percentage. However, by Year 4, this percentage of children with an Exception profile had reduced to slightly below the percentage found in the DTWRP. Additionally, Mixed profiles at both time points within the ALP sample were much smaller than the percentage found in the DTWRP sample, which also reduced longitudinally. For both samples, the vast majority of children showed no profile of reading difficulty. It is important to note that the overall percentage of children without a reading difficulty profile was larger in the ALP sample than the DTWRP sample, which additionally grew by Year 4 when even fewer children were classified with reading difficulty profiles.

Table 9

Reading Difficulty Profiles within the ALP Sample compared with the DTWRP Standardisation Sample using DTWRP Classifications

	ALP Year 1 Percentages	ALP Year 4 Percentages	DTWRP Sample Percentages
Nonword Profiles	1.8%	1.4%	5%
Exception Profiles	9.6%	4.4%	5%
Mixed Profiles	9.1%	7.0%	17%
No Profile	79.5%	87.2%	73%

3.3.4 Question 1: Profiles of Reading Difficulty within the ALP sample determined through Z-Score classifications

The results presented so far have utilised the standardised DTWRP reading difficulty profile classifications, using DTWRP subtest age-adjusted stanine scores. The DTWRP profile classifications were useful as they provided a direct comparison to a pre-phonics sample. However, the ALP sample was larger than the DTWRP sample and therefore may provide a more sensitive estimate of the relative difficulties with nonwords and exception words in early primary school. Therefore, a decision was made to calculate relative difficulties using within-sample estimates as a cross-check of the previous findings (i.e., does the ALP sample show relatively fewer nonword than exception word difficulties, even when the profiles are classified within the same sample?). This section reports the results when reading difficulty profiles were assigned using classifications standardised within the ALP sample, rather than referring to the DTWRP sample.

To assign ALP standardised reading difficulty profiles to participants within the ALP sample, raw scores on both the DTWRP subtests of nonword and exception word reading were used. These raw scores were converted into z-scores based on the means and standard deviations for all participants across both measures. Resulting in each participant receiving two z-scores, one for nonword reading and one for exception word reading. A stringent classification system was then used, whereby profiles were allocated if one of the two z-scores fell below -1.5 standard deviation (e.g., a z-score of -1.9) whilst the other z-score fell within one standard deviation, above -1 (e.g., a z-score of 0.5). As such, Nonword profiles required a nonword z-score below -1.5 and an exception word z-score above -1, whilst an Exception profile required an exception word z-score below -1.5 and a nonword z-score above -1. Mixed profiles were allocated when both test z-scores fell below a standard deviation of -1.5. This analysis was informed by Peterson et al. (2013; 2014) who classified dyslexia subtypes in their work using in part z-scores of phonological and orthographic tasks. Whereby a "pure" reading deficit was classified as a z-score -1.5 standard deviation below the control group mean for one set of tasks, while a z-score for the alternative set of tasks was within one standard deviation of the control group mean (Peterson et al., 2013; 2014). Statistically, this -1.5 standard deviation cut off is equivalent to a stanine score of 2 and below, making this classification approach stricter than that applied in the DTWRP classification system.

The results of this reading difficulty profile classification for the ALP sample can be found in Tables 10 and 11 below. At Year 1, similarly to the DTWRP reading profile classifications, the majority of the ALP sample did not display a profile of reading difficulty, when the profile classifications were normed within the ALP sample. Interestingly, there were only two children classified as displaying a Nonword reading difficulty profile. Whilst the Exception reading difficulty profile contained the most children from the three word type profiles, with slightly fewer children displaying a Mixed reading difficulty profile. Overall, when compared to the DTWRP reading profiles allocated to the ALP sample at Year 1, there were fewer children falling into each profile of reading difficulty, although Nonword profiles remained the smallest group and Exception profiles remained the largest group.

Table 10

Reading Difficulty Profiles in the ALP Year 1 Sample using within-sample Classifications

Year 1 Profiles	Frequencies in ALP at Year 1	Percentages in ALP at Year 1	*Percentages in ALP at Year 1 using DTWRP Profiles*
Nonword Difficulties	2	0.28%	*1.8%*
Exception Difficulties	19	2.61%	*9.6%*
Mixed Difficulties	12	1.65%	*9.1%*
No Difficulties	694	95.46%	*79.51%*

At Year 4, we once again find that the majority of the ALP sample did not display a profile of reading difficulty. However, the frequencies and percentages of children within the ALP sample displaying a reading difficulty profile has increased across the Nonword (4.98%) and Mixed profiles (2.21%). The Nonword profile group is now the largest group out of the three difficulty profiles, whilst the Exception profile group is now the smallest group, displaying a reversal of the Year 1 results. Overall, when compared to the DTWRP reading profiles allocated to the ALP sample at Year 4, there are fewer children falling into the Exception and Mixed profiles, but interestingly, this is not the case for the Nonword profiles.

Table 11

Reading Difficulty Profiles in the ALP Year 4 Sample using within-sample Classifications

Year 4 Profiles	Frequencies in ALP at Year 4	Percentages in ALP at Year 4	*Percentages in ALP at Year 4 using DTWRP Profiles*
Nonword Difficulties	30	5.26%	*1.4%*
Exception Difficulties	13	2.28%	*4.4%*
Mixed Difficulties	22	3.86%	*7%*
No Difficulties	505	88.6%	*87.2%*

3.3.5 Question 3: Longitudinal Stability of Reading Difficulty Profiles within the ALP sample

The following sections discuss the stability of the reading difficulty profiles found within the ALP sample from Year 1 to Year 4. Below, the stability of the profiles as allocated through DTWRP classifications and then classifications normed within the ALP sample are examined.

3.3.5.1 Longitudinal Stability of DTWRP Classification Profiles within the ALP Sample

This section presents the longitudinal stability findings for reading difficulty profiles allocated using the DTWRP stanine score classifications within the ALP sample. The matrix of longitudinal DTWRP classification reading difficulty profiles within the ALP sample is shown in Table 12 below. It is important to note that due to participant attrition, 157 participants tested at Year 1 within the ALP sample did not participate in Year 4, therefore stability data is missing for 37 participants who were classified as having a reading difficulty profile at Year 1.

Of those participants who retained their reading profile longitudinally, from the original group of 13 children with a Nonword profile at Year 1, none of these children retained the Nonword profile at Year 4. Interestingly, 12 of these children were classified as having no profile of reading difficulty at Year 4, while one child was classified as having a Mixed profile of reading difficulty.

From the group of 70 children with an Exception profile at Year 1, 49 were tested in Year 4 and 11 retained their Exception profile longitudinally. Demonstrating reading profile stability within 22% of the Exception profile group. Of the Exception profile group who did not retain their profile longitudinally at Year 4, 1 of these children was classified as displaying a Nonword profile, 13 children were classified as displaying a Mixed profile and 24 children were classified as having no profile of reading difficulty.

From the group of 66 children with a Mixed profile at Year 1, 50 were tested in Year 4 and 17 retained their Mixed profile longitudinally. Demonstrating reading profile stability within 34% of the Mixed profile group. Of the Mixed profile group who did not retain their profile longitudinally at Year 4, 3 children were classified as having a Nonword profile, 10 children were classified as having an Exception profile and 20 children were classified as having no profile of reading difficulty.

Additionally, of the 578 children with no profile of difficulty in Year 1, 458 were tested in Year 4 and 441 retained a no profile classification longitudinally, demonstrating stability within 96.29% of the group.

Interestingly across all three reading difficulty profile groups, the largest movements from one group to another were often from the Year 1 profile to no profile of reading difficulty. Additionally, a small percentage of children with no profile at Year 1, gained a profile of reading difficulty later in Year 4 (3.71%).

Table 12

Matrix of Longitudinal Reading Difficulty Profiles within the ALP Sample with DTWRP Classifications

Year 1 Profile (N)	**Year 4 Profile (N)**				
	Nonword (8)	**Exception (25)**	**Mixed (40)**	**None (497)**	**Missing at Year 4 (157)**
Nonword (13)	**0**	0	1	12	0
Exception (70)	1	**11**	13	24	21
Mixed (66)	3	10	**17**	20	16
None (578)	4	4	9	**441**	120

3.3.5.2 Inferential test of the Longitudinal Stability of DTWRP-classified Profiles of Reading Difficulty

To investigate the stability of these profiles of reading difficulty further, Cohen's Kappa was calculated, similar to the stability analyses of Peterson et al. (2014). These results can be found in Table 13 below. Firstly, participants without a profile of reading difficulty were compared to participants with any kind of profile of reading difficulty within the ALP sample. This was the most liberal interpretation of having reading difficulties (e.g., a child could move between different types of reading difficulty, but this would still be construed as a stable reading difficulty). This resulted in a significant agreement, with a moderate Cohen's Kappa of 0.53, $p < 0.01$. Therefore, this test found significant stability between children who remain without a profile of difficulty throughout and children who consistently have any type of profile of difficulty longitudinally. This first analysis shows moderate stability.

Since there is little stability within the Nonword profile group (as shown in Table 13), participants with a nonword reading difficulty were compared to all remaining participants, including those with and without profiles of reading difficulty. Even this broad comparison resulted in a non-significant agreement, with a Cohen's Kappa of -0.017, further demonstrating that the Nonword profile group lacks stability.

The final check was to consider whether there may be something unique about Exception or Mixed word difficulties, even if there was little stability in the Nonword profiles. The final Cohen's Kappa analysis therefore compared Exception and Mixed reading difficulty profiles against Nonword and no reading difficulty profiles (i.e., participants with Nonword profiles were grouped with participants with no profile of reading difficulty, while participants with Exception and Mixed profiles were grouped together). This resulted in a significant agreement, with a moderate Cohen's Kappa of 0.56, $p < 0.01$. This result demonstrates more profile stability than when those with a reading difficulty profile are compared to those without a profile of reading difficulty. Indicating that some of the instability within the sample is due to the Nonword profiles of reading difficulty, which is often classified as no profile of reading difficulty at Year 4. Meanwhile, Exception and Mixed profiles display some stability.

Table 13

Cohen's Kappa Results for Reading Difficulty Profiles within the ALP Sample using DTWRP Classifications

Factors	Cohen's Kappa	Significance
Profile VS No Profile	0.53	$p < 0.01$
Nonword VS All Profiles	-0.017	$p = 0.69$
Nonword and No Difficulty VS Exception and Mixed Profiles	0.56	$p < 0.01$

3.3.5.3 Longitudinal Stability of Z-Score Profiles, classified within the ALP Sample

This section presents the longitudinal stability findings for reading difficulty profiles allocated using z-score classifications within the ALP sample. As discussed in the section above, 157 participants tested at Year 1 within the ALP sample did not participate in Year 4. Resulting in missing stability data for 11 participants classified as having a profile of difficulty at Year 1 within the ALP sample, using z-score classifications. The matrix of longitudinal z-score classification reading difficulty profiles within the ALP sample is shown in Table 14 below.

Of the participants who retained their profile of reading difficulty longitudinally, 1 child was classified as displaying a Nonword profile while 1 child was classified as displaying no reading difficulty profile longitudinally. Resulting in 50% profile stability, however this was from a very small group of 2 children.

From the group of 19 children with an Exception profile, only 1 child retained this profile longitudinally, demonstrating reading profile stability within 5.26% of the Exception profile group. Of the Exception profile group who did not retain their profile longitudinally at Year 4, 2 children were classified as displaying a Mixed profile of difficulty, while 9 children displayed no profile of reading difficulty.

From the group of 12 children with a Mixed profile, 6 children retained this profile longitudinally, demonstrating reading profile stability within 50% of the Mixed profile group. Of the Mixed profile group who did not retain their profile longitudinally at Year 4, 1 child was classified as displaying a Nonword profile of difficulty, while 1 child was classified as displaying no profile of reading difficulty.

From the group of 694 children with no profile of reading difficulty, 494 children retained this profile longitudinally, demonstrating stability within 71.18% of the no reading difficulty profile group. Interestingly, 54 children who had no reading difficulty profile, gained a profile of reading difficulty at Year 4. The largest proportion of these children gained a Nonword profile classification at Year 4 (51.85%), while 22.22% of the group gained an Exception profile classification and 25.93% of the group gained a Mixed profile classification.

Table 14

Matrix of Longitudinal Reading Difficulty Profiles within the ALP Sample with Z-Score Classifications

Year 1 Profile (N)	**Year 4 Profile (N)**				
	Nonword (30)	**Exception (13)**	**Mixed (22)**	**None (505)**	**Missing at Year 4 ()**
Nonword (2)	**1**	0	0	1	0
Exception (19)	0	**1**	2	9	7
Mixed (12)	1	0	**6**	1	4
None (694)	28	12	14	**494**	146

3.3.5.4 Inferential test of the Longitudinal Stability of Z-score-classified Profiles of Reading Difficulty

Following the same procedure as the Cohen's Kappa analyses in Section 3.3.4.2, Cohen's Kappa was conducted to determine the stability of these z-score classification profiles within the ALP sample further. These results can be found in Table 15 below. Following the grouping options present in the analyses in Section 3.3.4.2, participants classified as having a profile of reading difficulty with z-scores were compared to participants classified as having no profile of reading difficulty with z-scores. This resulted in a significant agreement, with a slight Cohen's Kappa of 0.207, $p < 0.01$. This test found significant stability between children who remain without a z-score profile of difficulty and children who have a consistent z-score profile of difficulty longitudinally, although this stability is slight.

The second analysis compared those with a z-score profile of Nonword reading difficulty compared to all other z-scores of reading difficulty profiles, including those without a reading difficulty profile. This resulted in a non-significant agreement, with a slight Cohen's Kappa of 0.056. Demonstrating that the Nonword profile of reading difficulties using z-score classifications has a chance level of stability, which is possible with only 2 participants classified with a z-score Nonword profile.

The final analysis combined z-score Nonword profiles of reading difficulty with no profiles of reading difficulty, compared to the z-score Exception and Mixed profile of reading difficulty groups. This resulted in a significant agreement, with a fair Cohen's Kappa of 0.296, $p < 0.01$. Similarly to the results of Section 3.3.4.2, this analysis found the strongest profile stability, more so than comparing those with a profile of reading difficulty versus those without a profile of reading difficulty. Combined with the non-significant findings of the Nonword profile versus all remaining profiles analysis, we can also determine that the z-score Nonword profile lacks longitudinal stability. In contrast, there is a fair degree of stability within the other z-score profiles of reading difficulty. It is important to note that the Cohen's Kappa values found in these results ranged from slight to fair, with less agreement found than Section 3.3.4.2.

Table 15

Cohen's Kappa Results for Reading Difficulty Profiles within the ALP Sample using Z-Score Classifications

Factors	Cohen's Kappa	Significance
Profile VS No Profile	0.207	$p < 0.01$
Nonword VS All Profiles	0.056	$p = 0.46$
Nonword and No Difficulty VS Exception and Mixed Profiles	0.296	$p < 0.01$

3.3.5.5 Overall Stability of ALP Word Type Reading Performance over Time (ALP Z-Score Data)

The results above have indicated limited stability for z-score profiles of reading difficulty within the ALP sample, especially for Nonword profiles. This may in part be due to the stringent cut off values used to determine the z-score profiles and the small groups classified with profiles. The following Variance Components Analysis therefore sought to determine the overall stability of the ALP sample z-scores across nonwords and exception word reading, to determine which other factors contributed variance amongst the sample, which may contribute to a lack of profile stability. This was a much more powerful analysis of stability than those presented above, as the full ALP sample was utilised.

This analysis used the dependent variable of nonword and exception word raw scores (0-30) which had been converted into z-scores. Therefore, performance is relative to the ALP sample data only. Fixed factors included Word Type (2 levels: nonwords versus exception words) and Time (2 levels: Years 1 and 4). Random factors included Participant, representing individual differences between each child. The results of this analysis are shown in Table 16 below.

The largest source of variance amongst the word reading test scores was due to participant's individual differences, which contributed 55% of the variance.

The next largest contributor of variance was the interaction between Participants and Time, with 17% variance, suggesting that the rank order of participants based on z-scores, changed between Year 1 and Year 4. Therefore, the relative performance of participants fluctuated in relation to each other over time. This suggests some instability in participants' overall performance.

The main variable of interest was WordType and its interactions. WordType alone accounted for 0% of the variance in the ALP sample's reading scores, as expected because these were calculated using z-scores, therefore the same distribution (with a mean of 0) would be expected for both word types. Similarly, the interaction between WordType and Time contributed 0% of the variance in this sample, which is explained by the fact that z-scores are not expected to change over time.

Alternatively, the other interactions with WordType showed some instability. The interaction between WordType and Participants explained 7% of the variance in test scores. Since this is a very low proportion of the variance, this can be interpreted as suggesting that for most participants, their relative word type performance was similar. For example, a participant who had good nonword reading would typically have good exception word reading. This suggests stability in the participant's relative ability for reading different word types.

In contrast, the interaction between WordType, Participant and Time contributed 22% of the variance in the ALP sample. This was the second largest contributor of variance in test scores and indicates that this interaction is causing instability. Specifically, participants' relative performance on nonwords and exception words was not stable, and instead fluctuated over time.

Overall, individual differences between participants contributed the largest source of variance (as expected for a heterogeneous sample). Additionally, participants' rank-order performance over time also fluctuated and so did participants' relative performance for each type of word. This is consistent with the finding that profiles of reading difficulty were unstable over time. Similarly, across the whole sample, participants were not consistent in their relative advantage for different word types over time.

Table 16

Variance Components Analysis of ALP Z-scores

Component	Percentage of Variance
WordType	0%
Participant	55%
Participant*Time	17%
WordType*Participant	7%
WordType*Time	0%
WordType*Participant*Time	21%

3.4 Discussion

Similarly to the findings presented in Chapter 2, the findings from this chapter are also interpreted in the context of intensive synthetic phonics teaching. Including the potential impact this approach to teaching reading may have had on emerging profiles of reading difficulty amongst young readers, within early primary school. The research questions for Study 1b (Chapter 3) were as follows:

Question 1: What is the prevalence of each word type reading difficulty profiles within the ALP post-phonics sample?

Question 2: How does the prevalence of each word type reading difficulty profile in the ALP sample compare to a pre-phonics sample?

Question 3: How stable are these profiles of early reading difficulty?

Before discussing the results from this study in relation to each research question, it is important to discuss the results of the initial validity check within Section 3.2.4.1, which was conducted as a validity check for the DTWRP classification system of profiles of reading difficulty, compared to performance across the ALP orthographic and phonological choice measures. As part of this validity check, median scores across the non-standardised measures of Orthographic and Phonological Choice at Year 4 were compared for early readers classified within the three profiles of reading difficulty. Results were consistent with predictions for the profiles of exception and mixed word reading difficulty. In contrast, the early readers classified with a nonword reading difficulty performed better than expected on the measure of Phonological Choice. Additionally, this group outperformed the other two groups with a reading difficulty across both choice tasks. However, there were only 8 participants in this profile of reading difficulty, which may have been too small a sample size to produce reliable results.

To address the first research question, profiles of reading difficulty within the ALP post-phonics sample were examined. The first approach utilised the standardised classification system of the DTWRP to allocate profiles of reading difficulty along the two subtests of nonword and exception word reading, using age-standardised stanine scores. Results demonstrated that within the ALP sample, there were few profiles of Nonword reading difficulty within both Year 1 and Year 4. Demonstrating that overall, this systematic, synthetic phonics instruction is having the intended consequences of improving the Non-lexical route to reading and improving nonword reading amongst early readers, as few readers displayed difficulty when reading this word type. Within this classification system, the ALP sample displayed greater amounts of Exception and Mixed reading difficulties, compared to nonword reading within both Year 1 and Year 4. Nevertheless, a greater amount of exception word reading difficulties is to be expected at Year 1, as the Lexical route to reading may not yet be developed as early readers are still consolidating their phonological decoding skills, which will then lead to independent reading experience.

At Year 4, the number of early readers classified with a profile of reading difficulty within the ALP sample reduced across all three profiles of difficulty when classified with the DTWRP. This was especially striking for exception word reading difficulties, which could be attributed to the reading development of these early readers, namely the Lexical route for whole-word reading of exception words. Meanwhile, Mixed profiles reading difficulties remained the largest group at Year 4, with movement between the Exception and Mixed profiles of reading difficulty. Specifically, by Year 4, some early readers who had an Exception profile at Year 1, were classified with a Mixed profile at Year 4, as their nonword reading may have fallen behind their peers and manifested as a "relative" Mixed difficulty. Alternatively, movement between profiles occurred in the opposite direction, whereby early readers with Mixed profiles at Year 1, were classified with Exception profiles at Year 4, as their nonword reading may have improved and instead manifested as closer to a "pure" exception word reading difficulty.

To examine this disparity between the number of early readers classified with profiles of nonword and exception word reading difficulties within the ALP sample further, McNemar analyses at Years 1 and 4 were conducted. These results determined that there was an unequal risk at both time points, with early readers more likely to be classified with an exception reading difficulty than a nonword reading difficulty. Further demonstrating a nonword reading advantage within the post-phonics ALP sample.

To investigate reading difficulties within the ALP sample further, a classification system informed by the work of Peterson et al. (2013; 2014) was utilised. Whereby, reading difficulty profiles were based on z-score classifications normed within the ALP sample. When normed within the ALP sample, there were very few profiles of nonword reading difficulty at Year 1, reflecting the nonword reading advantage provided through the systematic, synthetic phonics teaching. Additionally, there were also few profiles of reading difficulty overall at Year 1. One possible explanation is that the z-score classification system used was too stringent to capture the full range of reading difficulties within the ALP sample, as this represented a stanine score of 2 or below, considerably stricter than the DTWRP stanine scores of 4 or below.

Using this within-sample classification at Year 4, the results demonstrated a larger number of early readers classified with Nonword and Mixed profiles of reading difficulty. Interestingly, profiles of exception word reading reduced at Year 4 using the within-sample classification, possibly reflecting reading development along the Lexical route which addressed this word type reading difficulty for a proportion of the early readers. The increase in Nonword and Mixed profiles of difficulty was mostly found within early readers who originally were not classified with a profile of reading difficulty at Year 1. Therefore, one explanation is that these early readers fell behind the performance of their peers on one word type or more, as found in Steacy et al. (2014), to be classified with a profile of reading difficulty at Year 4. These findings also support the results presented in Chapter 2, which indicate that an initial nonword reading advantage within the ALP sample at Year 1, was lost at Year 4.

To address the second research question, profiles of reading difficulty within the post-phonics ALP sample were compared to the profiles of reading difficulty found within the DTWRP pre-phonics standardisation sample. Striking, the post-phonics sample showed relatively fewer children with reading difficulties overall, compared to the pre-phonics sample. This was particularly striking for nonword reading difficulties which were much rarer than expected given the DTWRP norms. Conversely, the other striking finding was the relatively frequent profiles of exception word reading difficulties, which were higher in the post-phonics sample than expected at Year 1, indicating that there is limited Lexical route development within early synthetic phonics teaching. It is important to note that phonics has not caused an increase in these exception word reading difficulty profiles, as this profile of reading difficulty declines in Year 4 to below what was found within the pre-phonics sample. Potentially due to reading development occurring between Years 1 and 4 which includes Lexical route development. While Exception profiles of reading difficulty are initially higher than expected at Year 1, these should not be a large cause for concern as they decrease by Year 4. To address those early readers with a consistent Exception profile over time, who do not improve with reading development, further research is required to determine how best to support these readers, through determining potential causes for exception word reading difficulties (e.g., poor vocabulary or limited home literacy environment).

Nevertheless, synthetic phonics teaching and natural reading development over time did not address all reading difficulties. While Mixed profiles of reading difficulty within the post-phonics sample were below the expected number according to the pre-phonics sample, this profile of difficulty contained a greater number of early readers than Nonword and Exception profiles at both time points. Indicating that this group may require interventions to address their reading difficulties along both pathways to reading, such as McArthur et al. (2015) who recommend a combination of both phonics and sight word training for overall poor readers. It is important to note that the largest movements within this group between Year 1 and Year 4 were from Mixed profiles to no profile of reading difficulty and to Exception profiles of reading difficulty. Indicating that some early readers did improve their reading performance significantly from Year 1 to Year 4, or improved their Non-lexical route to reading, which manifested as a reading difficulty resembling a "pure" exception word reading difficulty than a "relative" difficulty on both word types.

To address the third research question, the longitudinal stability of the profiles of reading difficulty within the post-phonics ALP sample was investigated. The results presented in this chapter demonstrate that overall, there was a lack of longitudinal stability within emerging profiles of reading difficulty, between Years 1 and 4 of primary school. Particularly, there was very little stability for profiles of nonword reading difficulty, whereby most of the early readers with this profile at Year 1 were classified with no profile of reading difficulty at Year 4. This indicates that the Phonics Screening Check is not diagnostic, while it may indicate early readers who require immediate support for nonword reading and development of the Non-lexical route to reading, it does not give an accurate indication of longer term needs. Especially as the Phonics Screening Check assesses nonword reading and decodable regular word reading, which is the least stable reading difficulty.

Results of Cohen's Kappa analyses with the profiles of reading difficulty found within the post-phonics sample at Year 1 and Year 4 indicated moderate to slight stability, depending on the classification system used. Specifically, greater stability was found on the first and third analyses, as results were significant, which indicated a stability higher than chance (Perera et al., 2011). In the first analysis, early readers classified with a profile of reading difficulty were contrasted with early readers not classified with a profile of reading difficulty. The third analysis compared early readers classified with no profile of reading difficulty, combined with a Nonword profile (as this reading difficulty closer resembled no difficulty due to the lack of stability) and early readers classified with either Mixed or Exception profiles. As the Cohen's Kappa values were moderate when using the DTWRP classifications, they suggest some stability within these reading profiles. With overall results indicating that Nonword profiles are unstable, whilst Mixed and Exception word difficulties show some stability, although this is relatively low.

Interestingly, there was some longitudinal stability within the post-phonics sample overall, when examined outside of these profiles of reading difficulty. Through a variance components analysis, it was determined that there was stability within the relative word type reading performance for most participants, as this interaction contributed a small amount of variance. Additionally, the instability found within the sample may be related to the interaction between word type, participant and time, which indicated that over time, participant's relative word type performance fluctuated. Corresponding with the results presented in Chapter 2, whereby the sample had a nonword reading advantage at Year 1 which changed to an exception word reading advantage at Year 4. Furthermore, this may represent the early readers who moved from one profile of reading difficulty at Year 1 to a different profile of reading difficulty at Year 4.

As Study 1b utilised the same DTWRP and ALP samples as Study 1a, it is impacted by similar limitations. The ALP sample has provided a longitudinal sample of word type reading performance amongst a large sample of early readers within England, which facilitated analysis of longitudinal profiles of reading difficulty and the respective stability of these profiles between Years 1 and 4. Unfortunately, 37 participants who were originally classified with profiles of reading difficulty at Year 1 within the ALP sample, using the DTWRP classification system, were not tested at Year 4. This attrition of participants with reading profile difficulties led to a loss of information regarding the longitudinal stability of these reading difficulties within the ALP sample. Related to this was the cross-sectional nature of the DTWRP sample. Whilst the DTWRP sample and the published figures regarding their profiles of reading difficulty provided a useful pre-phonics comparison group compared to the post-phonics ALP sample, there was no longitudinal data available for the DTWRP sample. This meant that the longitudinal stability of the profiles of reading difficulty within the DTWRP sample could not be compared with the ALP sample, to determine if the ALP sample had differing longitudinal outcomes or if the lack of profile stability within the ALP sample was novel.

Additionally, as all of the DTWRP sample was not available for analysis, it was not possible to identify how many profiles of reading difficulty were allocated to each year group within the pre-phonics DTWRP sample. To provide a comparison to the post-phonics ALP sample, the percentages of profiles of reading difficulty for the pre-phonics DTWRP sample were gathered from the DTWRP manual (DTWRP; Forum for Research into Language and Literacy, 2012). Unfortunately, these published percentages span the entire DTWRP sample (Reception to Year 7) and are not separated by year group, specifically Years 1 and 4 to compare to the data from the ALP sample. Therefore, the percentage comparisons of profiles of difficulty within the ALP and DTWRP samples are not a direct, age matched comparison.

It is also important to consider the role of extraneous variables as contributing to the lack of word type reading stability within the ALP sample. The Variance Components Analysis estimated the contributions of variance from factors such as participant, time, and word type, however without further information about the home literacy environment and the school environment that these participants were exposed to, it is difficult to eliminate the role of extraneous variables. For example, between Year 1 and Year 4, certain participants may have been selected for reading interventions either conducted in school or at home which may have improved their word type reading abilities and therefore resulted in a lack of longitudinal reading difficulty profile stability.

Despite the limitations considered above, this study has provided a valuable insight into profiles of reading difficulty amongst poor readers within mainstream classrooms in England. Firstly, through determining the frequency of profiles of word type reading difficulty within a post-phonics sample of poor readers and comparing this to profiles of word reading difficulty within a pre-phonics sample. Secondly, through investigating the longitudinal stability of these profiles of word type reading difficulty across Years 1 and 4 of primary school. Mainstream poor readers have been the focus of the study, without a Dyslexia classification, as the existing research regarding this group of poor readers is limited. Furthermore, the existing literature has not considered the role of systematic, synthetic phonics education in relation to mainstream poor readers, and whether this early phonics-based education has an impact on both profiles of word reading difficulty in regard to type and frequency, and the longitudinal stability of these profiles.

3.4.1 Conclusions and next steps

In conclusion, the results reported in this study demonstrate that within the post-phonics mainstream sample of early readers, without a dyslexia classification within the ALP, there were very few profiles of nonword reading difficulty. Therefore, this indicates that systematic, synthetic phonics is working to improve both the nonword reading of early readers and develop the Non-lexical route to reading required for phonological decoding. Conversely, a greater number of early readers were classified with exception and mixed profiles of reading difficulty, which while expected at Year 1 due to the potential lack of Lexical route development, may indicate a wider issue at Year 4. One possibility is that exception word reading is not being addressed through this systematic, synthetic phonics instruction or that this whole-word teaching is inconsistent due to a lack of formal criteria and vary from programme to programme.

However, when compared to a pre-phonics sample, overall, the post-phonics sample of early readers displayed fewer profiles of reading difficulty than expected, at both Year 1 and Year 4. Further indicating that this increased focus on systematic, synthetic phonics instruction has positive literacy outcomes for early readers, through reducing emerging reading difficulties. Especially addressing early nonword reading difficulties. This direct improvement to the Non-lexical route has potential positive outcomes for the Lexical route to reading, through an indirect improvement. Whereby the phonological decoding skills consolidated through synthetic phonics provide early readers with a “self-teaching” mechanism to independent reading, orthographic knowledge growth and the development of orthographic representations required for Lexical whole-word reading. Providing one explanation for the reduced number of exception word reading difficulties found within the post-phonics sample than the pre-phonics sample.

Additionally, all profiles of reading difficulty decreased from Year 1 to Year 4 within the ALP sample (except when classified with z-score profiles). Therefore, perhaps emerging reading difficulties within Year 1 have limited predictive validity, as it may be too early to allocate longitudinal reading difficulties as reading development is expected to change over time. This also applies to the role of the Year 1 Phonics Screening Check, which may have limited predictive validity for determining later reading difficulties amongst early readers. Notably, the majority of early readers demonstrating exception word reading difficulties at Year 1, did not retain this profile of difficulty at

Year 4. Demonstrating potential Lexical route growth through independent text experience and growth in orthographic knowledge and representations required for exception word reading.

However, some profiles of exception word reading difficulties did remain at Year 4, combined with the mixed profile of reading difficulty, which remained the largest profile group at Year 4. This is the most difficult profile of reading difficulty to address, as it reflects a combined reading difficulty across both the Non-lexical and Lexical pathways to reading, and represents the poorest readers of the sample, which was demonstrated through median Orthographic and Phonological Choice scores at Year 4. At Year 4, there was some evidence of Non-lexical improvement within this group, as some early readers moved from a mixed profile of reading difficulty to an exception word profile of reading difficulty.

Overall, there was a lack of stability found within these emerging profiles of reading difficulty within the post-phonics sample, especially for nonword profiles of difficulty. Interestingly contrasting with predictions from dyslexia and reading difficulties literature, whereby phonological difficulties display longitudinal stability (Spector, 2005; Steacy et al., 2014; Talcott et al., 2013). Indicating that systematic, synthetic, phonics instruction is successfully addressing nonword reading difficulties. In contrast, moderate to fair longitudinal stability was found within exception and mixed profiles of reading difficulty, suggesting that these emerging reading difficulties are not being entirely addressed through systematic, synthetic phonics teaching. Additionally, contrasting with the wider literature which predicts a lack of stability within exception word profiles of difficulty which represent a developmental delay which may improve with time. As these profiles of reading difficulty remain over three years of primary school education, these profiles may require interventions beyond synthetic phonics teaching.

The results presented in this chapter indicate that there are few nonword reading difficulties within a sample of mainstream, early readers who have been exposed to systematic, synthetic phonics instruction within early primary school. Combined with fewer profiles of exception word and mixed reading difficulties than expected within a pre-phonics sample, these findings suggest that systematic, synthetic phonics teaching is successfully improving reading difficulties amongst early readers, specifically along the Non-lexical route. Which as a result may indirectly improve the Lexical route to reading and hence difficulties with both exception word reading and mixed word difficulties. Now that the impact of this increased focus on synthetic phonics teaching has been examined in terms of word type reading performance and emerging reading difficulties, the next step is to examine the fundamental principles of synthetic phonics teaching. In particular, examining the focus on GPC knowledge and the development of independence when reading, through the phonological recoding process.

4 Study 2: Can Early Readers Independently Learn Novel GPCs Post-Phonics Instruction?

Abstract

This study investigated whether children in the first and second year of phonics instruction, with some GPC knowledge, were able to learn new GPCs through exposure to whole-words in sentence reading. 126 children were recruited for a small within-subjects training study to determine whether early readers who have undergone systematic, synthetic phonics teaching, can begin to “self-teach” new GPCs through whole word reading and apply their newly acquired GPC knowledge to novel words. Through training nonword and real word whole-words in sentences, context effects were also investigated, to determine if potential context effects impacted GPC learning. Participants were trained on whole-words containing the target GPCs first through sentence reading, followed by individual flashcard exposure to the target words. Participants were presented individual flashcards of the training and generalisation items to read at post-test. This study also examined which literacy-related skills provide early readers with an advantage for learning GPCs through whole-word reading. Literacy-related skills such as the amount of existing GPC knowledge, vocabulary, phoneme awareness and home literacy-environment were investigated for their potential to influence GPC learning. Findings from this study suggest that early readers are able to learn target GPCs without explicit phonics instruction, with no context effects present on GPC learning. The literacy-related skills of existing GPC knowledge, phoneme awareness and vocabulary proved important for early readers to detect, learn and generalise novel GPCs when reading. In future research, these skills could be targeted in early readers who may be struggling with their existing GPC knowledge or their ability to read independently.

4.1 Introduction and rationale

4.1.1 Early Reading and Phonics Instruction

As discussed in the introductory chapter of this thesis, Share’s (1995) influential Self-Teaching Hypothesis (1995) emphasises the process of phonological recoding as the essential first step for learning to read independently. During the process of phonological recoding, words are read by matching each letter to their corresponding sound (i.e., forming grapheme to phoneme correspondences [GPCs]). Once early readers have acquired some GPC knowledge, they can then start to use this knowledge to read through phonological recoding: sounding letters out and blending these together to pronounce a novel word. Share (1995) hypothesised that through each phonological recoding of a word, the reader develops increasingly sophisticated orthographic representations that include detailed knowledge about words, parts of words and GPCs that include for example, knowledge about the relative position of letters in words.

Revisiting this hypothesis, Share (1995) stated that phonological recoding acts as a "self-teaching" mechanism, whereby early readers who have existing GPC knowledge can decode a wide variety of items through the phonological recoding process, which can therefore expand their access to print. Through this growing experience with text, early readers develop both their orthographic knowledge and create orthographic representations, resulting in early readers gradually relying less on phonological recoding for familiar words, as orthographic representations are utilised for faster reading along the Lexical pathways (Grainger et al., 2012).

Since the introduction of the Phonics Screening Check in 2012, GPC knowledge and phonological decoding skills have proven central to systematic, synthetic phonics teaching, which is widely taught across primary schools in England, as part of the English National Curriculum. Beginning from Reception, children are taught 40+ GPCs to prepare them for both the Year 1 Phonics Screening Check, but also to begin this phonological recoding reading process (Department for Education, 2013). Nevertheless, there are many more GPCs available than can be taught in schools through systematic synthetic phonics alone. Studies such as Solity and Vousden (2009) found that children can read 90% of the words that they encounter, if they know 64 of the most common GPC mappings combined with the 100 most frequent words for these mappings in English. However, the question remains, how do early readers transition from reading with a limited set of taught GPCs, to eventually learning novel GPCs which are not taught to them? Additionally, as schools cannot teach every available GPC in the English language, can early readers develop the skills to "self-teach" themselves novel GPCs outside of traditional classroom phonics?

Studies such as Pritchard et al. (2016) through modelling the GPC-LM (Grapheme to Phoneme Correspondence Learning Mechanism) found that a computational model could learn GPCs with no existing GPC or phonics knowledge. However, the researchers acknowledge that existing GPC knowledge taught through a format such as synthetic phonics would have given the model an advantage in novel GPC learning (Pritchard et al., 2016). If the GPC-LM would benefit from prior GPC and phonics knowledge in learning new GPCs from whole words, can we apply this to early readers? Once children are aware of simple GPCs and know how to decode and blend words through synthetic phonics, does this knowledge influence their ability to learn new GPCs from whole words? In this case, Reception aged children already have existing GPC knowledge and phonics decoding skills taught to them in the Early Years Foundation Stage, so they may possess the abilities to "self-teach" themselves novel GPCs from whole-words, akin to the GPC-LM, outside of classroom phonics. There has been some evidence that early readers can teach themselves novel GPCs. Studies such as Apfelbaum et al. (2013) found that early readers can learn GPCs through whole-word training, however these GPCs were better learnt if surrounded by variable letter frames.

4.1.2 Important Skills for Self-Teaching

As part of determining if early readers can “self-teach” themselves novel GPCs, we must consider other literacy-related skills which may contribute to this ability. Spencer et al. (2015) found that statistical learning ability which Apfelbaum et al. (2013) cited as being the mechanism for implicit GPC learning, influenced oral vocabulary and phonological processing. In this study, Spencer et al. (2015) found that two statistical learning tasks, a word segmentation task and a visual sequence learning task, accounted for a unique proportion of variance amongst literacy-related skills. The visual sequence learning task predicted phonological processing skills while the word segmentation task predicted oral language and vocabulary knowledge (Spencer et al., 2015). Demonstrating that there is a relationship between this statistical learning ability and literacy and language skills (Spencer et al., 2015). If early readers can “self-teach” themselves GPCs, either through statistical learning or another route, literacy-related skills which may aid this process, or be improved through this process, should also be investigated.

As discussed in the introductory chapter of this thesis, through Sections 1.3-1.3.4, there are literacy-related skills which have been shown to impact reading development, which could also be explored to determine their potential role in a GPC “self-teaching” process. These include: phoneme awareness, letter-sound knowledge, vocabulary and home literacy environment. Furthermore, phoneme awareness and vocabulary have been linked to Share’s Self-Teaching Hypothesis (1995), as fulfilling different roles in reading acquisition. Phoneme awareness facilitates perceiving and manipulating phonemes as part of the phonological recoding process, while vocabulary provides top-down information in phonological recoding, especially when reading exception words which cannot be phonologically recoded alone for a correct pronunciation. What we do not know at this time, is which of these literacy-related skills, if any, contribute to a potential GPC “self-teaching” process, to provide early readers with an advantage when learning GPCs through exposure in whole-words.

4.1.3 The Role of Context

It is important to remember that independent reading development does not happen only through words presented in isolation, but often novel words are presented in the context of sentences. When early readers are independently reading though a text using phonological recoding, the context of the sentence in which a novel word is presented in, may affect how the child reads and subsequently learns, that novel word or GPC. A study by Landi et al. (2006) found that during the “Self-Teaching” period, early readers read words more accurately in the context of a sentence than presented in isolation. Landi et al. (2006) suggested that the context of a sentence provides top-down semantic information which may provide the child with the semantic knowledge that allows for evasion of GPC mapping. Which has also been stated by Ehri (2014), who suggests that unfamiliar words can be read through prediction by early readers, whereby the reader utilises contextual clues from the sentence to anticipate the next word, matching the spelling of the word presented to the sounds of the anticipated word. In the Landi et al. (2006) study, the children displayed better word retention for words learnt in isolation. While the initial reading of the word in a context was beneficial in the beginning, this did not result in long-lasting word learning (Landi et al., 2006). The authors suggested that the role of the sentence context drew the early reader's attention away from the orthography and phonology of the novel word, whereas these elements could be attended to when presented in isolation (Landi et al., 2006). This isolation resulted in better word retention, as the orthographic and phonological information for the novel word was stored, so when encountered again, this information could be used to read the word again (Landi et al., 2006).

A study by Stuart et al. (2000) also investigated the role of context on word learning amongst 30 five-year-old children within three training conditions: words presented in isolation on flashcards, words presented in context in a book and a mixed condition. Results demonstrated that children were able to learn words more successfully in the flashcard condition, which lacked context, than in the mixed and book conditions which included context (Stuart et al., 2000). Notably, the authors indicated that some of the difficulty within the context based conditions may be due to requiring a pre-existing concept of the word in order to predict it from the sentence context; if the word is unknown to the child both semantically and phonologically, the context may not be helpful to predict the upcoming word (Stuart et al., 2000). Whereas in the context-free flashcard condition, the child's attention is directed to the single word on the flashcard, without requiring an existing concept of the word (Stuart et al., 2000).

Wang et al. (2011) examined orthographic learning and the role of context amongst 19 children, aged 6 to 8. Specifically, this study examined whether orthographic learning was successful when occurring in the context outlined by Share's Self-Teaching Hypothesis (1995) whereby the sentence context in which a novel word is presented is important for when only a partial decoding of the word can occur (i.e., exception words). When this partial decoding occurs, the context of the sentence can be utilised to anticipate the novel word and therefore activate the semantic and phonological representations of the target word, along with the orthographic exposure (Wang et al., 2011). In contrast to the work of Landi et al. (2006), Share (1995; 1999) proposed that the context of a surrounding word facilitates orthographic learning, rather than hindering it. Through investigating the reading of regular and irregular novel words in context, it was discovered that the context of irregular items provided a facilitation effect, with stronger initial readings and acquisition of orthographic representations of these novel irregular words (Wang et al., 2011). A similar supporting context effect was also discovered for regular words, however orthographic representations for these items were not as well retained as the irregular word items (Wang et al., 2011). Whilst these findings support Share's

Self-Teaching Hypothesis (1995) and contrast with the retention findings of Landi et al. (2006) and Stuart et al. (2000), it presents an interesting question surrounding how words are taught in the "self-teaching" period, and whether context supports or hinders orthographic learning. Therefore, when examining if children have this GPC "self-teaching" ability based on their initial GPC and phonics experience, it would be interesting to determine if a sentence context, which mimics naturalistic independent reading, provides a disadvantage when "self-teaching" and retaining novel GPC information.

The evidence presented above has demonstrated that words learnt in isolation result in better word retention, due to attention being given to the orthographic and phonological information of the word, which is then retained and used to read the word when presented later (Landi et al., 2006; Stuart et al., 2000). Whereas context supports orthographic learning when the items to be read and retained are irregular and require context to activate the phonological and semantic information to supplement the partial phonological recoding of these items (Wang et al., 2011).

What is not known, is if these context effects play a role in how GPCs are learnt and retained in a GPC "self-teaching" process. Whereby GPCs may not be retained in words presented in a context, as early readers are using contextual clues to read the word and therefore do not allocate attention to the GPC, which is then not retained for future use when reading an unfamiliar word in isolation.

4.1.4 Research Questions

This study attempted to address the gaps in knowledge outlined in the introductory sections above, through investigating if early readers with limited GPC knowledge and synthetic phonics experience can "self-teach" themselves GPCs through exposure to whole words alone. Additionally, this study investigated if this newly acquired GPC knowledge is retained over time, so that the learnt GPC can then be applied to reading novel words. Research with both early readers (Apfelbaum et al., 2013) and computational models (Pritchard et al., 2016) have demonstrated that GPCs can be learnt through exposure when presented in whole-words, without utilising phonics instruction. Early readers will be familiar with other GPCs and the phonological recoding process used to decode words, through systematic synthetic phonics taught in Reception and Year 1 of primary school. Therefore, they may be able to detect GPCs in whole words and "self-teach" themselves the novel GPC based on hearing, repeating and independently reading the word across four training sessions.

It was hypothesised that early readers may be able to learn novel GPCs through exposure to whole-words, without explicit phonics instruction across training sessions. Those who then retain this newly acquired GPC knowledge will then be able to generalise this GPC knowledge when reading novel words at post-test.

Secondly, this study investigated literacy-related skills which may impact this GPC "self-teaching" ability, including existing GPC knowledge, phoneme awareness, vocabulary and home literacy environment. Predictions regarding early reading were guided by Share's (1995), Self-Teaching Hypothesis, which phoneme awareness and vocabulary have both been linked to, as potential factors which may aid the phonological recoding process. While early readers may be able to "self-teach" themselves novel GPCs from whole-words, this ability may be limited by their other literacy-related skills.

It was hypothesised that these literacy-related skills and factors may influence the ability to learn new GPCs from whole-words. Specifically, early readers with wider vocabularies and advanced phoneme awareness, both of which may be related to a varied home literacy environment, may demonstrate better "self-teaching" of GPCs. Alternatively, early readers with limited vocabularies and phoneme awareness may find it difficult to detect and learn GPCs from whole-words alone, as their ability to phonologically recode words and rely on top-down semantic vocabulary information may be limited.

Thirdly, this study investigated if context effects play a role in the GPC "self-teaching" process, through investigating the difference between GPCs learnt in real words, versus GPCs learnt in nonwords. In this design, contextual effects may be present when learning GPCs from real words presented in context, when compared to learning GPCs from nonwords in context, which cannot be anticipated or read through contextual clues. Research has demonstrated that the role of context can affect the accuracy of word learning (Landi et al., 2006). This current study was designed whereby one GPC was taught using real words and another GPC was taught using nonwords, with both word types appearing in sentences. In this case, the context of the sentence when using real words may distract from retaining the GPC information which is being taught, as early readers are using contextual clues from the sentence to anticipate the upcoming word, for example, "*The enchanted harp made a wonderful **sound***". In contrast, the GPC taught through nonwords may be better attended to, as the context of the sentence does not provide contextual clues, so the reader cannot guess the upcoming word. Therefore, the nonword word type may be given more attention to the phonological and orthographic information of the GPC, so that GPC is better retained after the training sessions have been completed.

It was hypothesised that the GPC which was taught through nonwords exclusively across the training sessions, would be better retained in the post-testing session. This includes the GPC being accurately applied to read novel generalisation words. Meanwhile the GPC taught through real words may not be retained in the post-testing session, due to less attention being allocated to the orthographic and phonological information of the GPC during the training sessions; as contextual information was being used to anticipate and read the word. Therefore, word reading with this GPC may be less accurate and this GPC may not be accurately applied to novel generalisation words.

The three research questions addressed in this study are as follows:

1. Can early readers learn GPCs through exposure to whole-words, in the context of sentences?

2. Which literacy-related skills are the most critical for early readers to be able to learn new GPCs from whole-words?

3. Are early readers better able to generalise from real words or nonwords, using new GPCs learnt from whole words in the context of sentences?

4.2 Method

4.2.1 Participants

Participants were recruited through a voluntary sample. A selection of Birmingham, UK, primary schools were contacted individually and the head teacher and class teachers were asked if their school would like to participate in the study. Once their school had agreed to participate, parents of the Reception and Year 1 class pupils were provided documents outlining the purpose and details of the study, GDPR information and parental consent forms. If the parent wished for their child to participate, they were asked to sign and complete the parental consent form which was returned to the researchers. Any child with a parental consent was invited to participate (126 children across 5 schools; 38 in the Summer Term of Reception and 88 in the Autumn Term of Year 1; 70 Female and 56 Male). One participant withdrew from the study due to illness and an additional 3 children did not want to take part, resulting in a total of 122 participants who completed all four days of the study. Of these, 6 had Special Educational Needs (SEN) and 9 had English as an additional language (EAL). No participants were excluded from the analyses presented below.

4.2.2 Power Calculations

This study utilises a prospective, cross-sectional sample of early readers and as part of determining the required sample size for this study, a priori analyses were conducted to ensure that the analyses within the study would have an adequate sample size to investigate the research questions, through detecting effects, if present.

To address the first research question, a priori power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007, 2009) to determine the minimum sample size required to achieve 0.8 power for detecting a medium effect (Cohen's $f = 0.25$) at a significance criterion of $\alpha = 0.05$ for repeated measures ANOVAs with training data, which produced a result of $N = 16$. An additional a priori analysis was conducted to determine the minimum sample size required to achieve 0.8 power for detecting a medium effect size (Cohen's f = 0.25) at a significance criterion of α = 0.05 for repeated measures ANOVAs with post-test data, which produced a result of N = 24.

To address the second research question, a priori power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007, 2009) to determine the minimum sample size required to achieve 0.8 power for detecting a medium effect (Cohen's f = 0.25) at a significance criterion of α = 0.05 for a repeated measures ANCOVA with training data, which produced a result of N = 179. An additional a priori analysis was conducted to determine the minimum sample size required to achieve 0.8 power for detecting a medium effect size (Cohen's f = 0.25) at a significance criterion of α = 0.05 for a repeated measures ANCOVA with post-test data, which produced a result of N = 128.

Unfortunately Study 2 was only able to recruit a total of 122 participants, therefore the training data ANCOVA is underpowered, achieving 0.6 power for detecting a medium effect (Cohen's f = 0.25) at a significance criterion of α = 0.05 as reported by G*Power version 3.1.9.7 (Faul et al., 2007, 2009) with a sample size of N = 122 and achieving 0.8 power for detecting a large effect (Cohen's f = 0.4) at a significance criterion of α = 0.05, with a required sample size of N = 73. Additionally, the post-test data ANCOVA is slightly underpowered, with a minimum sample size required of N = 128 as stated above, compared to the recruited sample size of N = 122.

To address the third research question, a priori power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007, 2009) to determine the minimum sample size required to achieve 0.8 power for detecting a medium effect (Cohen's dz = 0.5) at a significance criterion of α = 0.05 for a paired samples T-Test, which produced a result of N = 27.

While the largest recommended sample size of N = 179 was used a target sample size, Study 2 recruited a smaller sample of N = 122. Thus, the sample size recruited for this study (N = 122) is adequate to power all but two analyses within this study, underpowering one analysis by 6 participants and the other analysis by 57 participants for 0.8 power for detecting a medium effect. As discussed above, this sample size was suitable for detecting large effect sizes (Cohen's f = 0.4) at 0.8 power.

4.2.3 Design

A short within-subjects training study was conducted, to measure increases in GPC knowledge within this group of participants. Participants were trained on GPCs they had not yet been taught based on the programme delivery guidelines of the Letters and Sounds (DfES, 2007) and Read Write Inc. Speed Sounds (Miskin, 2006) phonics programmes. In training, target words were presented in the context of sentences. Training words were presented both in isolation and in the context of sentences, with sentences based around fantasy and sci-fi themes; in the hopes that the participants would remain interested in the task. The use of the sentence design also allowed for examination of context effects; specifically, real words may be read through contextual clues whereas this contextual information would not provide top-down semantic clues for anticipating and reading the nonwords.

Participants were trained on two diagraph GPCs ("*ou*" as in "*sound*" and "*ea*" as in "*dream*") across the study, through a combination of real words and nonwords. Participants were randomly assigned to either Group 1 or Group 2 when their parental consent was received. In total, there were 64 participants within Group 1 and 62 participants within Group 2. This design was implemented to counterbalance the stimuli across participants. In this case, we could measure the nonword versus real word training effect on both GPCs.

Group 1 learnt the "*ou*" GPC using real words only and the "*ea*" GPC as nonwords only (see Table 1, Appendix 1). While Group 2 had the opposite stimuli pattern, learning the "*ou*" GPC as nonwords as the "*ea*" GPC as real words (see Table 2, Appendix 1). These two diagraph GPCs were selected for several reasons. Firstly, the "*ea*" GPC proved to be successfully used in a training study by Apfelbaum et al. (2013) with children of a similar age to this study. Additionally, popular UK phonics programmes such as "Letters and Sounds" (DfES, 2007) and "Read Write Inc. Speed Sounds" (Miskin, 2006) were used to determine which GPCs would have been taught to children at this point in primary school, as GPCs were required which were novel to participants. Both the "*ou*" and "*ea*" GPCs are promoted in these phonics programmes to be taught throughout Year 1. Reception teachers from two of the participating primary schools also confirmed that these GPCs had not yet been taught to their pupils at the proposed time of the study.

4.2.4 Measures

4.2.4.1 GPC Screener

The study began with participants undertaking a screener to ascertain their reading performance when reading real words containing the two target GPCs. The screener required the participant to read six high frequency words aloud, three containing the "*ou*" GPC and three containing the "*ea*" GPC (as displayed in Table 3, Appendix 2). These screener items were selected as being the highest frequency, 5 letter words, when entering possible options into the SUBTLEX-UK database (van Heuven et al., 2014). The rationale for this measure was that if early readers struggled to read the highest frequency words for each target GPC, which they may have encountered before in the classroom or children's texts, then they may equally struggle with the lower frequency words which form the training stimuli. Therefore, their inability to read the higher frequency words may indicate that they do not know the target GPCs, may not be able to read the training words and might benefit from the training sessions to learn the target GPCs. All participants were included in the main study, regardless of their performance on the screener. Instead of excluding participants who knew the target GPCs, screener performance was included in the analyses presented below.

4.2.4.2 GPC Knowledge: LeST (Letter Sound Test)

Following the GPC Screener, participants were asked to complete two additional measures at pre-test. This began with the Letter Sounds Test (LeST) by Larsen et al. (2015). This was a measure of the participant's existing GPC knowledge, through assessing participant's ability to read 51 of the most important GPCs for learning to read (Larsen et al., 2015). In this measure, the amount of GPCs the participant could currently read were examined, including the "*ou*" and "*ea*" diagraphs, as a measure of their target and wider GPC knowledge. This measure was included to later analyse if the number of GPCs an early reader currently knows, has any effect on their ability to learn new GPCs. For example, if knowing many GPCs makes it easier to acquire new GPCs, as you are aware of the other letter sounds which form a word, which enables you to deduce or guess the novel GPC sound when encountered. Additionally, this test was successfully administered in a study of kindergarten children of a similar age to this study, who did not display floor or ceiling effects (Larsen et al., 2015).

4.2.4.3 Oral Vocabulary

Following the LeST (Larsen et al., 2015), participants completed an oral vocabulary assessment. In this assessment, the participant was asked to provide a verbal definition for the 10 real words which they would later be trained on. These 10 words differ between the two groups (as displayed in Tables 1 and 2, Appendix 1), as real words were only taught for one GPC per group. The words participants were asked to define were read aloud to the participant, to avoid pre-exposing them to the printed version of the word, ahead of the training sessions. As this assessment was conducted to determine if the participant had the real word in their existing oral vocabulary, they did not need to view the printed word at this stage. Definition tasks with this design have been used in similar studies to assess the vocabulary of children (Dyson et al., 2017; Nation & Cocksey, 2009; Ouelette & Beers, 2010; Ricketts et al., 2016). Determining if the participant had an oral vocabulary representation for these training words allowed for later analysis of the influence of oral vocabulary on the ability to learn new GPCs.

4.2.4.4 Home Literacy Environment

A home literacy questionnaire was sent to parents of participants for completion as part of their parental information pack, to determine the home literacy environment of the participating child (see Appendix 3). This questionnaire used a subset of items taken from the PIRLS "Leaning to Read Survey" (Martin, Mullis & Kennedy, 2007). The original PIRLS survey, akin to this study, aimed to gather information from parents and guardians regarding their own literacy-related activities and home literacy resources available (Martin et al., 2007). This subset of questions from the PIRLS survey was selected as the survey itself is well-established and was carefully created through multiple reviews and field tests (Martin et al., 2007). The questionnaire itself asked parents or guardians of participants how often they read for their own enjoyment, how much time they spend reading at home, which reading related activities their child has been exposed to and how often, how many books and children's books are in their home and their agreement with general statements about reading, for example, "I enjoy reading", as well as the highest qualification gained by either parent or guardian (see Appendix 3 for the full questionnaire). The questionnaire used various response formats, from binary yes/no responses, multiple choice responses and 4-point Likert scale responses. Through gathering home literacy environment information, this data could be analysed to determine if a varied home literacy environment influenced participant's ability to learn new GPCs. For example, if a participant who reads more at home with their parents and has a wider range of texts to access, is better able to learn new GPCs through whole words.

A factor analysis was conducted to derive factor-scores to use in analysis. Principal Axis Factoring using equamax rotation revealed 3 factors corresponding to: Factor 1: Parents' Reading Enjoyment and Participation (e.g., "I like to spend my time reading" and other questions relating to parents' own reading enjoyment and time spent reading for pleasure), Factor 2: Reading Resources (questions relating to how many books in the home and parents' qualifications) and Factor 3: Parental Involvement (e.g., the amount they engaged their child in writing, reading and alphabet-related activities). The factor scores from this 3-factor solution were used in the analyses presented below.

4.2.4.5 YARC Phoneme Awareness

The YARC Phoneme Awareness task of Phoneme Deletion (Snowling et al., 2009) was also administered. This task assessed the participant's phoneme awareness through their ability to remove individual phonemes from whole words across 19 items, such as removing the "g" sound from the word "goat". This assessment was selected due to its success in the Aston Literacy Project (ALP). This same task was used with Reception age children in the Spring term in 2011, which was a similar age group as recruited in this study. The results from the ALP revealed a wide variation of scores amongst Reception aged children, avoiding the risk of all children performing at floor or ceiling. Similar Phoneme Deletion tasks have been used in studies such as Muter et al. (2004) which have shown participant's phoneme awareness to be independent from letter knowledge. Through testing the participant's phoneme awareness, this data could be analysed to determine if phoneme awareness influenced participant's ability to learn new GPCs.

4.2.4.6 Training Word Stimuli

As described above, participants were trained on words which contained the "*ou*" and "*ea*" diagraphs, over four training sessions. This included 10 real words and 10 nonwords for each group. Nonwords and real words were used in the training sessions to investigate potential context effects when using both word types in the context of sentences, building upon the work of Landi et al. (2006). Additionally, GPC knowledge is often assessed using nonwords, as these items must be read using GPC rules and phonological decoding alone, as the participant has no prior knowledge with reading these words (Larsen et al., 2015).

Words for the training lists were selected in a variety of ways, firstly, commonly used UK Primary phonics programmes; "Letters and Sounds" (DfES, 2007) and "Read Write Inc" (Miskin, 2006) were consulted to select words that were not explicitly taught in class during the Reception year. These programmes did not provide specific guidance or word lists for nonwords, so were consulted to aid real word selection only. The target GPC diagraphs for this study were to be taught in Year 1 academic year according to both programmes, so a portion of the words selected for training originated from these phonics programmes word lists for Year 1. By using resources for the Year 1 age group, this aimed to select words which participants were *not* familiar with in Reception or the beginning of Year 1, yet were soon to be learnt and therefore within reach.

As part of creating two separate word lists for the two training groups, it was ensured that both lists were matched as closely as possible. In the training lists, both groups were trained on three words which were 5 letters long with a high frequency, then three words which are 5 letters long with a low frequency. Both groups were then trained on two words which were 6 letters long with a high frequency and two words which were 6 letters long with a low frequency. In total, all participants were trained on 5 high frequency words and 5 low frequency words. All word frequencies were determined and matched using the SUBTLEX-UK database (van Heuven et al., 2014).

As the nonword items do not have word frequency information, they were matched on word length only. Both groups were trained on six nonwords which were 5 letters long and four nonwords which were 6 letters long (matching the item lengths in the real word training words). The training words for each group can be found in Tables 1 and 2, Appendix 1.

4.2.4.7 Generalisation Word Stimuli

For the post-testing session, additional word lists were created for the post-test reading assessment, which included both the training words discussed above, and novel generalisation words. The aim of this post-testing assessment was to determine if the participants had learnt the words that they were trained on, including the "*ou*" and "*ea*" GPCs. In order to test if they had learnt the trained

GPCs, novel words were presented as part of the assessment. For participants to read these novel words accurately, they must generalise their knowledge of the trained GPCs to the novel words. Therefore, if participants read the generalisation items incorrectly, then participants may not have learnt or retained the trained GPCs.

The post-test reading assessment consisted of 40 items to be read overall (as displayed in Tables 4 and 5, Appendix 4). 20 of these items consisted of the original 10 real words and 10 nonwords that each group was trained on. The additional 20 novel words included 5 novel real words and 5 novel nonwords which corresponded to the GPCs taught in the groups. For example, Group 1 received 5 novel real words using the "*ou*" GPC and 5 novel nonwords using the "*ea*" GPC, corresponding with their training sessions. These items assessed if the participants can apply their GPC knowledge from the training sessions to words which closely resemble the training word lists.

The final 10 items included in this assessment consisted of 5 novel real words and 5 novel nonwords which included the GPC order taught in the opposite group. For example, Group 1 received 5 novel real words using the "*ea*" GPC, which they were only taught in the context of nonwords. This allowed us to determine to what extent they had learnt the GPC and whether this was limited by the context of the word versus nonword used to teach the GPC. For example, if Group 1 could not read the "*ou*" nonwords, as they had only been taught this GPC in the context of real words.

These generalisation words included in the post-test assessment were selected in a variety of ways. Firstly, to create generalisation items from trained real words, two real words from the training list were selected that were 5 letters long, one of these words had high frequency and the other word had low frequency according to the SUBTLEX-UK database (van Heuven et al., 2014). These words were then matched to similar real words with a maximum difference in two letters, which were matched for their word length and frequency (e.g., the word "*mouth*" was matched with the word "*south*"). The same procedure was then applied to two 6 letter real words, one with a low frequency and one with a high frequency (e.g., the word "*around*" was matched with the word "*ground*"). Lastly a 5-letter real word from the training list that was high frequency (e.g., "*sound*" was matched with a 5-letter low frequency word, such as "*hound*"). This procedure was applied to the generalisation word lists for both groups, to ensure the assessment at post-test was matched as closely as possible.

Secondly, to create generalisation items from the trained nonwords, three nonwords were selected from the training list which were 5 letters long and two words were selected which were 6 letters long. These nonwords were then altered by up to 2 letters to create 5 new generalisation nonwords (e.g., "*pream*" was changed to "*tream*").

Thirdly, to create generalisation items from the opposite GPC training pattern (which had the opposite pattern to the GPCs taught in the real words and nonwords category for the respective group), these items were selected from the opposite groups training list. For example, these items created for the post-test assessment of Group 1 included novel "*ea*" real words, which included five real words selected from the training list for Group 2, with no additional changes. These included three words with a 5-letter word length and two words with a 6-letter word length. The same procedure was also applied to the nonword items in this category.

In the post-test reading and generalisation assessment task, each participant was asked to read aloud 40 words in total, presented one by one on individual flashcards, which displayed only the target word to be read.

4.2.5 Procedure

The study occurred over four days; the first day was the participant's pre-test session, followed by four training sessions across two separate days and a post-testing session on the final day. Resulting in six sessions in total (see Figure 1, Appendix 5). 15 participants were not able to have their sessions on consecutive days due to absence, illness or time constraints. These participants had a maximum of 7 days between one session and another, for example, 7 days between the pre-test session and training session 1. All pre-testing, training and post-testing sessions occurred on a one-to-one basis with a research assistant and one participant, in a separate room outside the classroom. At the end of each session, the participant was thanked for their participation and taken back to their classroom.

4.2.5.1 Pre-Test Session

To begin the first session, the participant was read aloud an information sheet by the research assistant and invited to take part using a standardised script. If the child was happy to participate, they completed a child-appropriate consent form. The participant was then given a session specific sticker reward chart to track their progress, as a sticker was given to the participant for completing each task. The first task administered was the GPC Screener, the six screener items were presented in a printed list to the participant, and the participant was asked to read each word aloud, moving through the list until completion. No corrective feedback was given by the research assistant. The research assistant recorded the participant's score using a laptop and participant spreadsheet, with a score of 0 for an incorrect pronunciation and a score of 1 for a correct pronunciation for each item, until all items had been read aloud or attempted by the participant following multiple scripted prompts. A total accuracy raw score ranging from 0-6 was generated from each participant's responses.

The LeST (Larsen et al., 2015) was then administered, with the research assistant first reading the standardised instructions to the participant, outlining the task and then placing the testing booklet in front of the participant. The participant was asked to read aloud 51 graphemes, one by one, as presented in the testing booklet. The research assistant recorded their response (a score of 0 for an incorrect pronunciation or no response, a score of 1 for a correct pronunciation), until the participant had read or attempted all 51 graphemes. No corrective feedback was given by the research assistant after the first three items, only prompts for the participant to read the grapheme after five seconds if they have not attempted to read the item. If they did not attempt to read the grapheme after another three seconds, the child was asked to read the next grapheme. A total accuracy raw score ranging from 0-51 was generated from each participant's responses.

The Oral Vocabulary assessment was then administered. When administering this task, the research assistant referred to the different stimuli lists for Group 1 and Group 2, to match the participant's recorded group with the assessment items. One by one, each real word was read aloud to the participant, who was then asked by the research assistant to provide a definition for that word, using prompts such as, "Can you tell me what the word ***dream*** means?". This assessment was audio recorded using either an inbuilt laptop microphone or an audio recorder for scoring purposes. Participants were made aware of this audio recording at the beginning of the task. The research assistant later coded if the participant gave an accurate definition of each word (allocating a score of 0 for an incorrect definition or no definition given, 1 for a partially correct definition and 2 for a complete correct definition, into the participant spreadsheet). This scoring system follows those of other vocabulary definition tasks (Dyson et al., 2017; Nation & Cocksey, 2009; Ricketts et al., 2016).

The Home Literacy Environment questionnaire was completed by the parent/guardian of each participant and sent to the Principal Investigator for scoring, along with their parental consent form prior to the primary school data collection.

4.2.5.2 Training Sessions

The next day, participants were invited to begin the training sessions, for a total of 4 sessions (two on day 2 and two on day 3 of the study). During the four training sessions, across two separate days, the participant was trained on 20 words in total, 10 real words and 10 nonwords. The real words and nonwords differed on the GPC used (either "*ou*" or "*ea*"), depending on the group that the participant was allocated to. Group 1 learnt 10 real words for the "*ou*" GPC and 10 nonwords for the "*ea*" GPC, while Group 2 had word lists with the opposite GPC pattern (see Tables 1 and 2, Appendix 1).

4.2.5.2.1 Sentence Reading Task

The first task administered within the training sessions was the individual sentence reading task. In this task, a flashcard was placed on the table in front of the participant. The flashcard displayed a sentence, written in small font, with each final word in the sentence displaying a training item, which was displayed in bigger, bolded, font (e.g., "*When you fall asleep you often* ***dream***."). The participant was told that it was the research assistant's task to read the small words in the sentence. The participant's task was to read the larger, bold, word at the end of the sentence. The research assistant read the sentence, stopping at the training word to prompt the participant to read the word. The research assistant recorded a score for each training word read through each sentence, a 1 for a correct pronunciation of the target word and a 0 for an incorrect pronunciation. No phonics or decoding instructions were given to the participant, only corrective feedback if the participant failed to read the training word correctly. The same procedure was followed for all 20 training words, using sentences for both the real words and nonwords (as displayed in Tables 7-10, Appendix 6).

4.2.5.2.2 Single Whole-Word Reading Task

Once all 20 flashcard sentences had been read, all 20 flashcards were reversed to reveal a design which only displayed the single training whole-word, with no surrounding sentence. The flashcards were then shuffled and presented to the participant one by one in a random order, to avoid recency effects. The participant was asked to read aloud the single training whole-word only, one by one, moving through the flashcards until all 20 words had been read aloud. The research assistant recorded a score for each response, a 1 for a correct pronunciation and a 0 for an incorrect pronunciation. Participants were again given corrective feedback if they failed to read the training word correctly, but no decoding or phonics instructions. Participants were then given a one-hour break to return to their class.

After one hour, the participant returned for their second training session of the day and again completed the shuffled single whole-word reading task until all 20 flashcards had been completed. Following the previous procedure, scores were recorded and corrective feedback was given for an incorrect pronunciation. This task was included twice to increase the participants' exposure to the training words and unfortunately due to the time constraints of a school day, it was not possible to repeat the sentence reading task. This same procedure was followed the next day, in the participant's second training day.

4.2.5.3 Post-Test Session

Once all four training sessions had been completed, the following day the post-test session commenced, beginning with the participant completing the Post-Test reading assessment of training and generalisation words. One by one, single flashcards containing a target word were presented to the participant; once the participant had read the target word aloud, the following flashcard was presented until all 40 items had been read aloud. These flashcards were blocked so that participants read their 10 real word training words first, in the same order they were presented in the training sessions. These items were then followed by the five generalisation real words with the same GPC pattern taught in the group (e.g., "*ou*" real words for Group 1). The participant then read their 10 training nonwords, followed by five generalisation nonwords with the same GPC pattern taught in the group (e.g., "*ea*" nonwords for Group 1). Participants were then presented with a novel category of words which did not reflect the GPC patterns taught in their training sessions for their group, including five generalisation nonwords with the opposite GPC pattern and five generalisation real words with the opposite GPC pattern.

No corrective feedback was given during this post-test reading assessment and the next flashcard was presented regardless of if the participant's response was correct. This assessment was audio recorded for accuracy and a research assistant recorded two scores for each item. The first score was a whole word score, where a 1 was given if the participant read the whole word correctly and a 0 was given if the participant read the whole word incorrectly. The second score was a GPC only score, where a score of 1 was given if the participant read the target GPC within the word correctly, even if they read the surrounding letters wrong (e.g., "bream" read for "dream" would be a score of 1 as the GPC was read correctly). A score of 0 was given if the GPC was read incorrectly. If the participant corrected their own mispronunciation within one attempt with no prompting from the researcher, it was then coded as a correct response instead of an incorrect response. The scoring of this task was triple checked by other research assistants and the principal investigator through reviewing the audio recording of the task and correcting any incorrect scores.

Following the reading assessment of training and generalisation items, the YARC Phoneme Awareness task (Snowling et al., 2009) was administered. This task presented participants with 19 individually printed pictures depicting items such as animals or food. The participant was read aloud the word corresponding to the picture, for example, "goat", by the research assistant and was then asked to repeat the word back to the researcher. After the repetition the participant was then asked to remove a sound from the word (e.g., remove "*g*" from "*goat*"). The first four items asked participants to remove syllables from the items (e.g., "*ice*" from "*ice-cream*") and the remaining items asked participants to remove individual phonemes. The researcher scored the participant for each of the items, with a score of 0 given for an incorrect answer and a score of 1 given for a correct answer. All 19 items were scored, and no stop rule was used, as this study reports raw phoneme deletion scores and did not use the accompanying YARC standardised scores. As this data was used for within-sample analyses, standardised scores were not required.

Once they had completed both tasks, the participant was given their final participation sticker for their chart, thanked for their hard work throughout the study and given a debrief sheet. This debrief was read by the research assistant to the participant to explain what the study had been investigating and what happens next. After providing them with the chance to ask any additional questions, the participant was taken back to their classroom.

4.3 Results

4.3.1 Data Analysis Strategy and Exploration of Data

The analyses for this study varied according to which of the three research questions were addressed. The first research question asked: can early readers learn GPCs through exposure to whole-words, in the context of sentences? This question was addressed through firstly examining reading performance across the four training sessions, through a repeated measures ANOVA. In order to determine if reading performance of training items changed over time and additionally, if this performance differed by word type (nonwords versus real words). This test was chosen as ANOVA with repeated measures would determine whether raw whole-word reading scores differed across the four training sessions. This was to determine if the training sessions were successful and if this differed between nonwords and exception words within the same group of participants. This was included to determine if training was limited by lexicality. Secondly, reading performance at post-test was examined through a repeated measures ANOVA, which indicated whether participants were able to read both trained and generalisation items at post-test and therefore learnt the trained GPCs and applied these to novel words. This repeated measures ANOVA was chosen to determine whether raw GPC reading scores across generalisation and trained items differed within the sample, who had not received training on the generalisation items. This difference was investigated to determine whether the sample of early readers were able to learn the GPC from within the training items and apply these to generalisation items. Alternatively, if there was a large difference between the trained and generalisation items with a reading advantage for trained items, GPC learning may not have occurred. Lexicality was also investigated within this repeated measures ANOVA, to determine if post-test reading performance of trained and generalisation items was limited by the lexicality of the items (nonwords or exception words).

The second research question asked: which literacy-related skills are the most critical for early readers to be able to learn new GPCs from whole-words? This question was addressed through analysing both the training data and post-test data separately. Firstly, a series of ANCOVAs were conducted, to establish which literacy-related skills contributed to GPC learning throughout the training sessions. This approach was chosen to systematically analyse the impact of each literacy-related skill on GPC “self-teaching” ability, to ensure that initially significant effects were not missed, which may have occurred if added to a singular ANCOVA for training and post-test data respectively. As part of ensuring that literacy-related skills were investigated thoroughly, these significant literacy-related skills were confirmed through positive, significant Pearson's correlation coefficients with total whole-word scores across training sessions and post-test GPC reading scores.

Once significant literacy-related skills within the training data had been determined, these were entered into a repeated measures ANCOVA to examine these literacy-related skills further, through examining interactions with the variables of Time and Lexicality (nonwords versus real words). This analysis revealed which literacy-related skills effected GPC learning within the four training sessions. A repeated measures ANCOVA was selected as the appropriate analysis to determine the individual contribution of each literacy-related skill to whole-word reading across the training sessions within the whole sample. This ANCOVA also revealed if there were interactions between these literacy-related skills and the variables of Time and Lexicality.

Following this analysis, GPC reading performance at post-test was analysed in combination with the literacy-related skills. In order to determine which literacy-related skills effected GPC reading and generalisation at post-test. A series of ANCOVAs were conducted, to establish which literacy-related skills significantly affected post-test GPC reading scores. These significant covariates were then entered into a repeated measures ANCOVA to examine interactions with the variables of Lexicality and Word Type (trained or generalisation items). This analysis revealed which literacy-related skills effected GPC reading and generalisation at post-test. A repeated measures ANCOVA was selected as the appropriate analysis to determine the individual contribution of each literacy-related skill to post-test GPC reading within the whole sample. This ANCOVA also revealed if there were interactions between the literacy-related skills and the variables of Lexicality and Word Type.

The third research question asked: are early readers better able to generalise from real words or nonwords, using new GPCs learnt from whole words in the context of sentences? To address this research question, a paired samples T-Test was conducted to establish whether participants performed better at post-test with generalisation words containing the GPC learnt through nonwords or real words during their training sessions. This analysis determined if there were lexicality and therefore context effects present during the training sessions which affected GPC generalisation at post-test.

4.3.2 Question 1: Does Reading Performance improve across Training Sessions? (Training Data)

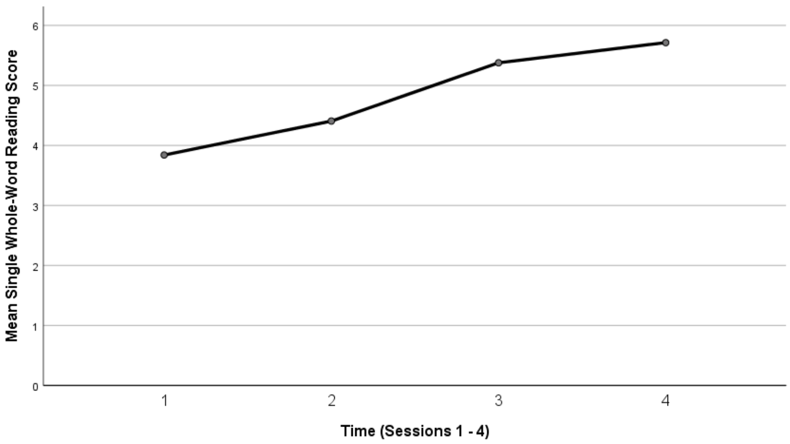
This first analysis used the dependent variable of reading performance scores (0-10 for each word type) across the four training single whole-word reading tasks. The first independent variable in this analysis was Time, to establish if reading performance scores differed across the four instances of the task and address if GPC learning had occurred as proposed in research question one. The second independent variable was Lexicality, to establish if reading performance scores differed for real words and nonwords across the four tasks as proposed in research question three.

Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $x^2(5) = 40.58$, $p = 0.01$. Since sphericity was violated, ($\varepsilon = 0.817$), Huynh-Feldt corrected results are reported.

A repeated measures ANOVA showed that the main effect of Time was significant, $F(1.73, 209.46) = 74.74$, $p < 0.001$, with a large effect size of $\eta^2 = 0.38$. As shown in Figure 4.1 below, participant's reading scores increased across the four training sessions. The main effect of Lexicality was significant, $F(1, 121) = 32.06$, $p < 0.001$, with a large effect size of $\eta^2 = 0.21$. Participants achieved higher scores for real words ($M = 21.47$, $SD = 14.43$) than nonword items ($M = 17.20$, $SD = 15.07$) across the four training sessions. There was no significant interaction between Lexicality*Time, $F(2.51, 303.11) = 2.09$, $p = 0.113$, $\eta^2 = 0.017$.

Figure 4.1

Single Whole-Word Reading Performance across Training Sessions



4.3.3 Question 1: Did Early Readers learn GPCs and apply this knowledge to Generalisation Items? (Post-Test Data)

The post-test analysis used the dependent variable of post-test GPC reading scores (0-40), meaning that an item was scored as correct if the target GPC inside the item was read correctly, for example, a pronunciation of "bream" for the item "dream" would be marked as correct, as the error was not the target GPC. The independent variables were Lexicality (real words and nonwords) and Word Type (trained items and generalisation items) to establish whether participants were able to generalise their new GPC knowledge to novel items they had not been exposed to during the training sessions. Mauchly's Test of Sphericity was not available due to the requirement of three conditions to conduct the test. Therefore, as no information regarding sphericity was available, the Greenhouse-Geisser correction was used and these corrected results are reported below.

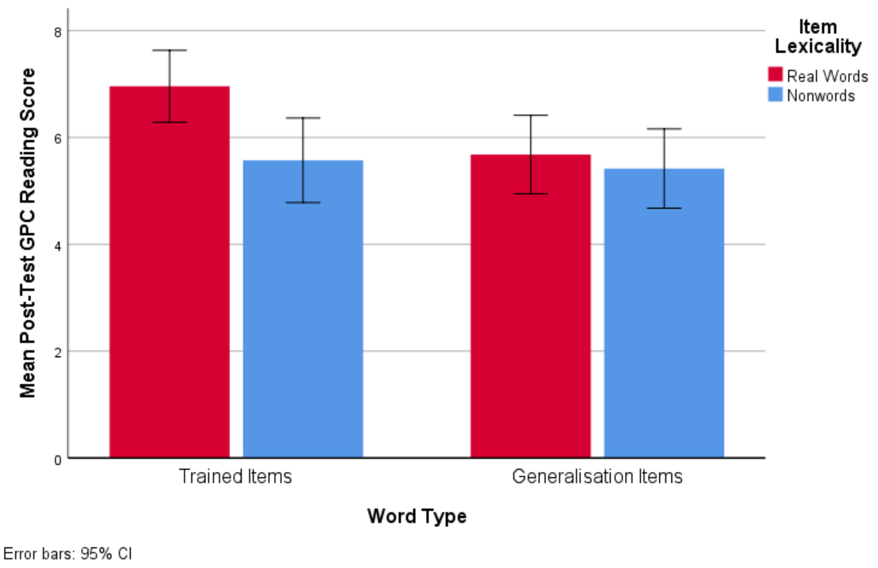
Participants achieved higher scores for real word (M = 12.64, SD = 7.46) than nonword items at post-test (M = 11, SD = 8.37), a repeated measures ANOVA showed that the main effect of Lexicality was significant, $F(1, 121) = 26.68$, $p < 0.001$, with a large effect size of $\eta^2 = 0.18$.

Participants achieved higher scores for trained items (M = 12.53, SD = 7.51) than novel generalisation items (M = 11.1, SD = 8.15), presented for the first time at post-test. The main effect of Word Type was significant, $F(1, 121) = 37.32$, $p < 0.001$, with a large effect size of $\eta^2 = 0.24$.

There was also a significant interaction between Lexicality*WordType, $F(1, 121) = 12.65$, $p = 0.001$, with a medium effect size of $\eta^2 = 0.095$. This pattern is explored further below, as it relates to Question 3. As shown in Figure 4.2 below, the lexicality effect is apparent for trained items, as accuracy was higher for real words than nonwords that were trained. Whereas accuracy for real words and nonwords were similar for generalisation items. This suggests that the significant lexicality effect reported above was driven by the trained words.

Figure 4.2

Post-Test Reading for Trained and Generalisation Real Word and Nonword Items



4.3.4 *Question 2: Which literacy-related skills effect GPC learning across Training Sessions? (Training Data)*

Nonetheless, there may be other literacy-related factors that affect GPC "self-teaching", as proposed in the second research question. These literacy-related factors were not included in the previous analyses so their role in GPC "self-teaching" has been undetermined. To address this, a series of ANCOVAs were conducted following the procedure of the previous ANOVA in Section 4.3.2., except in each instance one literacy related factor was included as a covariate.

Covariates were as follows: Knowledge of target GPCs (Screener Total Score), Knowledge of wider GPCs (LeST Total Score), Phoneme Awareness (YARC Total Score), Vocabulary (Vocabulary Total Score) and Home Literacy Environment (3 factors; Parents' Reading Enjoyment and Participation; Reading Resources; Parental Involvement).

The only literacy-related covariates which had significant effects on training reading performance scores were Screener Total Score, LeST Total Score, YARC Total Score and Vocabulary Total Score. These were also confirmed with Pearson's correlation coefficients to assess the relationship between each literacy-related skill and the raw whole-word reading training score from across the four training sessions. These are reported as follows: Screener Total Score, $r(120) = .79$, $p = 0.01$, LeST Total Score, $r(120) = .74$, $p = 0.01$, YARC Total Score, $r(120) = .76$, $p = 0.01$ and Vocabulary Total Score, $r(120) = .32$, $p = 0.01$.

Next, all of these significant literacy-related factors were added into a Repeated Measures ANCOVA, with a dependent variable of reading performance scores across training sessions and independent variables of Time (reading performance scores from the four training sessions) and Lexicality (real words and nonword item scores).

Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $x^2(5) = 38.53$, $p = 0.01$. Since sphericity was violated, ($\varepsilon = 0.82$), Huynh-Feldt corrected results are reported.

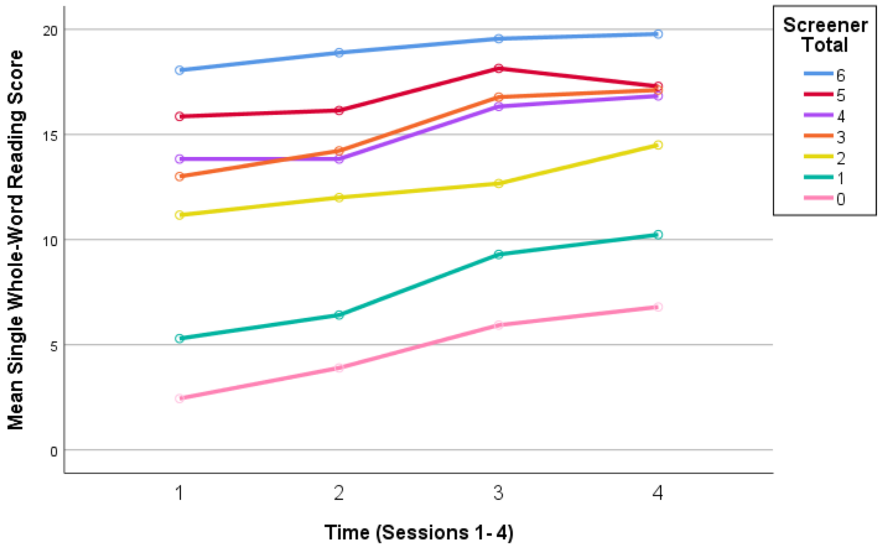
Once these covariates were included, the main effect of Time became non-significant $F(1.89, 220.62) = 0.17$, $p = 0.83$, $\eta^2 = 0.001$ as did Lexicality $F(1, 117) = 1.23$, $p = 0.27$, $\eta^2 = 0.01$. There also was no significant interaction between Lexicality*Time $F(2.6, 303.13) = 0.35$, $p = 0.76$, $\eta^2 = 0.003$.

All covariates had a significant effect on training reading performance scores except for Vocabulary as follows: Screener scores: $F(1, 117) = 44.17$, $p < 0.001$, with a large effect size of $\eta^2 = 0.27$; LeST scores: $F(1, 117) = 12.09$, $p = 0.001$, with a medium effect size of $\eta^2 = 0.09$ and YARC Phoneme Deletion scores: $F(1, 117) = 26.8$, $p < 0.001$, with a large effect size of $\eta^2 = 0.19$. The overall pattern for these covariates was for higher scoring participants to generally score higher on reading performance during the training sessions.

There were significant interactions between Time and the following covariates on training reading performance: Screener: $F(1.89, 220.62) = 13.72$, $p < 0.001$, with a medium effect size of $\eta^2 = 0.105$ and YARC Phoneme Deletion: $F(1.89, 220.62) = 6.41$, $p = 0.003$, with a small effect size of $\eta^2 = 0.052$. As shown in Figure 4.3 below, participants scoring lower on the Screener made greater improvement during training. Ceiling effects were present for the highest scoring participants on the Screener, who did not improve greatly across the training sessions.

Figure 4.3

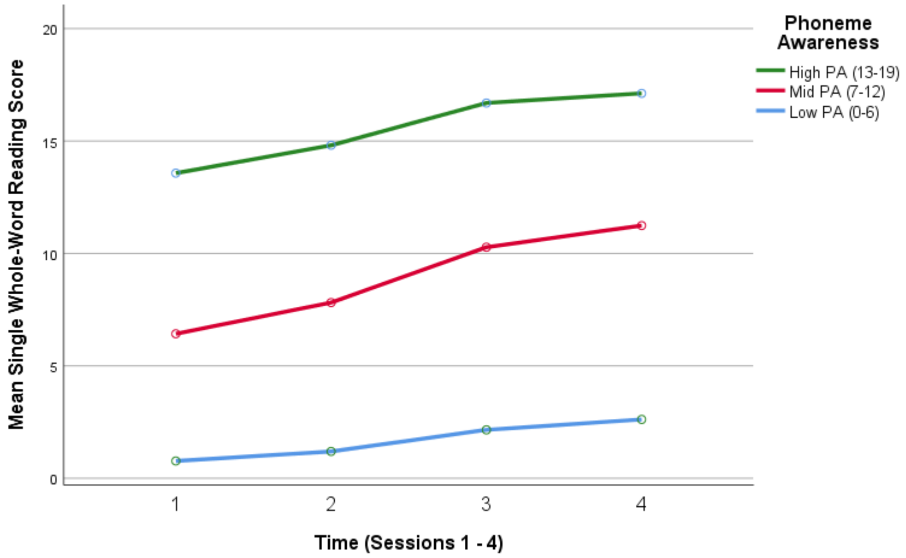
Profile Plot of Single Whole-Word Reading Scores by Screener Score



As shown in Figure 4.4 below, participants with low Phoneme Awareness made less progress than participants with average Phoneme Awareness across the training sessions. Participants with high Phoneme Awareness also made less progress across the training sessions but this may be potentially due to ceiling effects.

Figure 4.4

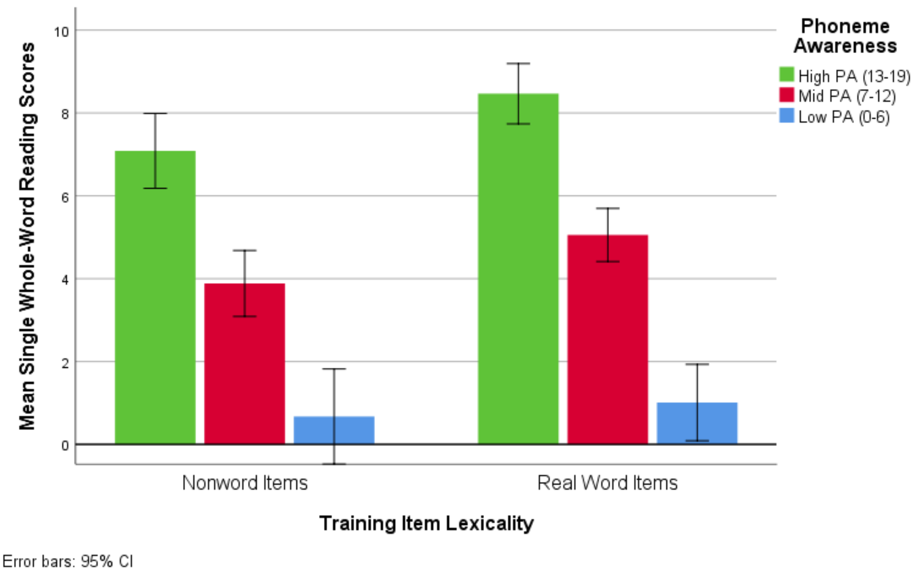
Single Whole-Word Reading Scores by YARC Phoneme Deletion Score



There was also a significant interaction between Lexicality and YARC Phoneme Deletion on training reading performance scores $F(1, 117) = 5.74$, $p = 0.018$, with a small effect size of $\eta^2 = 0.047$. As shown in Figure 4.5 below, participants with higher Phoneme Awareness showed a stronger advantage for real words than participants with low Phoneme Awareness.

Figure 4.5

Nonword and Real Word Reading across Phoneme Awareness Scores



4.3.5 *Question 2: Which literacy-related skills effect GPC reading and generalisation at Post-Test (Post-Test Data)*

This analysis examined whether individual differences in literacy-related factors effect GPC "self-teaching" and generalisation at post-test. A series of ANCOVAs were conducted using the independent variables of Lexicality (real words and nonwords) and Word Type (trained or generalisation items) and the dependent variable of post-test GPC reading scores (0-40). With the addition of each literacy related factor as a covariate.

Covariates were as follows: Knowledge of target GPCs (Screener Total Score), Knowledge of wider GPCs (LeST Total Score), Phoneme Awareness (YARC Total Score), Vocabulary (Vocabulary Total Score) and Home Literacy Environment (3 separate factors from an earlier Factor Analysis), including Factor 1: Parents' Reading Enjoyment and Participation, Factor 2: Reading Resources and Factor 3: Parental Involvement.

When added one at a time, the only literacy-related covariates which had significant effects on post-test GPC reading scores were Screener Total Score, LeST Total Score, YARC Total Score and Vocabulary Total Score. These were also confirmed with Pearson's correlation coefficients to assess the relationship between each literacy-related skill and the raw post-test GPC reading score. These are reported as follows: Screener Total Score, $r(120) = .66$, $p = 0.01$, LeST Total Score, $r(120) = .67$, $p = 0.01$, YARC Total Score, $r(120) = .66$, $p = 0.01$ and Vocabulary Total Score, $r(120) = .34$, $p = 0.01$.

Next, all of these significant literacy-related factors were added into a Repeated Measures ANCOVA, with a dependent variable of post-test GPC reading scores (0-40) and independent variables of Lexicality (real words and nonwords) and Word Type (trained or generalisation items).

Mauchly's Test of Sphericity was not available due to the requirement of three conditions to conduct the test. Therefore, as no information regarding sphericity was available, the Greenhouse-Geisser correction was used and these corrected results are reported below.

Once these covariates were included, the main effect of Lexicality became non-significant, $F(1, 117) = 0.39$, $p = 0.53$, $\eta^2 = 0.003$, as did the effect of Word Type: $F(1, 117) = 2.37$, $p = 0.13$, $\eta^2 = 0.02$. Similarly, the interaction between Lexicality*WordType became non-significant: $F(1, 117) = 0.89$, $p = 0.35$, $\eta^2 = 0.008$.

However, the following covariates had a significant effect on post-test GPC reading scores: Screener scores: $F(1, 117) = 11.07$, $p = 0.001$, with a medium effect size of $\eta^2 = 0.086$; LeST scores: $F(1, 117) = 7.95$, $p = 0.006$, with a medium effect size of $\eta^2 = 0.064$ and YARC Phoneme Deletion scores: $F(1, 117) = 11.63$, $p = 0.001$, with a medium effect size of $\eta^2 = 0.09$. The same pattern was shown for all of these covariates: in general, participants who scored highly on the covariates also scored highly on post-test GPC reading. This main effect is shown in Figures 4.6-4.9 below. Conversely, this was not the case for Vocabulary, which had no significant effect on post-test GPC reading scores: $F(1, 117) = 3.2$, $p = 0.077$, $\eta^2 = 0.027$.

There were no significant interactions between Lexicality and covariates on post-test GPC reading however the interaction between Lexicality*Screener was borderline significant, $F(1, 117) = 3.17$, $p = 0.056$, with a small effect size of $\eta^2 = 0.031$. In general, a higher screener score was associated with a smaller advantage for real words over non-words at post-test compared to lower screener scores, as shown in Figure 4.7 below.

Figure 4.6

Post-Test Mean GPC Reading Total Score by Phoneme Awareness Scores

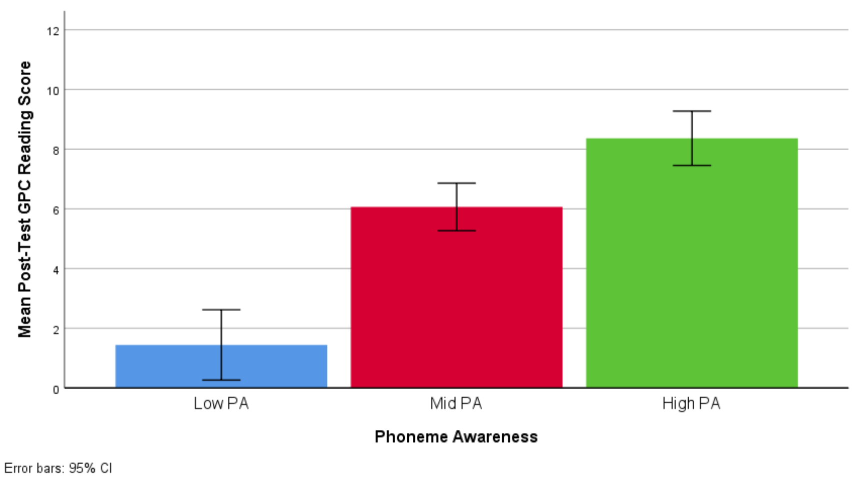
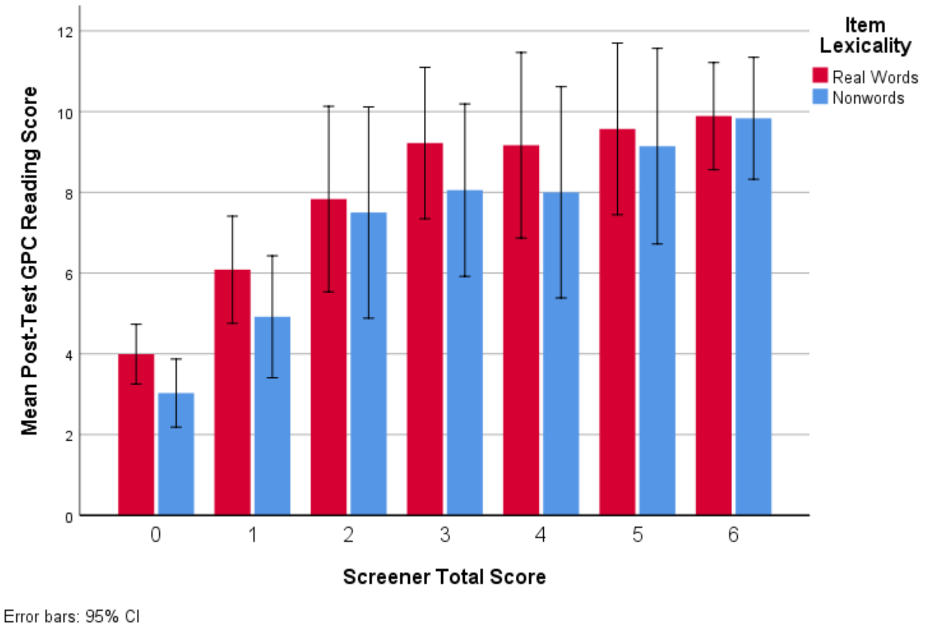


Figure 4.7

Post-Test Real Word and Nonword Reading by Screener Score



There were significant interactions between Word Type and the following covariates on post-test GPC reading: Screener $F(1, 117) = 15.92$, $p < 0.001$, with a medium effect size of $\eta^2 = 0.12$; LeST $F(1, 117) = 10.14$, $p = 0.002$, with a medium effect size of $\eta^2 = 0.08$ and Vocabulary $F(1, 117) = 4.32$, $p = 0.04$, with a small effect size of $\eta^2 = 0.036$. As shown in Figure 4.8 below, the advantage for trained over generalisation words was apparent only for participants scoring less than 3 on the GPC Screener. A similar pattern was shown for the LeST, with the effect of word type driven by lower scoring participants, as shown in Figure 4.9 below. Figure 4.10 below shows a similar pattern for Vocabulary, where participants scoring below 4 displayed a greater advantage for trained over generalisation words. Figure 4.11 also displays this trend within the YARC Phoneme Deletion measure, whereby a larger advantage for trained words compared to generalisation words was found amongst participants with low phoneme awareness scores, compared to participants with average to high phoneme awareness scores. Despite this, it is important to note that this finding reflects the general trend shown with the aforementioned literacy-related skills, as this this interaction was not significant ($F(1, 117) = 0.005$, $p = 0.95$, $\eta^2 = 0.00$).

Figure 4.8

Post-Test Trained and Generalisation Item Reading by GPC Screener Score

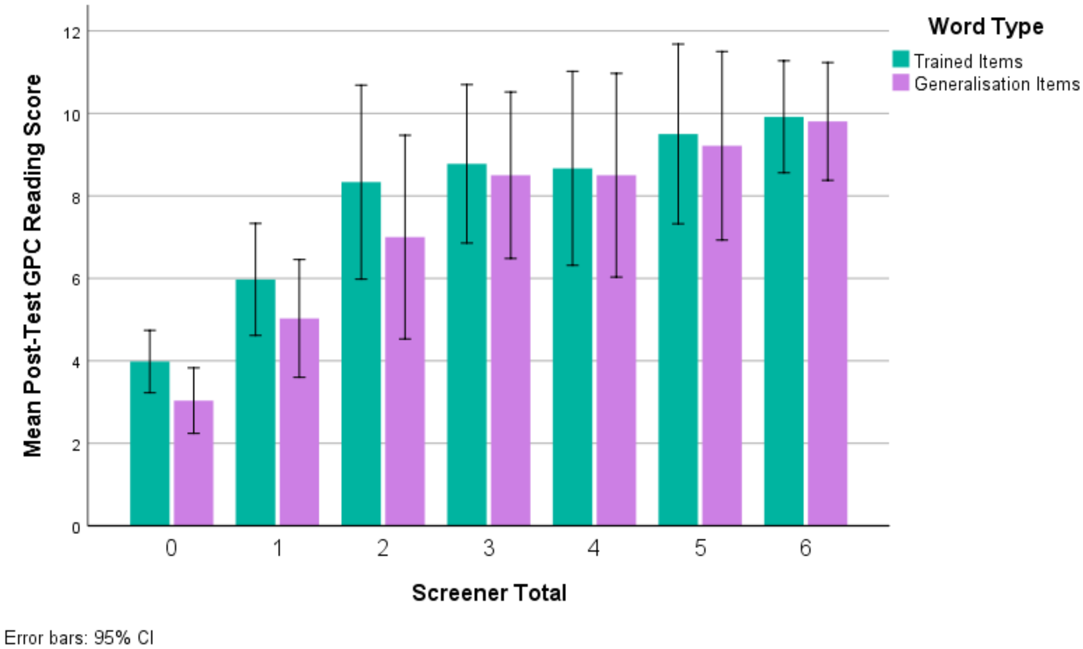
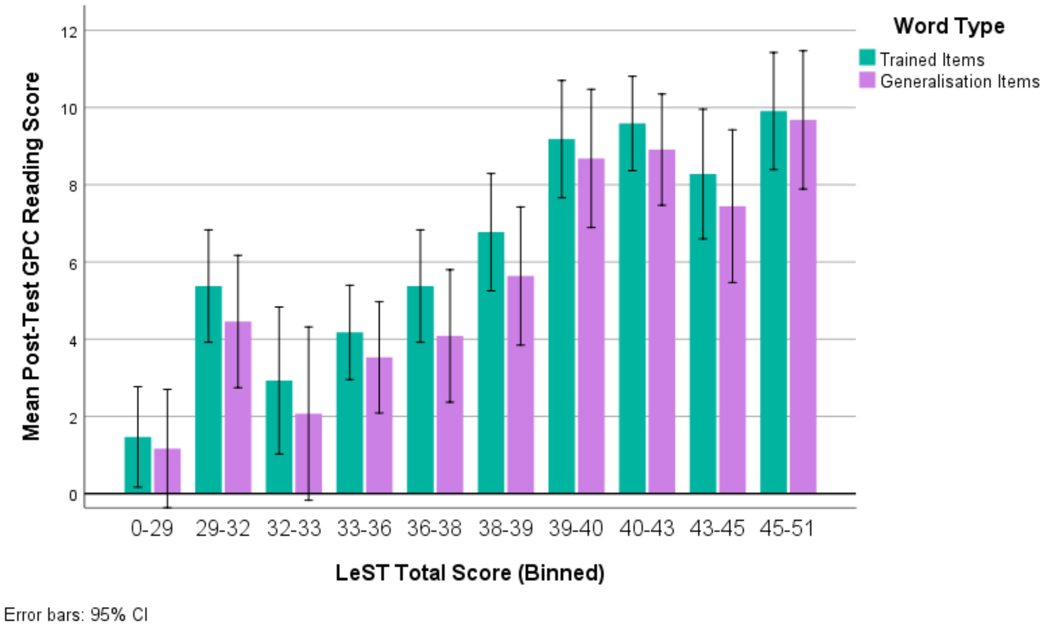


Figure 4.9

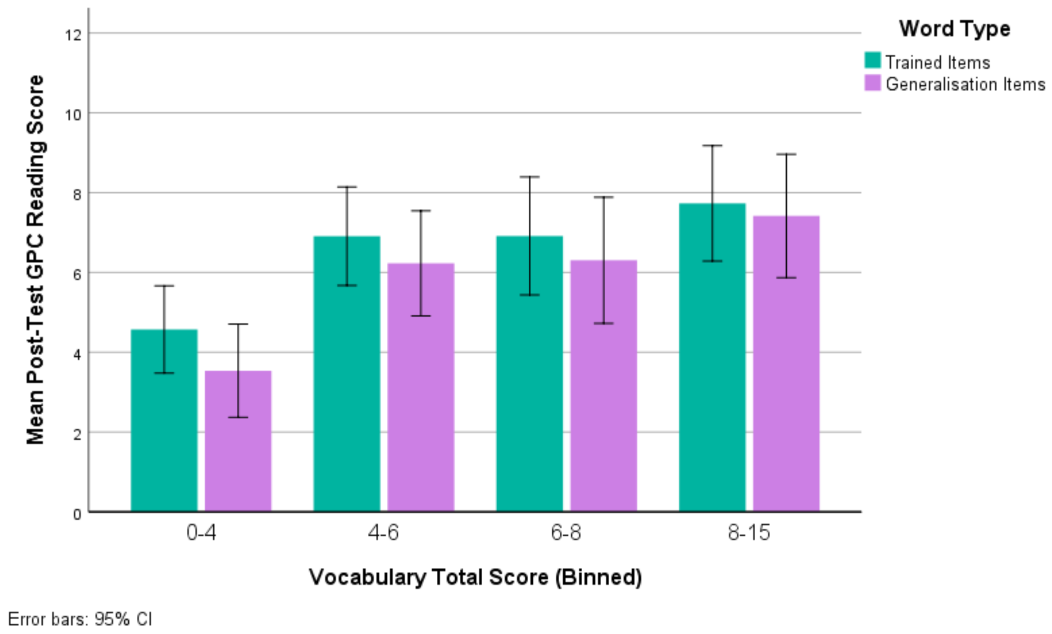
Post-Test Trained and Generalisation Reading scores by LeST Scores



Note. LeST Total Scores binned at equal 10% percentiles based on scanned cases.

Figure 4.10

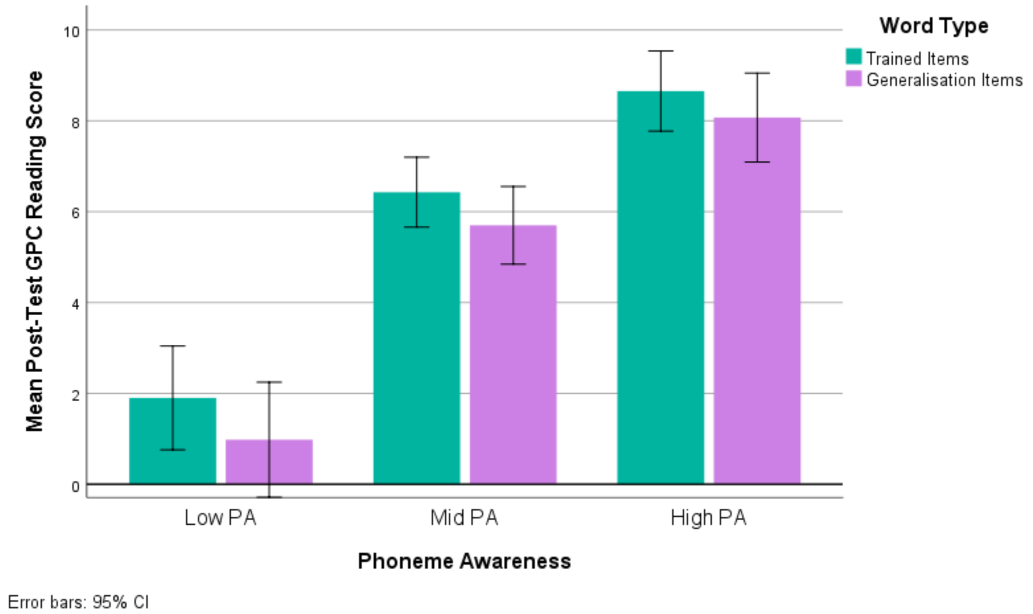
Post-Test Trained and Generalisation Item Reading Scores by Vocabulary Scores



Note. Vocabulary Total Scores binned at equal 25% percentiles based on scanned cases.

Figure 4.11

Post-Test Trained and Generalisation Item Reading Scores by YARC Phoneme Deletion Scores



Note. The findings presented above reflect an overall trend as this interaction was not significant ($p = 0.95$).

4.3.6 *Question 3: Are early readers better able to generalise from real words or nonwords, using new GPCs learnt from whole words?*

A paired samples T-Test was conducted to establish whether participants performed better at post-test when reading generalisation words with GPCs learnt in their nonword training pattern (/*ea*/ GPC for Group 1 and /*ou*/ GPC for Group 2) or words learnt in their real word training pattern (/*ou*/ for Group 1 and /*ea*/ for Group 2). In fact, the results found no significant effect of lexicality on generalisation word reading, and the mean accuracy was very similar for the GPCs trained as nonwords (M = 5.42, SD = 4.62) and as real words (M = 5.68, SD = 4.12); t(121) = .859, p = .392.

4.4 Discussion

This study has demonstrated that early readers are able to learn GPCs through exposure to whole words in sentence reading, without explicit phonics instruction. Interestingly, there were no context effects present from the sentence reading design, as GPCs learnt through real words and nonwords read in context displayed similar retention and generalisation at post-test. Additionally, literacy-related skills such as existing GPC knowledge, phoneme awareness and vocabulary influenced participant's ability to learn GPCs from whole words. This discussion section is separated to address each of the research questions independently. The research questions for Study 2 were as follows:

Question 1: Can early readers learn GPCs through exposure to whole-words, in the context of sentences?

Question 2: Which literacy-related skills are the most critical for early readers to be able to learn new GPCs from whole-words?

Question 3: Are early readers better able to generalise from real words or nonwords, using new GPCs learnt from whole words in the context of sentences?

In regards to the first research question, this study found that children's accuracy in reading the target GPCs increased significantly over the training sessions. Since they were only given exposure to these GPCs in the context of whole words, this suggests that early readers are able to learn target GPCs without explicit phonics instruction, presumably through GPC "self-teaching" (Share, 1995). As these children had already experienced systematic, synthetic phonics teaching within the first year of the primary school curriculum, which may have facilitated phonological recoding and provided the skills necessary for GPC "self-teaching". The amount of improvement within the training sessions was significantly related to the participant's pre-test Screener score and existing knowledge of the target GPCs: participants with lower initial target-GPC knowledge made greater improvement than participants with higher Screener scores of 5 or 6. This may reflect a ceiling effect for the higher scoring participants, for whom less improvement was possible.

Interestingly, the opposite pattern was found for phoneme awareness: participants with poor phoneme awareness displayed less improvement in reading performance across the four training sessions, than participants with average to high phoneme awareness. As discussed in the introduction of this thesis, phoneme awareness is vital for the phonological recoding stage of Share's Self-Teaching Hypothesis (1995). Without this knowledge of how graphemes correspond to phonemes within a word, a word cannot be phonologically recoded and this independent "self-teaching" mechanism cannot develop (Share, 1995). The results from this study support this notion, as perhaps the low phoneme awareness group struggled to decode the training items due to their lack of phoneme awareness, which also made it difficult to detect and learn the target GPCs. This accounts for both their lack of improvement across the training sessions and their overall poor GPC reading at post-test. As well as supporting the finding that participants with low phoneme awareness struggled to read the training items overall regardless of the lexicality of the item.

Whereas participants with mid and high phoneme awareness performed more highly for real word items within the training sessions. This suggests that a familiarity with the real word items was an advantage for the better performing readers and provides no evidence of any context effects: children were certainly not distracted by the context of the sentence during the training sessions (which would have resulted in better performance for nonwords).

Overall, these results demonstrate that participants could learn GPCs from whole word exposure, and interestingly children who had poorer GPC knowledge prior to training made greater improvement, than children with some existing target GPC knowledge. However, poor phoneme awareness was negatively associated with participant's ability to benefit from whole-word exposure.

To address the second research question, the results from this study suggest that post-test GPC reading of trained and generalisation items was related to three literacy-related skills. Namely, GPC knowledge, phoneme awareness and vocabulary. It was discovered that post-test reading performance of both trained and generalisation items was related to a participant's Screener score at pre-test and therefore their existing knowledge of the training GPCs prior to the training sessions. Participants with high Screener scores at pre-test had some knowledge of the training GPCs; these participants could then use this knowledge to accurately read both trained and generalisation items at post-test, resulting in no word type effect. In contrast, participants with low Screener scores may have had a word type advantage for trained items at post-test due to the amount of exposure they received to these items across the training sessions. However, there was only a small difference between the scores achieved for trained items compared to generalisation items amongst participants with low

Screener scores. This demonstrates that these participants benefitted from the training sessions and learning of the target GPCs from whole-word exposure did occur. This newly acquired GPC knowledge was then applied to enable these participants to read novel generalisation items they had not been exposed to prior to the post-test. If participants had not benefitted from the training sessions, a wider disparity between their scores for trained and generalisation items would be expected.

Secondly, these results suggest that post-test GPC reading of trained and generalisation items was related to participant's existing wider GPC knowledge, assessed through the LeST (Larsen et al., 2015). Participants with existing wider GPC knowledge could use this information to read both trained and generalisation words accurately, resulting in little word type effect. Participants with lower LeST scores and presumably less wider GPC knowledge displayed a small advantage for trained items, again potentially due to the amount of exposure participants had to these items across the training sessions. Nonetheless, these participants with limited wider GPC knowledge still read generalisation words accurately at post-test. Demonstrating that learning of the target GPCs had occurred throughout the training sessions and were then applied to post-test generalisation items.

Thirdly, these results suggest that post-test GPC reading of trained and generalisation items was related to participant's Vocabulary score at pre-test, over and above their initial GPC score. The majority of participants with high vocabulary scores, scored similarly for both generalisation and trained items at post-test. This finding may occur for various reasons, firstly, these participants may be using information from their wide vocabulary to read the generalisation items, resulting in little difference between their generalisation and trained item scores. Secondly, participants with high vocabulary scores may be proficient readers with greater independent reading experience, who also have wider GPC knowledge or existing knowledge of the target GPCs. Thirdly, referring back to the Lexical Restructuring Model (Walley et al., 2003), these readers with high vocabulary scores may have greater phonological recoding abilities, through the growth in phoneme awareness and specified phonological representations which are associated with vocabulary growth and independent reading. Therefore, these early readers may have successfully detected, learnt and applied the target GPCs at post-test, resulting in little word type reading differences. Whereas participants with low vocabulary scores may not have a vast vocabulary to draw information from and read the generalisation items. Alternatively, these readers may have difficulties in phonological recoding and accessing independent reading, resulting in a lack of vocabulary, phoneme awareness and phonological representation growth according to the Lexical Restructuring Model (Walley et al., 2003). Resulting in an advantage for trained items, potentially due to the amount of exposure given to these items across the training sessions. It is important to note that the difference in post-test scores between generalisation and trained items was small for participants with low vocabulary scores. Demonstrating that these participants were able to read some generalisation items accurately, showing that the target GPCs had been learnt across the training sessions.

The post-test GPC reading scores of participants were also found to be related to participant's phoneme awareness scores, however this literacy-related skill displayed no significant interaction with trained or generalisation items at post-test, unlike GPC knowledge and vocabulary. Similarly to the findings across the training sessions, at post-test, the low phoneme awareness group displayed poor GPC reading whilst the average to high phoneme awareness groups did demonstrate some GPC learning and overall higher GPC reading scores at post-test.

To address the third research question, the results from this study suggested that participants were not better able to generalise from GPCs learnt in real words or nonwords at post-test, demonstrating similar performance across generalisation words with their trained nonword and real word GPC patterns. This disproves the original prediction that there may be a context effect during the training sessions, with more attention given to nonword GPCs, as contextual effects were used to predict and read the real word GPCs. If this context effect was present, the results would have shown a significant difference in post-test reading of generalisation words, with better reading performance for generalisation words which contained the GPCs trained in nonwords. These findings do suggest that early readers are utilising phonological recoding both when reading in context and when reading in isolation, as described by Share's Self-Teaching Hypothesis (1995), until orthographic learning and orthographic representations develop. Therefore, as this study utilised GPCs which were not yet taught in the respective year groups, it would not be expected that orthographic representations of these GPCs and training words would have developed prior to the training. Demonstrating that the early readers who were able to "self-teach" themselves the target GPCs, utilised phonological recoding to facilitate this. Unfortunately, the specific prediction made by Share (1995) regarding the context of a sentence facilitating word reading for partial phonological recoding (i.e., exception words), was not directly assessed.

While this study has revealed important findings regarding GPC "self-teaching" ability amongst early readers with systematic, synthetic phonics experience, it is not without its limitations. Firstly, the design of the study was limited to an active training group only, which was split into two groups to counterbalance the training stimuli. Therefore, the study lacked a control group for between-groups comparisons, to determine the impact of the GPC whole-word training compared to typical classroom instruction. Future research may address this through conducting a randomised control trial in a setting where systematic, synthetic phonics is not the mandatory teaching approach within mainstream classrooms and conduct a similar training study with a control group for comparison. Secondly, the design of the training sessions included a mixture of contextual sentences and whole-word exposure, which was changed from the original proposed design due to the limited time available within a school day. Therefore, the role of context may be nonsignificant in this study, as participants did not receive substantial contextual exposure to the training items after only two trials. Additionally, the remainder of their exposure to the whole-words was through non-contextual flashcards across four trials before the post-test. Future research may address this through increasing the exposure to contextual stimuli within the training sessions or through providing training with resources such as short stories designed for early readers. Thirdly, the literacy-related skill of vocabulary could have been measured through a standardised test of vocabulary depth compared to the oral vocabulary definitions task given to participants. Whilst similar measures of oral vocabulary have been utilised in existing literature with early readers (Dyson et al., 2017; Nation & Cocksey, 2009; Ouelette & Beers, 2010; Ricketts et al., 2016), this measure lacked age standardised information. This measure reflected whether participants had the word in their oral vocabulary, rather than a wider measure of vocabulary depth and breadth, which may have impacted the role that vocabulary had on training and post-test reading scores, as a literacy-related skill. Future studies could include standardised measures of receptive vocabulary such as the British Picture Vocabulary Scale (BPVS) as well as an oral vocabulary definition task to measure both facets of vocabulary. Finally, this study did not control for SES and FSM data across the participating schools. Whilst the primary schools who participated did span various geographical locations across Birmingham, it cannot be determined that there were varying levels of SES and FSM within the sample. This lack of SES and FSM data combined with the geographical location limited to Birmingham, UK, limits the generalisability of the results across England. Future research could seek to conduct a similar training study or as suggested, a randomised control trial across a wider geographical location to recruit a large sample reflective of various socioeconomic settings.

4.4.1 Conclusions and next steps

In conclusion, this study has demonstrated that GPC “self-teaching” is possible amongst early readers, following exposure to systematic, synthetic phonics instruction within the first two years of the English National Curriculum. This evidence provides support for the fundamental principles of Share’s Self-Teaching Hypothesis (1995), which states that the ability to phonologically recode through utilising existing GPC knowledge and phoneme awareness facilitates independent reading, leading to orthographic learning and reading “by-sight”. The results demonstrate that early readers with phonological recoding experience through synthetic phonics teaching are able to use this experience as a “self-teaching” mechanism, as explicit phonics instruction was not given during this study, yet participants were able to detect, learn and apply the target GPCs to novel items.

The results also indicated that wider literacy-related skills contributed to this GPC “self-teaching” ability. Specifically, phoneme awareness played a key role, as early readers with poor phoneme awareness did not demonstrate as much improvement throughout the training and scored below the readers with average to high phoneme awareness on GPC reading at post-test. GPC knowledge also proved to be important, as early readers with greater GPC knowledge tended to perform better overall. Readers with limited GPC knowledge did demonstrate some improvement, indicating that a lack of GPC knowledge alone does not stop GPC “self-teaching”, however these readers did not score as highly as readers with greater GPC knowledge. Vocabulary also proved important during GPC reading at post-test but only for trained versus generalisation word types, rather than overall performance. Early readers with higher vocabulary scores displayed less of a word type effect, potentially indicating that they were overall better readers. Early readers with low vocabulary scores achieved higher scores for trained items than generalisation items at post-test, however displayed a difficulty within GPC reading of generalisation items. To determine the underlying cause of this GPC reading difficulty amongst early readers with low vocabulary scores, further research is required. Specifically, to determine if these early readers have a poor vocabulary due to having less independent reading experience or whether they are poor readers overall who cannot access independent reading through difficulties within phonological recoding and therefore struggled to detect, learn and generalise the target GPCs. Additionally, the role of home literacy environment did not appear to significantly affect GPC “self-teaching”.

Interestingly, no context effects were found between the training items of real words and nonwords in sentences. Indicating that early readers were not using contextual clues to anticipate and read the training words and are instead allocated equal attention to both real word and nonword items. These findings indicate that when teaching early readers target GPCs, these can be presented both in sentence context (e.g., decodable books) and in isolation (e.g., flashcards and word lists), for successful GPC “self-teaching”. However, to investigate the direct predictions made by Share’s Self-Teaching Hypothesis (1995) an alternative version of the training study could be conducted to determine if this lack of context effect remains when irregular word items are trained; as only regular words and nonwords which could be phonologically recoded were included in this study. As Share’s Self-Teaching Hypothesis (1995) directly discusses the role of context as facilitation for activating phonological and semantic representations of a target word, to “boot-strap” pronunciation when only partial decoding can occur due to irregularities.

5 General Discussion

This thesis investigates longitudinal reading performance, longitudinal reading difficulties and reading self-teaching ability present in early readers from English primary schools, following the increased focus on systematic, synthetic phonics teaching. The reading performance and profiles of reading difficulty found amongst early readers who have received systematic synthetic phonics instruction have also been compared with a sample of early readers who were assessed before the introduction of the mandatory Phonics Screening Check in 2012. An experimental study was then conducted with early readers in the Reception and Year 1 to determine their ability to "self-teach" GPCs through whole-word exposure and apply these to novel words, following the phonics instruction they received as part of the National Curriculum. This thesis includes three empirical chapters and the 11 key findings from these studies are outlined and discussed below.

5.1 Summary of findings

In Chapter 2, the reading performance of nonwords and exception words within a post-phonics sample of early readers was investigated longitudinally at Years 1 and 4 of primary school. The first key finding was that this post-phonics sample had an initial advantage for nonword reading at Year 1, compared to exception word reading. In contrast, at Year 4, there was an advantage for exception word reading compared to nonword reading, as the curriculum placed less focus on phonics teaching and early readers began reading whole-words. However, at both time points, reading across both word types was average and above, indicating that systematic, synthetic phonics teaching had not caused an initial nonword reading advantage at the expense of exception word reading.

Additionally, the reading performance of this post-phonics sample was compared to a pre-phonics sample collected during the DTWRP measure standardisation, before the Year 1 Phonics Screening Check was introduced in England. The second key finding emerged from comparisons of the two samples, which revealed that the post-phonics sample had an overall reading advantage, demonstrating better reading performance across both word types, at both Year 1 and Year 4, compared to the pre-phonics sample. The covariate of age significantly contributed to reading performance within Year 1, however by Year 4 this covariate was non-significant.

In Chapter 3, the emerging profiles of reading difficulty within the ALP post-phonics sample were investigated. The third key finding indicated that when utilising a standardised classification system (DTWRP) and a within-sample classification system, the post-phonics sample had very few Nonword profiles of reading difficulty at Year 1 and a greater amount of Exception word and Mixed reading difficulty profiles. The fourth key finding indicated that at Year 4, all profiles of reading difficulty had reduced within the post-phonics sample when using a standardised classification system, reflecting a similar pattern to Year 1 with greater numbers of Exception and Mixed profiles than Nonword profiles.

Within Chapter 3, these emerging profiles of reading difficulty were compared to emerging profiles of reading difficulty found within the aforementioned pre-phonics sample, using the DTWRP classification system. The fifth key finding demonstrated that at both time points, the post-phonics sample displayed fewer emerging profiles of reading difficulty than expected, compared to the pre-phonics sample. The sixth key finding demonstrated that this smaller amount of emerging reading difficulty profiles was especially striking for Nonword profiles of reading difficulty, which was much smaller within the post-phonics sample compared to the pre-phonics sample at both time points.

Additionally, the longitudinal stability of these emerging profiles of reading difficulty within the post-phonics sample was examined in Chapter 3. The seventh key finding demonstrated that across Year 1 and Year 4, the Nonword profile of reading difficulty had the least stability and more closely resembled a "no difficulty" profile. While the eight key finding demonstrated that across Year 1 and Year 4, the Exception and Mixed profiles of reading difficulty displayed some stability, albeit moderate to fair. These findings also occurred when profiles were allocated with the DTWRP classification system, or the within-sample classification system.

In Chapter 4, the GPC "self-teaching" ability of early readers who had received systematic, synthetic phonics instruction within the early years of primary school were investigated. The ninth key finding demonstrated that early readers with experience of systematic, synthetic phonics teaching, were able to detect, learn and generalise novel GPCs through whole-word training, as a form of "self-teaching".

As part of this training study, the role of literacy-related skills were examined, namely, phoneme awareness, letter-sound knowledge, vocabulary and home-literacy environment. The tenth key finding demonstrated that the literacy-related skills of phoneme awareness, letter-sound knowledge and vocabulary significantly contributed to GPC "self-teaching". Home-literacy environment was found to be non-significant.

Finally, the role of context on this GPC "self-teaching" ability was investigated, to determine if GPCs learnt within nonwords, as part of sentences, would be better learnt, retained and generalised than GPCs learnt within real words, as part of sentences. The eleventh key finding demonstrated that there was no context effect present on GPC "self-teaching", when early readers learnt GPCs from whole-words, as GPCs leant within real words and nonwords were equally generalised.

5.2 Theoretical implications of these findings

For the purpose of this thesis, the main theoretical predictions regarding early reading and reading development are taken from the dual-route reading perspective. This includes both the DRC (Coltheart et al., 2001) and the Triangle Model (Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg & Patterson, 1996; Seidenberg & McClelland, 1989). Despite the fact that these two models differ considerably in their design and implementation, they both describe two routes to reading which reflect a Non-lexical route and Lexical route and therefore share the same predictions from a dual-route reading perspective. Additionally, Share's Self-Teaching Hypothesis (1995) can also be viewed as consistent with a dual-route reading perspective, as this hypothesis provides a developmental account of how early readers learn a Non-lexical route and then transition from a dependence on the Non-lexical route towards developing sufficient orthographic knowledge and detailed representations to support a Lexical route.

The Non-lexical route is associated with reading novel words, which is especially relevant for early readers who are less familiar with print and therefore utilise phonological decoding to "sound-out" GPCs within a target word and blend these sounds together to produce a pronunciation of the target word (Stuart & Stainthorp, 2015). As a pre-requisite for conducting this phonological decoding, early readers must master the alphabetic principle (Bryne, 1998, as cited in Hulme et al., 2012) which requires two key underlying skills: phoneme awareness and letter knowledge (Hulme et al., 2012).

Over time as reading develops, orthographic representations of whole words are created and words are instead read "by sight", through this direct Lexical route from orthography to phonology, which does not require labour intensive phonological decoding. Additionally, as oral vocabulary develops over time, this facilitates the storage of semantic information which is then linked to orthographic and phonological representations of known words (Stuart & Stainthorp, 2015). Which as a result, facilitates reading along the Lexical-Semantic route present in both the DRC and Triangle Model, whereby words are read through an indirect route from orthography to semantics to phonology (Stuart & Stainthorp, 2015).

The findings reported in this thesis have a number of theoretical implications consistent with a dual-route perspective of reading. Firstly, the results presented in Chapter 2 regarding word type reading over time support the notion of dual reading routes and the transition from the Non-lexical route to the Lexical route of reading, as reading develops. Specifically, Chapter 2 demonstrated that at Year 1, there was a nonword reading advantage in the post-phonics ALP sample when compared to exception word reading, providing support for the dual-route prediction that early readers utilise the Non-lexical pathway and read novel items through phonological recoding. As nonwords do not have existing semantic or orthographic representations, these items must be read exclusively through phonological recoding along the Non-lexical pathway (Coltheart, 2005).

By Year 4, this initial nonword reading advantage changed into an advantage for exception word reading, although nonword reading performance was average to above average. This later advantage for exception word reading also provides support for reading development from a dual-route perspective. Specifically, as reading develops, readers are creating orthographic representations through orthographic learning which then facilitates direct reading “by sight” along the Lexical route to reading. Therefore, providing an advantage for exception word reading, as this word type is read “through sight” as whole-words, either directly from orthographic to phonological representations, or indirectly through semantic information, as this word type cannot be phonologically recoded along the Non-lexical route, without producing regularization errors (Stuart & Stainthorp, 2015). Additionally, these developing readers who are utilising the Lexical pathway to reading are relying less on the early reading Non-lexical pathway and therefore the initial advantage for nonword reading is diminished. However, theoretically the two routes operate in tandem, as the Non-lexical pathway is utilised when novel words without orthographic representations are encountered and therefore these participants were still able to read the nonword items successfully; providing further support for the dual-route approach to reading.

In terms of overall reading performance across the nonword and exception word types, the largest contributors to variance within the ALP post-phonics sample also corresponded with predictions from the dual-route perspective. Firstly, individual differences were prominent, which would be expected as every early reader has made differing levels of progress within their reading development. Secondly, the interaction between Participants and Time reflected an overall change of rank performance between participants from Years 1 and 4, as some participants showed greater relative improvement whereas others fell behind their peers. Thirdly, the interaction between WordType, Participant and Time indicated that participants’ relative performance on each word type changed over time. This suggests that participants’ advantage for one word type is not stable and is consistent with the finding reported earlier, that an initial nonword advantage at Year 1, changed to an exception word advantage at Year 4: reflecting development of the Lexical pathway through orthographic learning and less reliance on the Non-lexical pathway. Additionally, there was very little variance from the interaction of WordType and Participant, indicating that participants’ reading ability was fairly consistent across both word types. Specifically, those who performed well on one word type, tended to perform well on the other word type. Conversely, participants who performed poorly on one word type tended to show difficulties with the other word type (although as discussed next, there were profiles of children who showed more specific difficulties).

The findings in Chapter 3 also support this same pattern of Non-lexical dominance in early reading, whereby there were few Nonword profiles of reading difficulty at Year 1 compared to greater numbers of Exception word profiles of difficulty, when classified both with a standardised classification system (DTWRP norms) and a within-sample classification system (within-sample z scores). This lack of Nonword profiles of reading difficulty within early readers demonstrates that the Non-lexical pathway required to read nonwords had developed successfully in most participants who were able to phonologically decode nonword items above what was expected for their age group, with few participants displaying a difficulty when reading this word type.

Conversely, the greater than expected proportion of early readers at Year 1 displaying an Exception word reading difficulty can also be explained from a dual-route perspective, as these early readers were potentially relying on the Non-lexical pathway to phonologically decode items, which resulted in difficulties with exception word reading due to the regularization errors. Additionally, these early readers were still growing their GPC knowledge through systematic, synthetic phonics teaching in preparation for the Year 1 Phonics Screening Check, which as a result would facilitate phonological decoding and allow for orthographic learning through independent reading experience. From a dual-route perspective, these early readers may have had limited orthographic learning experience and therefore limited use of the Lexical pathway. Resulting in greater numbers of Exception word profiles of reading difficulty, as these words could not be successfully phonologically recoded along the Non-lexical pathway or read along the limited Lexical pathway these early readers had, if they did not have existing orthographic representations for the exception word items.

At Year 4, these profiles of difficulty allocated with the DTWRP classification system reflected improvement across all profiles of difficulty, potentially reflecting general reading development as both pathways to reading improved. Especially the Exception word profiles of difficulty, which from a dual-route perspective, may have improved due to early readers developing the Lexical pathway to reading and creating orthographic representations through orthographic learning between Year 1 and Year 4 in primary school. Meanwhile, Mixed profiles of difficulty demonstrated less improvement. From a dual-route perspective, Mixed profiles reflect difficulties along both pathways to reading, as readers display reading difficulties with both nonwords and exception words which utilise different routes to reading, operating in tandem. This Mixed profile of reading difficulty with limited improvement may reflect that these early readers are overall "poor readers", who may not be expected to improve their reading performance through reading experience and orthographic learning alone, due to difficulties along both pathways. Similar to the notion of a "relative" deficit found in dyslexia literature such as Peterson et al. (2013), whereby these early readers with a Mixed profile of reading difficulty demonstrate difficulties along both the Non-lexical and Lexical pathways. Meanwhile, Nonword and Exception profiles of reading difficulty instead reflect difficulties within only one of the two pathways to reading and could be classified as a "pure" deficit.

When using a within-sample classification system for profiles of reading difficulty, the theoretical implications are consistent with the discussion above at Year 1 but less clear at Year 4. At Year 1, the ALP sample displayed a lack of Nonword reading difficulty profiles and greater proportion of Exception word reading difficulty profiles, which is in accordance with the dual-route perspective. Mixed profiles of reading difficulty displayed the most stability from the three profiles of reading difficulty, which as discussed above, would be expected according to the dual-route reading perspective as these overall poor readers with relative difficulties would not be expected to improve with time and reading experience alone. In contrast, at Year 4, there were more Nonword profiles of reading difficulty than Exception profiles of reading difficulty, which would not be expected according to the dual-route perspective. One possibility was that from Year 1 to Year 4, the initial nonword reading advantage was diminished due to the reduced focus on phonics instruction and movement towards developing the Lexical route to reading. As a result, readers who may not have been detected as having a profile of Nonword reading difficulty at Year 1, as they did not meet the stringent "cut-off" criteria, fell behind their peers by Year 4, to reveal a Nonword profile of reading difficulty.

In terms of Share's Self-Teaching Hypothesis (1995), this was explored in Chapter 4 and supporting evidence was found that supported the hypothesis. Specifically, that early readers can "self-teach" themselves novel GPCs once they have had experience with phonological recoding and phonics through systematic, synthetic, phonics teaching. As proposed in the Self-Teaching Hypothesis (Share, 1995), phonological recoding is the facilitator to independent reading experience, through allowing early readers to access text independently through novel word decoding. As results indicate in Chapter 4, this independent reading is extended even further, if following phonics experience and instruction, early readers are able to "self-teach" novel GPCs; which would further facilitate phonological recoding through utilising wider GPC knowledge. The results of Chapter 4 therefore have potential theoretical implications, through building upon Share's existing hypothesis (1995) and indicating that early readers can "self-teach" GPCs, once they achieve a threshold of GPC knowledge and experience with phonological recoding through phonics. Nevertheless, the results of Chapter 4 are unable to determine the exact thresholds of knowledge required to facilitate GPC "self-teaching". These findings also have implications for phonics teaching, as schools cannot currently teach every single GPC mapping within the English language and studies have found that early readers can access a wide variety of novel words with limited GPC knowledge combined with irregular word knowledge (Solity & Vousden, 2009). These findings indicate that as early readers can GPC "self-teach" once a threshold of GPC knowledge and phonological recoding experience is achieved, phonics teaching could be simplified once this threshold of knowledge is established.

Chapter 4 also found that phoneme awareness was important for this GPC "self-teaching", which also corresponds with Share's discussion of phoneme awareness as part of the Self-Teaching Hypothesis (1995). Specifically, early readers must acquire knowledge of how letters correspond to sounds (graphemes and phonemes) in order to conduct phonological recoding, through identifying sounds contained within a target word and blending these together for pronunciation (Share, 1995). With the results of Chapter 4 demonstrating that poor phoneme awareness was associated with poor GPC learning and generalisation, as early readers were not able to identify the target GPC sounds within training words to be "self-taught".

Chapter 2 also present results which correspond with Share's Self-Teaching Hypothesis (1995) as a developmental account of learning to read. As reading performance within the post-phonics ALP group displayed an initial nonword advantage at Year 1, which Share's hypothesis (1995) would attribute to a focus on phonological recoding as part of their systematic, synthetic phonics instruction within the first two years of primary school. As the fundamental elements of UK government approved phonics programmes are that GPCs are taught to early readers, combined with a reading approach which utilises identifying letter-sounds within a target word and blending these together to produce the sound of the whole word. As reading develops according to the Self-Teaching Hypothesis (Share, 1995), early readers begin orthographic learning through exposure and experience with wider texts, read through this phonological recoding process. This ultimately leads to whole word reading, "by-sight" instead of segmenting a word into its constituent sounds as orthographic representations are created, which was apparent at Year 4 within the ALP sample, who had an advantage for exception word reading compared to nonword reading. At this point in their reading development, the ALP sample had created orthographic representations to read exception word items as whole-words, corresponding with Share's (1995) hypothesis. However, nonword reading could still occur at Year 4, as phonological recoding is required for reading of novel words without an existing orthographic representation.

Regarding literacy-related skills and their relation to reading, the results of Chapter 4 have some theoretical implications. Firstly, Chapter 4 demonstrated that phoneme awareness contributed to GPC "self-teaching", and this ability within early readers was hindered if phoneme awareness was poor, whilst the opposite was found for average phoneme awareness and above. Contributing to the existing literature which demonstrates a relationship between phoneme awareness and reading performance (Ball & Blachman, 1991; Elbro & Petersen, 2004; Hatcher et al., 2002; Melby-Lervåg et al., 2012). Which can be explained in relation to GPC "self-teaching" through the notion that early readers could not detect and learn target GPCs, if they did not understand how letters and sounds are associated, resulting in poor phoneme awareness skills and poor GPC “self-teaching”. Secondly, Chapter 4 demonstrated that letter-sound knowledge contributed to GPC “self-teaching” and this was separate from the influence of phoneme awareness alone. Supporting literature which indicates that the literacy-related skills of letter-sound knowledge and phoneme awareness, while potentially related, contribute separately to reading (Blaiklock, 2004; Clayton et al., 2020; Hulme et al, 2005).

In accordance with the alphabetic principle (Bryne, 1998, as cited in Hulme et al., 2012), reading development requires both phoneme awareness and letter-sound knowledge, both of which were found to significantly contribute to GPC “self-teaching” in Chapter 4. Providing overall support for the proposal of the alphabetic principle and its account of how both factors contribute independently to the reading development of early readers, in this case, through facilitating GPC “self-teaching” and phonological recoding. However, Chapter 4 did not establish if these two factors were related to each other, as existing literature indicates that phoneme awareness may predict letter-sound knowledge (Foy & Mann, 2006; Huang et al., 2014).

Vocabulary was also found to significantly contribute to GPC “self-teaching” in Chapter 4, with early readers with high vocabulary scores at pre-test, scoring highly at post-test on both trained and novel items, while readers with poor vocabulary scores at pre-test scored the lowest at post-test, especially for novel items. Theoretically, this could have two explanations, however it is beyond the scope of the training study within Chapter 4 to conclude which occurred within GPC “self-teaching”. These potential theoretical explanations and future research which could be conducted to determine the cause of this vocabulary effect on GPC “self-teaching” are discussed in Section 5.4 below but to briefly summarise, poor vocabulary growth may cause a lack of phoneme awareness growth, per the Lexical Restructuring Model (Walley et al., 2003) resulting in poor phoneme awareness, which has implications for GPC “self-teaching” as discussed above. Alternatively, a poor vocabulary score may indicate phonological processing difficulties and as a result, difficulties with GPC “self-teaching” and accessing text independently to widen vocabulary. Additionally, due to the design of the training study in Chapter 4, the role of vocabulary as providing top-down information may have been minimised. According to Share (1995), vocabulary provides top-down information for words which are partially recoded, however, all of the items within the study were designed to be decodable to facilitate GPC “self-teaching”. Therefore, this theoretical contribution of top-down information accessed through vocabulary was not examined.

Interestingly, home-literacy environment did not significantly contribute to GPC “self-teaching” in the results of Chapter 4. However as suggested by DeBaryshe et al. (2000), a varied home literacy environment may provide learning opportunities for an early reader such as familiarity with a range of texts and to independently explore literacy, which may facilitate independent reading and reading development. Therefore, it is difficult to conclude that home-literacy environment does not contribute to GPC "self-teaching" indirectly.

Overall, these studies provide support for the models of reading discussed throughout the thesis, through demonstrating that early readers utilise a dual-route approach to reading, which is impacted by systematic, synthetic phonics teaching. In regard to each model of reading, the SVR (Gough & Tunmer, 1986) proposes that decoding is essential for reading comprehension when combined with linguistic comprehension. This decoding ability should theoretically be enhanced by phonics teaching, enhancing fundamental skills for decoding, such as phoneme awareness and GPC knowledge. The results within Study 1a can be interpreted as reflecting this reading advantage for children taught through systematic, synthetic phonics teaching, which improved their decoding ability as a fundamental skill for reading and speculatively, as a contributing factor to reading comprehension, although this was not directly measured.

Share's Self-Teaching Hypothesis (1995) also suggests that phonological decoding forms the basis of early reading and facilitates reading development, through providing early readers with the ability to "self-teach" through independent experience with text. Share's Self-Teaching Hypothesis (1995) is the only developmental approach to reading, compared to the SVR (Gough & Tunmer, 1986), DRC (Coltheart et al., 2001) and Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989), which proposes that over time and independent experience with text, the Lexical pathways to reading develop through the lexicalisation process. The results from Study 1a can be interpreted as reflecting this developmental process of reading, through a dependence on decoding ability within early reading, which is facilitated through systematic, synthetic phonics teaching. These readers displayed an initial nonword reading advantage which relied on phonological decoding alone and out-performed a sample of pre-phonics readers when comparing reading performance of the same word types. Longitudinally, this post-phonics sample displayed the reading development pattern predicted by Share (1995), through moving from an initial reliance on decoding and nonwords, to a whole-word reading strategy along the Lexical pathways to reading, reflected by an advantage for exception words which cannot be successfully read through decoding alone. Results from Study 1b can also be interpreted as reflecting this proposed reading development. Initially, early readers display more profiles of reading difficulty with exception words, due a lack of developed Lexical pathways to reading, which longitudinally decreases as this pathway develops through lexicalisation. In contrast, early readers who have received systematic, synthetic, phonics teaching display few nonword profiles of reading difficulty, due to the initial reliance on the Non-lexical pathway which is supported by phonics teaching.

This Self-Teaching Hypothesis (Share, 1995) was extended by the results of Study 2, which indicated that early readers can "self-teach" themselves GPCs following systematic, synthetic, phonics teaching experience to facilitate this transitional period of reading development. These results can once again be interpreted as reflecting the theoretical impact of systematic, synthetic, phonics teaching. Specifically, decoding ability is fundamental to early reading and if decoding is improved through phonics teaching, then reading development and improvement amongst the associated skills of phoneme awareness and letter-sound knowledge are facilitated. Study 2 demonstrates that Share's (1995) original hypothesis can be extended to include a transitional period within "self-teaching", whereby early readers with some phonics experience can detect, learn and generalise novel GPCs without explicit instruction. These findings and this extension of the original hypothesis also support the existing literature which suggests that dual-route computational models (Pritchard et al., 2016) and early readers (Apfelbaum et al., 2013) can detect and learn novel GPCs from whole word exposure, although speculatively as suggested by Pritchard et al. (2016), phonics teaching may facilitate this ability.

The DRC (Coltheart et al., 2001) and the Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) both have a dual-route reading approach although they differ in their conceptualisation. The different word type reading abilities and associated profiles of reading difficulty within Studies 1a and 1b are interpreted as providing evidence for these dual-route approaches to reading, as they reflect that these word types require distinct underlying processes and routes to reading. These parallel routes to reading are specialised for different word types, as discussed in Section 1.2.2, with nonwords read along the Non-lexical route through phonological decoding, and exception words read along the Lexical routes through whole-word reading, to avoid regularization errors. Word type reading results within Study 1a and profiles of reading difficulties within Study 1b reflect these separate processes, as word type reading abilities varied between pre- and post-phonics groups and longitudinally as part of the reading development process outlined above by Share (1995). If reading occurred along one continuous route, word type reading, and emerging profiles of reading difficulty would not be expected to differ. Additionally, systematic, synthetic phonics would improve all word reading difficulties, which contrasts with the findings of Study 1b, whereby exception and mixed profiles of reading difficulty displayed longitudinal stability, albeit moderate to fair.

These Study 1b results of longitudinal profiles of reading difficulty within a sample of mainstream poor readers who have received systematic, synthetic phonics instruction are interpreted as providing support for the existing literature regarding tailored interventions for reading difficulties. Specifically, interventions should be designed to address the underlying processes and skills associated with each type of difficulty (phonological or orthographic), (Fiorello et al., 2006; Gustafson et al., 2007; Rose, 2009). Whilst systematic, synthetic phonics may have addressed phonological difficulties for some of these early readers, it is not effective for those who displayed longitudinal exception and mixed profiles of reading difficulty. As part of this creation of tailored interventions, further research is required to determine the underlying cause of exception word reading difficulties, if these are not based within phonological difficulties addressed through systematic, synthetic phonics teaching. Additionally, the literature regarding longitudinal profiles of reading difficulty from samples of dyslexic and mainstream poor readers is only partially supported by the results of Study 1b. Exception profiles of reading difficulty were found to partially reduce with time, interpreted as reflecting longitudinal reading development and replicating similar results that this reading difficulty is due to developmental delay and thus improves with time (Manis et al., 1996; Stanovich et al., 1997; Talcott et al., 2013; Wolff, 2009). Nevertheless, this type of reading difficulty profile did display fair to moderate longitudinal stability, indicating that this word reading difficulty is not due to delayed development alone, akin to the results of Peterson et al. (2013).

Study 2 demonstrated that the literacy-related skills of phoneme awareness, letter-sound knowledge and vocabulary contribute to GPC "self-teaching" ability of early readers. These results can be interpreted as confirming the requirement of early readers to master the alphabetic principle (Bryne, 1998, as cited in Hulme et al., 2012); consisting of phoneme awareness and letter knowledge skills, as a pre-requisite for phonological decoding. This phonological decoding ability combined with these underlying skills facilitate detecting novel GPCs within whole words for GPC "self-teaching". Therefore, providing further evidence that phoneme awareness and letter-sound knowledge form the fundamental skills for early reading, through facilitating phonological decoding. The role of vocabulary is interpreted as part of the reciprocal relationship between vocabulary and phoneme awareness, as suggested by the Lexical Restructuring Model (Walley et al., 2003).

5.2.1 Theoretical reading development

This thesis contributes original knowledge to the field of child reading development and reading difficulties, through providing research findings on the potential effects of systematic, synthetic, phonics teaching within mainstream primary schools in England. Specifically, how this increased focus on phonics teaching may have impacted word type reading, emerging profiles of reading difficulties and providing additional benefits to early readers, to facilitate "self-teaching" and independent reading. When the results from the three empirical studies outlined above are considered together, they inform us the following key points, in relation to the field of child reading development and reading difficulties. It is important to note that this is a speculative account of reading development, based on existing cognitive models and developmental hypotheses of reading, combined with insights from the empirical results within this thesis.

There is support for the dual-route theory of reading and reading development, outlined developmentally by Share (1995) as part of the Self-Teaching Hypothesis and as part of the DRC

(Coltheart et al., 2001) and Triangle Model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989). Developmentally, early readers display an initial reliance on the Non-lexical pathway to reading and display difficulties with word types which require the Lexical pathways to reading, as these are not yet developed. Over time, as Lexical pathways develop, performance with non-decodable word types improve, however this ability to decode words along the Non-lexical route does not decline. These findings provide further support for the developmental process of these two routes to reading, outlined by Share (1995) and the specialisation of each route to reading in relation to word types, outlined by the DRC (Coltheart et al., 2001).

These findings also indicate that early reading ability is influenced by the way in which reading is taught, as early readers taught through systematic, synthetic phonics, show early reading abilities associated with phonics teaching. Specifically, early readers taught through synthetic phonics display an advantage for word types which rely on phonological decoding rather than whole-word reading, as this ability and associated skills of phoneme awareness and GPC knowledge are solidified during phonics teaching. The process of reading development reflected within the longitudinal results of this thesis has implications for emerging reading difficulties, which displayed fair to moderate longitudinal stability. It appears that emerging reading difficulties amongst early readers are dynamic and may partially reflect this natural reading development process, whereby a proportion of word type reading difficulties are addressed over time with reading experience. In contrast, the profiles of reading difficulty which did display longitudinal stability are not addressed through phonics teaching alone or through developmental delay as they do not improve with time. Therefore, emerging reading difficulties within the later years of primary school may require tailored support to address the underlying processes associated with each route to reading.

Literacy-related skills are also important for early reading success, as proposed by the SVR (Gough & Tunmer, 1986) and Share (1995); phonological decoding forms the basis for successful early reading and facilitating reading development, through this "self-teaching" process. Thus, justifying the basis for synthetic phonics teaching, to improve phonological decoding skills as a core aspect of reading development. Additional literacy-related skills of phoneme awareness, letter-sound knowledge and vocabulary were also found to contribute to GPC "self-teaching" ability of early readers and thus, contributing to reading development through facilitating independent reading. Therefore, providing support for the alphabetic principle of reading (Bryne, 1998 as cited in Hulme et al., 2012) and its contribution to reading development. While the role of vocabulary was only confirmed in relation to GPC "self-teaching" ability, there is a possibility of a reciprocal relationship between phoneme awareness and vocabulary through the Lexical Restructuring Model (Walley et al., 2003). Alternatively, vocabulary may contribute to GPC "self-teaching" through providing top-down knowledge of the phonological forms of training items, therefore facilitating novel GPC detection.

These literacy-related skills also contribute to childhood reading development as part of establishing phonological decoding ability, which requires phoneme awareness and GPC knowledge. This phonological decoding is then utilised as part of the Non-lexical route to reading. This development of phonological decoding then facilitates "self-teaching" through independent reading experience, leading to the development of the Lexical routes to reading, through lexicalisation and growing orthographic knowledge (Share, 1995). Ultimately leading to orthographic representations of whole-words, to facilitate reading "by sight", especially for word types such as exception words which cannot be read accurately through phonological decoding alone.

5.3 Educational implications

The results presented in this thesis have potential implications for educational practice and policy regarding how reading is taught within English primary schools. Firstly, this thesis and the results from the three empirical chapters generally provide a positive message for systematic, synthetic phonics teaching. Overall, the results presented have demonstrated that phonics teaching is effectively providing early readers with the skills and knowledge required to read words through phonological decoding and therefore along the Non-lexical route from a dual-route reading perspective. The educational implications arising from this positive phonics message can be separated into three areas.

Firstly, systematic, synthetic phonics teaching within the early years of primary school education has had a positive effect on the reading ability and emerging difficulties of developing readers. In Chapter 2, the results demonstrated that early readers who had received systematic, synthetic phonics teaching, performed better than a pre-phonics standardisation sample on both nonword and exception word reading, in both Year 1 and Year 4 of primary school. Suggesting that overall, the group who had received systematic, synthetic phonics teaching were overall better readers. In Chapter 3, the results reported very few children who were classified as having a Nonword profile of reading difficulty; regardless of whether this was classified using standardised measures or a within-sample classification system. Additionally, there were fewer children classified with a Nonword profile of reading difficulty within the post-phonics sample when compared to a pre-phonics sample. Suggesting that the increased focus on synthetic phonics teaching in preparation for the Phonics Screening Check in its first year, had a positive impact for potential emerging reading difficulties. In Chapter 4, results demonstrated that phonics teaching provided early readers with the knowledge and skills to "self-teach" GPCs through whole-word exposure, without providing additional phonics instruction. This ability was also supported through wider literacy-related skills such as GPC knowledge and phoneme awareness, which are directly taught through phonics teaching. Therefore, there is support from these findings for the positive effects of phonics teaching, which should remain as part of the National Curriculum in order to provide early readers with the necessary tools for early reading. Additionally, phonics teaching provides early readers with the skills to begin independent reading, through this combination of phonological recoding and "self-teaching" ability.

Secondly, while these results have proven that phonics has positive outcomes for reading ability, the results contribute to the existing evidence base that the underlying processes proposed through phonics are working as intended. The results of Chapter 2 demonstrate that following phonics teaching, early readers at Year 1 have a nonword reading advantage compared to exception words. Suggesting that phonics teaching is working to provide early readers with GPC knowledge, phoneme awareness and the ability to phonologically recode novel words. The results of Chapter 4 also demonstrate that this phonics teaching provides early readers with the ability to "self-teach" GPCs, which can result in accessing a wide variety of children's texts and facilitating independent reading, along with phonological recoding, ultimately building orthographic representations. However, it is important to note that the successful underlying processes of phonics teaching and this improved Non-lexical, nonword reading, has not come at the expense of wider reading. Chapters 2 and 3 demonstrated that exception word reading performance in the post-phonics sample was average to above average and there were no more Exception and Mixed profiles of difficulty than expected, compared to the pre-phonics norms. While phonics may leave some inconsistencies with regards to exception word teaching, which vary by number and items across the various programmes, it has not created more difficulties than expected with this word type. Therefore, phonics teaching should remain as part of the National Curriculum, although the exception words taught through these programmes could be made consistent, or government guidance could be issued, in order to address difficulties with this word type which cannot be addressed through phonological recoding.

Thirdly, while this increased focus on phonics teaching has had positive outcomes on early reading overall, these results also indicate issues with longitudinal stability and the extent to which reading difficulties are detectable within early primary school. As shown in Chapter 3, the longitudinal stability of reading difficulty profiles amongst early readers is limited. Especially profiles of Nonword reading difficulty, which displayed the lowest stability. Therefore, whilst the Phonics Screening Check may be a valid measure of GPC knowledge and phonic decoding at the time of issue, it may not be able to predict reading difficulties longitudinally. As the limited number of children displaying Nonword reading difficulties at Year 1, did not maintain this profile of reading difficulty until Year 4.

While this lack of stability may demonstrate that phonics teaching is successfully addressing emerging Nonword profiles of difficulty, this was not the case for Exception and Mixed profiles of difficulty, which displayed increased longitudinal stability. Additionally, the longitudinal stability found within these groups may indicate that they are not being detected by the existing reading difficulty measures utilised within these schools, or they may not be receiving intervention to address their reading difficulties. Therefore, perhaps systematic, synthetic, phonics teaching isn't a complete solution for addressing all children with emerging profiles of reading difficulty. It is possible that these children may benefit from additional supplementary teaching which focuses on either increasing oral language competence (e.g., vocabulary) or print exposure (targeting orthography). It is beyond the scope of the evidence in this thesis to judge which of these approaches is more likely to be successful. While the Phonics Screening Check equally cannot capture all early readers with emerging profiles of reading difficulty, although this was not the design of the measure. The potential educational implications for these findings are that more frequent assessments or reviews of the reading ability of early readers, across multiple word types, are required to determine: which children have reading difficulties which cannot be addressed through phonics and which children have longitudinal reading difficulties. Through assessing longitudinal reading difficulties, these assessments or reviews can also determine which children demonstrate longitudinal reading difficulties which may require individual intervention, rather than a reading ability delay which has little longitudinal stability.

5.4 Potential limitations and next steps

Chapters 2 and 3 examine the same two samples of secondary data, one from the pre-phonics DTWRP standardisation sample and the other from the post-phonics ALP longitudinal sample. One limitation of these comparisons between the two samples is that the two samples were not perfectly matched. The ALP sample was larger than the DTWRP sample and while the ALP sample was selected to be representative of the UK population through Birmingham schools, participants were not selected as precisely as the DTWRP standardisation sample. As the DTWRP standardisation sample were explicitly selected to match all key characteristics across a wide geographic area, with requirements related to various factors such as a mixture of urban and rural environments, free school meals (FSM), English as an additional language (EAL) and Special Educational Needs (SEN).

Additionally, the ALP sample was longitudinal which aided in addressing research questions regarding reading performance and reading difficulty profiles over time. In contrast, the DTWRP sample was not longitudinal, which limited the ability to compare the two samples, as it would have been of interest to explore how this pre-phonics sample differed on their word type performance and profiles of reading difficulty over time, compared to the ALP post-phonics sample.

One further limitation was that participant recruitment for the ALP sample asked schools to volunteer to take part in the project. Therefore, an element of these results may be due to volunteer bias, whereby above average performing schools were confident in their pupils' reading ability and volunteered. In contrast, some poorer performing schools may not have agreed to take part, on the basis that it would emphasise their pupils' difficulties in word reading. However, the same limitation could also be applied to the DTWRP standardisation sample.

Within Chapter 3, there are some additional limitations. Firstly, when examining profiles of reading difficulty within the ALP sample, using either a standardised classification sample (DTWRP) or a within-sample classification (z-scores), the number of children displaying reading difficulty profiles with longitudinal stability was limited. One possibility may be that a much larger sample is required to detect stable reading difficulties and this lack of stability was due to the ALP samples limited longitudinal sample size (n = 570). Combined with participant attrition of 37 participants who were classified as having a reading difficulty at Year 1 using the DTWRP classification system, who did not participate at Year 4.

Furthermore, the within-sample classification system of reading difficulty profiles utilised in Chapter 3 may have been too stringent to capture all participants with emerging reading difficulties within the ALP sample. This within-sample classification, informed by the research of Peterson et al. (2013; 2014), allocated profiles of reading difficulty based on z-scores which were equivalent to stanine scores of 2 or below; which was much more stringent than the DTWRP classification system which required stanine scores of 4 or below for a profile of reading difficulty. However, it is important to note that this classification system was able to identify some participants with profiles of reading difficulty and their longitudinal stability, therefore one possibility is that the stringency of this classification system may explain why the percentages of the ALP sample classified as displaying reading difficulties was so low.

Chapter 4 provided interesting findings regarding early reader's ability to learn GPCs from whole-word exposure, however there are limitations that must be acknowledged. Firstly, this study only recruited 126 participants, which while an adequate number of participants to power analyses, a larger sample would have potentially allowed for more variability in the literacy-related skills, such as wider phoneme awareness and home literacy environments. This may in turn have provided more information regarding which literacy-related skills impact implicit GPC learning from whole-word exposure.

Secondly, participants were recruited through voluntary sampling, firstly through parental consent. Therefore, there may be a volunteer bias whereby parents of proficient readers volunteered their children to participate, as they thought they would be able to perform well in the tasks. This may explain why ceiling effects were found across the training sessions for 20% of the sample who scored between 5 and 6 on the GPC Screener, as parents of proficient readers with some GPC knowledge consented for these children to take part. However, this is not entirely the case, as the sample did contain participants with no target GPC knowledge, as shown by the number of children who scored 0 in the GPC Screener (49%). This sample did demonstrate that most participants had some limited target GPC knowledge (51% scoring above 0 on the Screener), while future research should recruit a larger sample of participants with no target GPC knowledge to examine GPC "self-teaching" in more detail. Addressing research questions such as, how much GPC knowledge do participants need to benefit from whole-word GPC training?

There are also some limitations in terms of the study design reported in Chapter 4. Firstly, it is important to acknowledge that the flashcard style of the training sessions, may not reflect how early readers read texts independently. To combat this, the study used the sentence design to attempt to replicate sentences that early readers may encounter in children's texts. Through using the context of sentences, it was hoped that this resembled independent reading more closely than long word lists, which early readers are unlikely to encounter in texts such as children's books. Despite this, it is acknowledged that sentence reading, and the training words used, did not entirely reflect the wide range of texts and words that early readers encounter during naturalistic independent reading. As this was a small-scale training study, these design limitations may be addressed in a wider subsequent study, which may allow for a wider variety of words, sourced from material that participants are currently reading (e.g., classroom texts or home reading assignments).

The three results chapters presented in this thesis have presented some possibilities for future research, addressing outstanding research questions. In Chapter 3, the results demonstrated some stability within Exception and Mixed word profiles of reading difficulty within the post-phonics sample (22% within Exception profiles and 34% within Mixed profiles using the standardised DTWRP classification system). Indicating some longitudinal difficulties when reading exception words which cannot be phonologically recoded, both as a single reading difficulty within the Exception profile and as a combined reading difficulty within the Mixed profile, who may instead represent overall poor readers. Due to the stability found within these profiles, it can be assumed that overall, these exception word reading difficulties were not being addressed between Years 1 and 4 of primary school education.

The difficulty of addressing these longitudinal exception word reading difficulties lies within the inconclusive literature regarding the cause of this word reading difficulty. As discussed in Section 1.4.1, the existing literature and theory regarding reading development proposes multiple explanations for why exception word reading difficulties may develop. These can be classified as either a relative influence on exception word reading difficulties such as a poor home literacy environment, or specific difficulties at child-level such as poor orthographic processing skills. Firstly, one possible explanation is that these children have limited orthographic representations and knowledge of exception words, as a result of a lack of independent reading and exposure to text, such as Griffiths and Snowling (2002). Secondly, these children are not creating orthographic representations for reading along the Lexical routes and instead rely on reading through phonological recoding, which results in regularization errors and exception word reading difficulties. Potentially due to difficulties within orthographic processing, as found by Johnston et al. (2014). Thirdly, these children may have poor vocabulary knowledge, which hinders their ability to use top-down semantic information to pronounce exception words (Steacy et al., 2017; Ricketts et al., 2007). Therefore, future research should be conducted to further investigate the origin of exception word reading difficulties, in order to determine how these difficulties should be addressed to avoid longitudinal reading difficulties.

Additionally, the results of Chapter 3 can conclude that the early synthetic phonics teaching experienced by the ALP sample, did not address these exception word difficulties longitudinally. The literature surrounding reading interventions, including those conducted with phonological and surface dyslexic participants indicate that intervention should be tailored to the child's area of reading difficulty (Gustafson et al., 2007; Fiorello et al., 2006; O'Brien et al., 2012; Rose, 2009). Whilst researchers have applied phonics as an intervention for overall reading difficulties (McArthur et al., 2015), the results from Chapter 3 would suggest that while this did reduce Nonword difficulty profiles, that phonics is not a complete solution to all reading difficulties. Further research is required to investigate not only the cause of these exception word reading difficulties, but also how best to support these readers (i.e., the use of tailored interventions and stability of difficulties post-intervention). As the stability results from Chapter 3 would also suggest, Exception and Mixed word reading difficulties do not display a delay in reading development, as indicated within research with surface dyslexic participants (Manis et al., 1996; Stanovich et al., 1997; Wolff, 2009) and Talcott et al.'s (2013) findings that children with exception word reading difficulties displayed a trajectory of developmental delay. As these profiles of reading difficulty remained at Year 4, three years after the initial classification at Year 1. Instead, these difficulties displayed more longitudinal stability than phonological, nonword reading difficulties.

In Chapter 4, the results demonstrated that early readers are able to learn and apply novel GPCs, following experience with systematic, synthetic phonics teaching, as a form of "self-teaching" in accordance with Share's (1995) hypothesis. However, the question remains, what threshold of phonics instruction does an early reader need to experience in order to be able to "self-teach" GPCs? Including how many GPCs do early readers need to initially learn through this systematic, synthetic phonics in order to facilitate later GPC "self-teaching"? Current phonics teaching aims to teach the 44+ GPCs found within the English language to early readers, in preparation for the Year 1 Phonics Screening Check. Perhaps this task for phonics teaching could be simplified if early readers are able to demonstrate GPC "self-teaching" once a threshold of existing experience with systematic, synthetic phonics and GPC knowledge is achieved. Currently, government approved phonics programmes in use within English primary schools can vary on the number of GPCs taught overall, for example, “Jolly Phonics” teaches 42 letter sounds whilst “Letterland” teaches 44 letter sounds. Therefore, through determining the threshold of phonics experience and knowledge early readers require to facilitate GPC “self-teaching”, recommendations can be made on where the “cut-off” point should be for the number of GPCs taught within phonics programmes, to provide further consistency. This would also help to ensure that phonics programmes are not using a “cut-off” point which may be before the skills and knowledge required for GPC “self-teaching” are consolidated.

Interestingly, the results of Chapter 4 found a training effect with only six exposures to target words across four training sessions. Whereas other sight word training studies such as Stuart et al. (2000) conducted eight to nine training sessions with a low rate of learning amongst participants. Specifically, learning and retention was low after 36 exposures to 16 target words (Stuart et al., 2000). This proposes an additional question for future research which could also be investigated: how many exposures to a whole-word, read by sight with no explicit phonics instruction, do early readers need to conduct GPC “self-teaching”, and what is the optimal amount to ensure accurate GPC “self-teaching”?

Furthermore, Chapter 4 demonstrated that there were no context effects present throughout the training study, as GPCs learnt through nonwords and real words through the context of sentences were similarly detected, learnt and generalised to novel items. However, as the design of the Chapter 4 training study utilised nonword and real words items, which were designed to be decodable through phonological recoding, it was unable to examine if this context effect would be present with exception word items, which could not be phonologically recoded. Share’s Self-Teaching Hypothesis (1995) predicted that the role of context is important for providing top-down information, specifically, for aiding reading when only partial phonological recoding is possible (i.e., exception words). Therefore, a similar training study could be conducted to determine if this context effect only occurs when top-down information is required for exception word reading, compared to a lack of context effect when reading decodable words which do not require top-down information. In order to further examine the predictions made in the Self-Teaching Hypothesis (Share, 1995).

Finally, further investigation into the relationship between vocabulary and phonological recoding ability could be conducted. Chapter 4 demonstrated that vocabulary only significantly contributed to GPC reading at post-test, in an interaction with Word Type (trained or generalisation items). Whereby early readers with lower vocabulary scores achieved higher scores for trained items than generalisation items, while early readers with higher vocabulary scores showed little word type effect at post-test. One possible explanation for this finding is that the early readers with poor vocabulary scores may have poor phonological recoding abilities, which would limit their access to independent reading and experience with text according to Share’s Self-Teaching Hypothesis (1995) and therefore impact on their vocabulary growth through experience with text. Whereby in this explanation, vocabulary is mediated through phonological processing skills (Ouelette, 2006). Therefore, these early readers were not able to detect, learn and generalise the training GPCs to novel items at post-test. However, this explanation may not be the only possibility, as it has been proposed that the relationship between decoding and vocabulary is reciprocal (Verhoeven et al., 2011). Instead, these poor vocabulary scores may reflect reduced vocabulary growth in general, which as a result did not facilitate growth in phoneme awareness and therefore phonological decoding abilities (Ouelette, 2006). As stated within the Lexical Restructuring Model (Walley et al., 2003), as vocabulary develops, phonological representations are revised with increased detail to distinguish them from other phonological representations. Therefore, early readers with poor vocabulary growth may lack this further development of phoneme awareness and phonological recoding skills. Therefore, further investigation into this relationship is required, to determine the underlying difficulties associated with the early readers with poor vocabulary scores and their ability to "self-teach" GPCs.

5.5 Conclusions

This thesis has examined the impact of an increased emphasis on systematic, synthetic phonics teaching on the reading performance, emerging reading difficulties and "self-teaching" ability of early readers within primary schools in England. The results reported in this thesis contribute to the existing literature regarding childhood reading and reading development, especially literature regarding the impact of phonics instruction. Specifically, this thesis provides new evidence supporting the conclusion that the increased emphasis on systematic, synthetic phonics teaching has had a positive outcome for the reading performance and emerging reading difficulties of early readers. Specifically, early readers who have received phonics instruction since the introduction of the 2012 Year 1 Phonics Screening Check, outperformed a pre-phonics sample on word reading across nonwords and exception words longitudinally and displayed overall less profiles of reading difficulty longitudinally when classified with the same measures. Additionally, early readers displayed a specific, initial nonword reading advantage compared to exception words and displayed less profiles of nonword reading difficulty longitudinally. Interestingly, this nonword reading advantage did not result in poor exception word reading, or more exception word reading difficulties than expected. Therefore, the nonword reading advantage provided through systematic, synthetic phonics teaching, was not at the expense of exception word reading. This systematic, synthetic phonics teaching also appears to provide early readers with the ability to "self-teach" novel GPCs, which can then be generalised to read novel words. However, literacy-related skills such as phoneme awareness, letter-sound knowledge and vocabulary also contribute to this GPC "self-teaching" ability.

However, the results presented in this thesis also raise some unanswered questions which should be addressed through future research. Firstly, exception word reading difficulties and mixed word type reading difficulties displayed some longitudinal stability throughout primary school. These results indicate that these profiles of reading difficulty may not simply reflect a "developmental delay" at this point, as these profiles of difficulty retained some stability over three years. Future research should investigate two questions in relation to this finding. Firstly, the origin of exception word difficulties from a dual-route reading perspective are disputed, with multiple theoretical explanations (e.g., a lack of exposure to print, difficulties within orthographic processing or poor vocabulary knowledge). Therefore, further research is required to determine the origin of exception word reading difficulties. This could then inform the second question to be examined, which is how best to support these children with exception and mixed word difficulties within primary school. Specifically, through further examining the stability of these profiles of reading difficulty and their origin, research could conclude whether these reading difficulties require tailored intervention due to specific difficulties with orthographic processing, or that these difficulties require broader support in terms of increasing exposure to written and spoken language at school and/or at home.

References

Apfelbaum, K. S., Hazeltine, E., & McMurray, B. (2013). Statistical learning in reading: Variability in irrelevant letters helps children learn phonics better. *Developmental Psychology, 49*(7), 1348-1365. https://doi.org/10.1037/a0029839

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26*(1), 49–66. https://doi.org/10.1598/rrq.26.1.3

Berninger, V. W., Vermeulen, K., Abbott, R. D., McCutchen, D., Cotton, S., Cude, J., Dorn, S., & Sharon, T. (2003). Comparison of three approaches to supplementary reading instruction for low-achieving second-grade readers. *Language, Speech, and Hearing Services in School, 34*(2), 101-116. https://doi.org/10.1044/0161-1461(2003/009)

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130*(6), 858–886. https://doi.org/10.1037/0033-2909.130.6.858

Blaiklock, K.E. (2004). The importance of letter knowledge in the relationship between phonological awareness and reading. *Journal of Research in Reading, 27*(1), 36-57. https://doi.org/10.1111/j.1467-9817.2004.00213.x

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., Götz, K., & Hulme, C. (2008). Improving early language and literacy skills: differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry, 49*(4), 422–432. https://doi.org/10.1111/j.1469-7610.2007.01849.x

Burgess, S.R., Hecht, S.A., & Lonigan, C.J. (2002). Relations of the home literacy environment (HLE) to the development of reading-related abilities: A one-year longitudinal study. *Reading Research Quarterly, 37*(4), 408-426. https://doi.org/10.1598/RRQ.37.4.4

Carson, K. L., Bayetto, A. E., & Roberts, A. F. B. (2019). Effectiveness of preschool-wide teacher-implemented phoneme awareness and letter-sound knowledge instruction on code-based school-entry reading readiness. *Communication Disorders Quarterly, 41*(1), 42–53. https://doi.org/10.1177/1525740118789061

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*(2), 149–180. https://doi.org/10.1016/0010-0277(93)90003-E

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91*(1), 77-111. https://doi.org/10.1016/S0010-0277(03)00164-1

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

Chang, Y. N., & Monaghan, P. (2019). Quantity and diversity of preliteracy language exposure both affect literacy development: Evidence from a computational model of reading. *Scientific Studies of Reading, 23*(3), 235–253. https://doi.org/10.1080/10888438.2018.1529177

Chang, Y. N., Taylor, J. S. H., Rastle, K., & Monaghan, P. (2020). The relationships between oral language and reading instruction: Evidence from a computational model of reading. *Cognitive Psychology, 123*, 101336. https://doi.org/10.1016/j.cogpsych.2020.101336

Clayton, F. J., West, G., Sears, C., Hulme, C., & Lervåg, A. (2019). A longitudinal study of early reading development: Letter-sound knowledge, phoneme awareness and RAN, but not letter-sound integration, predict variations in reading development. *Scientific Studies of Reading, 24*(2), 91–107. https://doi.org/10.1080/10888438.2019.1622546

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Coltheart, M. (2005). Modeling reading: The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 6–23). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch1

Coltheart, M. (2006). Acquired dyslexias and the computational modelling of reading. *Cognitive Neuropsychology, 23*(1), 96–109. https://doi.org/10.1080/02643290500202649

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256. https://doi.org/10.1037/0033-295x.108.1.204

Coltheart, M., Tree, J. J., & Saunders, S. J. (2010). Computational modeling of reading in semantic dementia: Comment on Woollams, Lambon Ralph, Plaut, and Patterson (2007). *Psychological Review, 117*(1), 256–272. https://doi.org/10.1037/a0015948

Conners, F. A., Loveall, S. J., Moore, M. S., Hume, L. E., & Maddox, C. D. (2011). An individual differences analysis of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 108*(2), 402–410. https://doi.org/10.1016/j.jecp.2010.09.009

Crawford, C., Dearden, L., & Meghir, C. (2007). *When you are born matters: The impact of date of birth on child cognitive outcomes in England.* The Institute for Fiscal Studies. https://ifs.org.uk/sites/default/files/output_url_files/born_matters_report.pdf

Crawford, C., Dearden, L., & Greaves, E. (2013). *When you are born matters: Evidence for England* (IFS Report R80). The Institute for Fiscal Studies. https://ifs.org.uk/sites/default/files/output_url_files/r80.pdf

Cunningham, A. J., Burgess, A. P., Witton, C., Talcott, J. B., & Shapiro, L. R. (2020). Dynamic relationships between phonological memory and reading: A five year longitudinal study from age 4 to 9. *Developmental Science, 24*(1), e12986. https://doi.org/10.1111/desc.12986

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: examining the role of self-teaching. *Journal of Experimental Child Psychology, 82*(3), 185–199. https://doi.org/10.1016/s0022-0965(02)00008-5

Cunningham, A. J., Witton, C., Talcott, J. B., Burgess, A. P., & Shapiro, L. R. (2015). Deconstructing phonological tasks: The contribution of stimulus and response type to the prediction of early decoding skills. *Cognition, 143*, 178–186. https://doi.org/10.1016/j.cognition.2015.06.013

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

DeBaryshe, B.D. (1995). Maternal belief systems: Linchpin in the home reading process. *Journal of Applied Developmental Psychology, 16*(1), 1-20. https://doi.org/10.1016/0193-3973(95)90013-6

DeBaryshe, B.D., Binder, J.C., & Buell, M.J. (2000). Mother's implicit theories of early literacy instruction: Implications for children's reading and writing. *Early Child Development and Care, 160*(1), 119-131. https://doi.org/10.1080/0030443001600111

Department for Education. (2010, October 1). *Phonics teaching materials: Core criteria and the self-assessment process*. The National Archives. https://webarchive.nationalarchives.gov.uk/ukgwa/20170623173920/https://www.gov.uk/government/publications/phonics-teaching-materials-core-criteria-and-self-assessment

Department for Education. (2013). *English programmes of study: Key stages 1 and 2. National curriculum in England*. (Report No. DFE-00181-2013). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335186/PRIMARY_national_curriculum_-_English_220714.pdf

Department for Education (2014, July 16). *National curriculum in England: English programmes of study*. GOV.UK. https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study/national-curriculum-in-england-english-programmes-of-study

Department for Education. (2022a, January 18). *Choosing a phonics teaching programme*. GOV.UK. https://www.gov.uk/government/publications/choosing-a-phonics-teaching-programme/list-of-phonics-teaching-programmes

Department for Education. (2022b). *The reading framework: Teaching the foundations of literacy*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1050849/Reading_framework_Teaching_the_foundations_of_literacy_-_July_2021_Jan_22_update.pdf

Department for Education. (2022c, January 18). *Validation of systematic synthetic phonics programmes: Supporting documentation*. GOV.UK. https://www.gov.uk/government/publications/phonics-teaching-materials-core-criteria-and-self-assessment/validation-of-systematic-synthetic-phonics-programmes-supporting-documentation

Department for Education and Skills (DfES). (2007). *Letters and Sounds: Principles and practice of high quality phonics*. (Report No. DFES-00281-2007). Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/190599/Letters_and_Sounds_-_DFES-00281-2007.pdf

Double, K. S., McGrane, J. A., Stiff, J. C., & Hopfenbeck, T. N. (2019). The importance of early phonics improvements for predicting later reading comprehension. *British Educational Research Journal, 45*(6), 1220–1234. https://doi.org/10.1002/berj.3559

Duff, F. J., Mengoni, S. E., Bailey, A. M., & Snowling, M. J. (2014). Validity and sensitivity of the phonics screening check: Implications for practice. *Journal of Research in Reading, 38*(2), 109–123. https://doi.org/10.1111/1467-9817.12029

Dyson, H., Best, W., Solity, J., & Hulme, C. (2017). Training mispronunciation correction and word meanings improve children's ability to learn real words. *Scientific Studies of Reading, 21*(5), 392-407. https://doi.org/10.1080/10888438.2017.1315424

Ehri, L. C. (1987). Learning to read and spell words. *Journal of Reading Behavior, 19*(1), 5–31. https://doi.org/10.1080/10862968709547585

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading, 18*(1), 5-21. https://doi.org/10.1080/10888438.2013.819356

Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Reading Research Quarterly, 55*(S1), S45–S60. https://doi.org/10.1002/rrq.334

Ehri, L. C. (2022). What teachers need to know and do to teach letter–sounds, phonemic awareness, word reading, and phonics. *The Reading Teacher, 0*(0), 1–9. https://doi.org/10.1002/trtr.2095

Elbro, C., & Petersen, D.K. (2004). Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia. *Journal of Educational Psychology, 96*(4), 660-670. https://doi.org/10.1037/0022-0663.96.4.660

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191

Fiorello, C.A., Hale, J.B., & Snyder, L.E. (2006). Cognitive hypothesis testing and response to intervention for children with reading problems. *Psychology in the Schools, 43*(8), 835-853. https://doi.org/10.1002/pits.20192

Forum for Research in Literacy and Language. (2012). *Diagnostic test of reading processes (DTWRP) [measurement instrument]*. GL Assessment.

Foy, J.G., & Mann, V. (2003). Home literacy environment and phonological awareness in preschool children: Differential effects for rhyme and phoneme awareness. *Applied Psycholinguistics, 24*(1), 59-88. https://doi.org/10.1017/S0142716403000043

Foy, J. G., & Mann, V. (2006). Changes in letter sound knowledge are associated with development of phonological awareness in pre-school children. *Journal of Research in Reading, 29*(2), 143–161. https://doi.org/10.1111/j.1467-9817.2006.00279.x

Gilchrist, J. M., & Snowling, M. J. (2018). On the validity and sensitivity of the phonics screening check: Erratum and further analysis. *Journal of Research in Reading, 41*(1), 97–105. https://doi.org/10.1111/1467-9817.12095

GL Education Support. (n.d.). *How are standardised test results described?* GL Education. https://support.gl-education.com/knowledge-base/guides/a-guide-to-standardised-tests/a-guide-to-standardised-tests/how-are-standardised-test-results-described/

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

Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition, 123*(2), 280–292. https://doi.org/10.1016/j.cognition.2012.01.003

Griffiths, Y. M., & Snowling, M. J. (2002). Predictors of exception word and nonword reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology, 94*(1), 34–43. https://doi.org/10.1037/0022-0663.94.1.34

Gustafson, S., Ferreira, J., & Rönnberg, J. (2007). Phonological or orthographic training for children with phonological or orthographic decoding deficits. *dyslexia, 13*(3), 211–229. https://doi.org/10.1002/dys.339

Gustafson, S., Samuelsson, C., Johansson, E., & Wallmann, J. (2013). How simple is the simple view of reading? *Scandinavian Journal of Educational Research, 57*(3), 292–308. https://doi.org/10.1080/00313831.2012.656279

Gustafson, S., Samuelsson, S., & Rönnberg, J. (2000). Why do some resist phonological intervention? A Swedish longitudinal study of poor readers in Grade 4. *Scandinavian Journal of Educational Research, 44*(2), 145–162. https://doi.org/10.1080/713696666

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*(3), 662–720. https://doi.org/10.1037/0033-295x.111.3.662

Hatcher, P.J., Hulme, C., & Snowling, M.J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry, 45*(2), 338-358. https://doi.org/10.1111/j.1469-7610.2004.00225.x

Hjetland, H. N., Lervåg, A., Lyster, S. A. H., Hagtvet, B. E., Hulme, C., & Melby-Lervåg, M. (2019). Pathways to reading comprehension: A longitudinal study from 4 to 9 years of age. *Journal of Educational Psychology, 111*(5), 751–763. https://doi.org/10.1037/edu0000321

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*(2), 127–160. https://doi.org/10.1007/bf00401799

Huang, F. L., Tortorelli, L. S., & Invernizzi, M. A. (2014). An investigation of factors associated with letter-sound knowledge at kindergarten entry. *Early Childhood Research Quarterly, 29*(2), 182–192. https://doi.org/10.1016/j.ecresq.2014.02.001

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science, 23*(6), 572–577. https://doi.org/10.1177/0956797611435921

Hulme, C., Caravolas, M., Málková, G., & Brigstocke, S. (2005). Phoneme isolation ability is not simply a consequence of letter-sound knowledge. *Cognition, 97*(1), B1–B11. https://doi.org/10.1016/j.cognition.2005.01.002

Hulme, C., Hatcher, P.J., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology, 82*(1), 2-28. https://doi.org/10.1006/jecp.2002.2670

Hulme, C., & Snowling, M. J. (2013). Learning to read: what we know and what we need to understand better. *Child Development Perspectives, 7*(1), 1–5. https://doi.org/10.1111/cdep.12005

Hulme, C., & Snowling, M. J. (2014). The interface between spoken and written language: developmental disorders. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 369*(1634), 20120395. https://doi.org/10.1098/rstb.2012.0395

Johnston, R., McGeown, S., & Moxon, G. E. (2014). Towards an understanding of how children read and spell irregular words: The role of nonword and orthographic processing skills. *Journal of Research in Reading, 37*(1), 51–64. https://doi.org/10.1111/jrir.12007

Johnston, R. S., McGeown, S., & Watson, J. E. (2012). Long-term effects of synthetic versus analytic phonics teaching on the reading and spelling ability of 10 year old boys and girls. *Reading and Writing, 25*(6), 1365–1384. https://doi.org/10.1007/s11145-011-9323-x

Johnston, R.S., & Watson, J.E. (2004). Accelerating the development of reading, spelling and phonemic awareness skills in initial readers. *Reading and Writing, 17*, 327–357. https://doi.org/10.1023/B:READ.0000032666.66359.62

Kendeou, P., Savage, R., & van den Broek, P. (2009). Revisiting the simple view of reading. *British Journal of Educational Psychology, 79*(2), 353–370. https://doi.org/10.1348/978185408x369020

Kirby, J. R., Desrochers, A., Roth, L., & Lai, S. S. V. (2008). Longitudinal predictors of word reading development. *Canadian Psychology, 49*(2), 103–110. https://doi.org/10.1037/0708-5591.49.2.103

Landi, N., Perfetti, C. A., Bolger, D. J., Dunlap, S., & Foorman, B. R. (2006). The role of discourse context in developing word form representations: A paradoxical relation between reading and learning. *Journal of Experimental Child Psychology, 94*(2), 114-133. https://doi.org/10.1016/j.jecp.2005.12.004

Larsen, L., Kohnen, S., Nickels, L., & McArthur, G. (2015). The letter-sound test (LeST): A reliable and valid comprehensive measure of grapheme-phoneme knowledge. *Australian Journal of Learning Difficulties, 20*(2), 129-142. https://doi.org/10.1080/19404158.2015.1037323

Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry, 51*(5), 612–620. https://doi.org/10.1111/j.1469-7610.2009.02185.x

Lundberg, I., Frost, J., & Petersen, O. P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23*(3), 263–284. https://doi.org/10.1598/rrq.23.3.1

Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Petersen, A. (1996). On the bases of two subtypes of development dyslexia. *Cognition, 58*(2), 157–195. https://doi.org/10.1016/0010-0277(95)00679-6

Martin, M. O., Mullis, I. V. S., & Kennedy, A. M. (2007). *PIRLS 2006 technical report*. TIMSS & PIRLS International Study Center, Boston College. https://timssandpirls.bc.edu/pirls2006/tech_rpt.html

McArthur, G., Kohnen, S., Jones, K., Eve, P., Banales, E., Larsen, L., & Castles, A. (2015). Replicability of sight word training and phonics training in poor readers: A randomised controlled trial. *PeerJ, 3*, e922. https://doi.org/10.7717/peerj.922

McArthur, G., Kohnen, S., Larsen, L., Jones, K., Anandakumar, T., Banales, E., & Castles, A. (2013). Getting to grips with the heterogeneity of developmental dyslexia. *Cognitive Neuropsychology, 30*(1), 1–24. https://doi.org/10.1080/02643294.2013.784192

McKague, M., Pratt, C., & Johnston, M. B. (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition, 80*(3), 231–262. https://doi.org/10.1016/s0010-0277(00)00150-5

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*(2), 322–352. https://doi.org/10.1037/a0026744

Miskin, R. (2006). *Read Write Inc. Phonics handbook*. Oxford University Press.

Monaghan, P., Chang, Y. N., Welbourne, S., & Brysbaert, M. (2017). Exploring the relations between word frequency, language exposure, and bilingualism in a computational model of reading. *Journal of Memory and Language*, *93*, 1-21. https://doi.org/10.1016/j.jml.2016.08.003

Mulatti, C., Peressotti, F., & Job, R. (2007). Zeading and reazing: Which is faster? The position of the diverging letter in a pseudoword determines reading time. *The Quarterly Journal of Experimental Psychology, 60*(7), 1005–1014. https://doi.org/10.1080/17470210600847842

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*(5), 665-681. https://doi.org/10.1037/0012-1649.40.5.665

Nation, K. (2017). Nurturing a lexical legacy: reading experience is critical for the development of word reading skill. *Npj Science of Learning, 2*(3). https://doi.org/10.1038/s41539-017-0004-7

Nation, K. (2019). Children's reading difficulties, language, and reflections on the simple view of reading. *Australian Journal of Learning Difficulties, 24*(1), 47–73. https://doi.org/10.1080/19404158.2019.1609272

Nation, K., & Cocksey, J. (2009). The relationship between knowing a word and reading it aloud in children's word reading development. *Journal of Experimental Child Psychology, 103*(3), 296–308. https://doi.org/10.1016/j.jecp.2009.03.004

Nickels, L., Biedermann, B., Coltheart, M., Saunders, S., & Tree, J. J. (2008). Computational modelling of phonological dyslexia: How does the DRC model fare? *Cognitive Neuropsychology, 25*(2), 165–193. https://doi.org/10.1080/02643290701514479

O'Brien, B. A., Wolf, M., & Lovett, M. W. (2012). A taxometric investigation of developmental dyslexia subtypes. *dyslexia, 18*(1), 16–39. https://doi.org/10.1002/dys.1431

Olson, R.K., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243-277). Brookes Publishing.

Ouelette, G.P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554-566. https://doi.org/10.1037/0022-0663.98.3.554

Ouelette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing, 23*(2), 189-208. https://doi.org/10.1007/s11145-008-9159-1

Ouellette, G., & Fraser, J. R. (2009). What exactly is a yait anyway: The role of semantics in orthographic learning. *Journal of Experimental Child Psychology, 104*(2), 239–251. https://doi.org/10.1016/j.jecp.2009.05.001

Perera, R., Heneghan, C., & Badenoch, D. (2011). *Statistics Toolkit* (1st ed.). Blackwell Publishing.

Peterson, R. L., Pennington, B. F., & Olson, R. K. (2013). Subtypes of developmental dyslexia: Testing the predictions of the dual-route and connectionist frameworks. *Cognition, 126*(1), 20–38. https://doi.org/10.1016/j.cognition.2012.08.007

Peterson, R. L., Pennington, B. F., Olson, R. K., & Wadsworth, S. J. (2014). Longitudinal stability of phonological and surface subtypes of developmental dyslexia. *Scientific Studies of Reading, 18*(5), 347–362. https://doi.org/10.1080/10888438.2014.904870

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56–115. https://doi.org/10.1037/0033-295x.103.1.56

Powell, D., Plaut, D., & Funnell, E. (2006). Does the PMSP connectionist model of single word reading learn to read in the same way as a child? *Journal of Research in Reading, 29*(2), 229–250. https://doi.org/10.1111/j.1467-9817.2006.00300.x

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2016). Modelling the implicit learning of phonological decoding from training on whole-word spellings and pronunciations. *Scientific Studies of Reading, 20*(1), 49-63. https://doi.org/10.1080/10888438.2015.1085384

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2018). A computational model of the self-teaching hypothesis based on the dual-route cascaded model of reading. *Cognitive Science, 42*(3), 722–770. https://doi.org/10.1111/cogs.12571

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25*(2), 482–503. https://doi.org/10.1037/0096-1523.25.2.482

Rayner, K., & Reichle, E. D. (2010). Models of the reading process. *WIREs Cognitive Science, 1*(6), 787–799. https://doi.org/10.1002/wcs.68

Ricketts, J., Bishop, D. V. M., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading, 15*(1), 47–70. https://doi.org/10.1080/10888438.2011.536129

Ricketts, J., Davies, R., Masterson, J., Stuart, M., Duff, F. J. (2016). Evidence for the semantic involvement in regular and exception word reading in emergent readers of English. *Journal of Experimental Child Psychology, 150*, 330-345. https://doi.org/10.1016/j.jecp.2016.05.013

Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading, 11*(3), 235-257. https://doi.org/10.1080/10888430701344306

Rodriguez, E. T., Tamis-LeMonda, C. S., Spellmann, M. E., Pan, B. A., Raikes, H., Lugo-Gil, J., & Luze, G. (2009). The formative role of home literacy experiences across the first three years of life in children from low-income families. *Journal of Applied Developmental Psychology, 30*(6), 677–694. https://doi.org/10.1016/j.appdev.2009.01.003

Romani, C., Di Betta, A. M., Tsouknida, E., & Olson, A. (2008). Lexical and nonlexical processing in developmental dyslexia: A case for different resources and different impairments. *Cognitive Neuropsychology, 25*(6), 798–830. https://doi.org/10.1080/02643290802347183

Rose, J. (2006). *Independent review of the teaching of early reading final report*. UK Department for Education and Skills. https://dera.ioe.ac.uk/5551/2/report.pdf

Rose, J. (2009). *Identifying and teaching children and young people with dyslexia and literacy difficulties*. (Report No. DCSF-00659-2009). DCSF Publications. https://dera.ioe.ac.uk//14790/

Rosen, M. (2021, October 9). *Dear Nadhim Zahawi, the tories vowed to 'eradicate illiteracy' years ago. What went wrong?* The Guardian. https://www.theguardian.com/education/2021/oct/09/dear-nadhim-zahawi-the-tories-vowed-to-eradicate-illiteracy-years-ago-what-went-wrong

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523–568. https://doi.org/10.1037/0033-295X.96.4.523

Shapiro, L. R., & Solity, J. (2016). Differing effects of two synthetic phonics programmes on early reading development. *British Journal of Educational Psychology, 86*(2), 182–203. https://doi.org/10.1111/bjep.12097

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition, 55*(2), 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*(2), 95–129. https://doi.org/10.1006/jecp.1998.2481

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an 'outlier' orthography. *Psychological Bulletin, 134*(4), 584–615. https://doi.org/10.1037/0033-2909.134.4.584

Smith, A. C., Monaghan, P., & Huettig, F. (2021). The effect of orthographic systems on the developing reading system: Typological and computational analyses. *Psychological Review*, *128*(1), 125-159. https://doi.org/10.1037/rev0000257

Snowling, M., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 346*(1315), 21–27. https://doi.org/10.1098/rstb.1994.0124

Snowling, M. J., & Hulme, C. (2012). Interventions for children's language and literacy difficulties. *International Journal of Language & Communication Disorders, 47*(1), 27–34. https://doi.org/10.1111/j.1460-6984.2011.00081.x

Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., & Nation, K. (2009). *YARC York Assessment of Reading for Comprehension Passage Reading [measurement instrument]*. GL Publishers.

Solity, J., & Vousden, J. (2009). Real books vs reading schemes: A new perspective from instructional psychology. *Educational Psychology, 29*(4), 469-511. https://doi.org/10.1080/01443410903103657

Spector, J. E. (2005). Instability of double-deficit subtypes among at-risk first grade readers. *Reading Psychology, 26*(3), 285–312. https://doi.org/10.1080/02702710590967834

Spencer, K. (2007). Predicting children's word-spelling difficulty for common English words from measures of orthographic transparency, phonemic and graphemic length and word frequency. *British Journal of Psychology, 98*(2), 305–338. https://doi.org/10.1348/000712606x123002

Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and Writing, 28*(4), 467-490. https://doi.org/10.1007/s11145-014-9533-0

Stainthorp, R. (2020). A national intervention in teaching phonics: A case study from England. *The Educational and Developmental Psychologist, 37*(2), 114–122. https://doi.org/10.1017/edp.2020.14

Standards and Testing Agency. (2011). *Year 1 phonics screening check. Pilot 2011: Technical report*. (Report No. STA/12/5791). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/368290/phonics_2011_technical_report.pdf

Standards and Testing Agency. (2019). *Phonics screening check: Administration guidance*. (Report No. STA/19/8111/e). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/798235/2019_phonics_screening_check_administration_guidance.pdf

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89*(1), 114–127. https://doi.org/10.1037/0022-0663.89.1.114

Steacy, L. M., Kearns, D. M., Gilbert, J. K., Compton, D. L., Cho, E., Lindstrom, E. R., & Collins, A. A. (2017). Exploring individual differences in irregular word recognition among children with early-emerging and late-emerging word reading difficulty. *Journal of Educational Psychology, 109*(1), 51–69. https://doi.org/10.1037/edu0000113

Steacy, L. M., Kirby, J. R., Parrila, R., & Compton, D. L. (2014). Classification of double deficit groups across time: An analysis of group stability from kindergarten to second grade. *Scientific Studies of Reading, 18*(4), 255–273. https://doi.org/10.1080/10888438.2013.873936

Stuart, M., Masterson, J., & Dixon, M. (2000). Spongelike acquisition of sight vocabulary in beginning readers? *Journal of Research in Reading, 23*(1), 12-27. https://doi.org/10.1111/1467-9817.00099

Stuart, M., & Stainthorp, R. (2015). *Reading Development and Teaching* (1st ed.) SAGE Publications Ltd.

Swan, D., & Goswami, U. (1997). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *Journal of Experimental Child Psychology, 66*(1), 18–41. https://doi.org/10.1006/jecp.1997.2375

Talcott, J. B., Witton, C., & Stein, J. F. (2013). Probing the neurocognitive trajectories of children's reading skills. *Neuropsychologia, 51*(3), 472–481. https://doi.org/10.1016/j.neuropsychologia.2012.11.016

Taylor, J. S. H., Plunkett, K., & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: An artificial orthography paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 60–76. https://doi.org/10.1037/a0020126

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin, 139*(4), 766–791. https://doi.org/10.1037/a0030266

Torgerson, C., Brooks, G., Gascoine, L., & Higgins, S. (2019). Phonics: Reading policy and the evidence of effectiveness from a systematic 'tertiary' review. *Research Papers in Education, 34*(2), 208–238. https://doi.org/10.1080/02671522.2017.1420816

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*. (Report No. RR711). The Department for Education and Skills. https://webarchive.nationalarchives.gov.uk/ukgwa/20130323020129/https:/www.education.gov.uk/publications/eOrderingDownload/RR711_.pdf

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34*(1), 33–58. https://doi.org/10.1177/002221940103400104

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

van Bergen, E., van Zuijen, T., Bishop, D., & de Jong, P. F. (2016). Why are home literacy environments and children's reading skills associated? What parental skills reveal. *Reading Research Quarterly, 52*(2), 147-160. https://doi.org/10.1002/rrq.160

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67*(6), 1176-1190. https://doi.org/10.1080/17470218.2013.850521

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading, 15*(1), 8–25. https://doi.org/10.1080/10888438.2011.536125

Walker, M., Sainsbury, M., Worth, J., Bamforth, H., & Betts, H. (2015). *Phonics screening check evaluation: Final report*. (Report No. DFE-RR418A). National Foundation for Educational Research. https://www.gov.uk/government/publications/phonics-screening-check-evaluation-final-report

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing, 16*, 5–20. https://doi.org/10.1023/a:1021789804977

Wang, H.-C., Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology, 109*(1), 39–57. https://doi.org/10.1016/j.jecp.2010.11.005

Wang, H.-C., Marinus, E., Nickels, L., & Castles, A. (2014). Tracking orthographic learning in children with different profiles of reading difficulty. *Frontiers in Human Neuroscience, 8*, Article 468. https://doi.org/10.3389/fnhum.2014.00468

Wang, H. -C., Nickels, L., Nation, K., & Castles, A. (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading, 17*(5), 369–384. https://doi.org/10.1080/10888438.2012.749879

Weigel, D. J., Martin, S. S., & Bennett, K. K. (2006). Contributions of the home literacy environment to preschool-aged children's emerging literacy and language skills. *Early Child Development and Care, 176*(3), 357-378. https://doi.org/10.1080/03004430500063747

Wolff, U. (2009). Phonological and surface subtypes among university students with dyslexia. *International Journal of Disability, Development and Education, 56*(1), 73–91. https://doi.org/10.1080/10349120802682083

Woollams, A. M., Lambon Ralph, M. A., Madrid, G., & Patterson, K. E. (2016). Do you read how I read? Systematic individual differences in semantic reliance amongst normal readers. *Frontiers in Psychology*, *7*, 1757. https://doi.org/10.3389/fpsyg.2016.01757

Woollams, A. M., Lambon Ralph, M. A., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review, 114*(2), 316–339. https://doi.org/10.1037/0033-295X.114.2.316

Wyse, D., & Goswami, U. (2008). Synthetic phonics and the teaching of reading. *British Educational Research Journal, 34*(6), 691–710. https://doi.org/10.1080/01411920802268912

Zevin, J. D., & Seidenberg, M. S. (2006). Simulating consistency effects and individual differences in nonword naming: A comparison of current models. *Journal of Memory and Language, 54*(2), 145–160. https://doi.org/10.1016/j.jml.2005.08.002

Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F. X., & Perry, C. (2008). Developmental dyslexia and the dual route model of reading: Simulating individual differences and subtypes. *Cognition, 107*(1), 151–178. https://doi.org/10.1016/j.cognition.2007.09.004

Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology, 12*(3), 413–430. https://doi.org/10.1080/09541440050114570

Appendices

Appendix 1

Study 2 Training Stimuli Word Lists

Table 1

Group 1 Training Stimuli

/OU/ Real Words	/EA/ Nonwords
Sound	Pream
Mouth	Meast
Count	Neach
Snout	Chean
Pouch	Jeach
Trout	Spean
Crouch	Streaf
Grouse	Fleach
Bounce	Keader
Around	Pleaky

Table 2

Group 2 Training Stimuli

/OU/ Real Words	/EA/ Nonwords
Dream	Nound
Beast	Routh
Reach	Sount
Cheat	Snoub
Peach	Mouch
Creak	Brout

Stream	Trouch
Breach	Prouse
Leader	Dounce
Sneaky	Bround

Appendix 2

Study 2 Pre-Test GPC Screener Task Word List

Table 3

/EA/ and /OU/ GPC Screener Word List

High Frequency /OU/ GPC	High Frequency /EA/ GPC
Found	Speak
Round	Least
House	Leave

Appendix 3

Study 2 Home Literacy Questionnaire Adapted from the PIRLS "Leaning to Read Survey" (Martin, Mullis & Kennedy, 2007).

Questions for parents and carers about reading

We want to find out what you think about reading and about the activities you did with your child when they were younger. This will help us understand what kinds of experiences at home are important for children's reading and vocabulary. Please answer all questions you feel comfortable answering. If you do not want to answer a question, just leave it blank. **Please put your completed questionnaire in the pre-paid envelope along with the consent form and post it back to us.**

Your responses to this survey are confidential. We will never share your personal data (e.g., names or contact details) with anyone outside the research team and all information will be processed in accordance with the provisions of the General Data Protection Regulation.

My child's name is:	
Their primary school is:	
Their secondary school will be:	

1 **This survey was completed by** (Circle one option)**:**

Female caregiver (mother, stepmother, grandmother, guardian)	Circle if Yes	Male caregiver (father, stepfather, grandfather, guardian)	Circle if Yes

2 **Before your child began primary school, how often did you or someone else in your home do the following activities with him or her?** Circle one answer for each line

a) Read books	Often	Sometimes	Never (or almost never)
b) Tell stories	Often	Sometimes	Never (or almost never)

c) Talk about what you had read	Often	Sometimes	Never (or almost never)
d) Write letters or words	Often	Sometimes	Never (or almost never)
e) Read aloud signs and labels	Often	Sometimes	Never (or almost never)
f) Play with alphabet toys	Often	Sometimes	Never (or almost never)

3 **In a normal week, how much time do you usually spend reading for yourself <u>at home</u>, including books, magazines, newspapers, and materials for work (in print or on a computer/ tablet)?** Circle one option

a) Less than 1 hour	b) 1-5 hours	c) 6-10 hours	d) More than 10 hours

4 **When you are at home, how often do you read <u>for your own enjoyment</u>?** Circle one option

a) Every day (or almost every day)	c) Once or twice a month
b) Once or twice a week	d) Never or almost never

5 **About <u>how many books</u> are there in your home? (Do not count ebooks, magazines, newspapers, or children's books.)** Circle one option

a) 0-10	b) 11-25	c) 26-100	d) 101-200	e) more than 200

6 **About how many <u>children's books</u> are there in your home? (Do not count children's ebooks, magazines, or school books.)** Circle one option

a) 0-10	b) 11-25	c) 26-100	d) 101-200	e) more than 200

7 **Please say how much you agree with the following statements about reading.** Circle one answer for each line

a) I read only if I have to	Agree a lot	Agree a little	Disagree a little	Disagree a lot
b) I like talking about what I read with people	Agree a lot	Agree a little	Disagree a little	Disagree a lot
c) I like to spend my spare time reading	Agree a lot	Agree a little	Disagree a little	Disagree a lot
d) Reading is an important activity in my home	Agree a lot	Agree a little	Disagree a little	Disagree a lot
e) I would like to have more time for reading	Agree a lot	Agree a little	Disagree a little	Disagree a lot
f) I enjoy reading	Agree a lot	Agree a little	Disagree a little	Disagree a lot

8. **What is the highest level of qualifications completed by the child's mother (or stepmother or female guardian) and father (or stepfather or male guardian)?** This question is important so we know the range of educational backgrounds of parents who have answered our questions.

Mother (circle the highest qualification)	Father (circle the highest qualification)
Higher degree (for example MA, MSc, PhD)	Higher degree (for example MA, MSc, PhD)
Undergraduate degree (for example BA, BSc, B.Ed)	Undergraduate degree (for example BA, BSc, B.Ed)

A Levels or equivalent	A Levels or equivalent
General Certificate of Secondary Education (GCSEs) or equivalent	General Certificate of Secondary Education (GCSEs) or equivalent
Other qualifications (for example vocational qualifications)	Other qualifications (for example vocational qualifications)
Not known	Not known

Many thanks for answering our questions. If you would like us to send you a £5 amazon voucher, and keep you informed about the project (occasional reports on our findings and invites to activities related to this project), please write your email address, mobile phone number, or postal address here. **We will never pass on your contact details to anyone outside the research team and we will only contact you about this project**:

Email/mobile/address:

Appendix 4

Study 2 Post-Test Word Stimuli Lists

Table 4

Group 1 Post-Test Reading Stimuli List (Excluding Trained Stimuli)

/OU/ Real Words	/OU/ Nonwords	/EA/ Real Words	/EA/ Nonwords
Hound	Nound	Dream	Tream
South	Routh	Beast	Reast
Clout	Sount	Peach	Chead
Slouch	Prouse	Stream	Gleach
Ground	Dounce	Bleach	Deader

Table 5

Group 2 Post-Test Reading Stimuli List (Excluding Trained Stimuli)

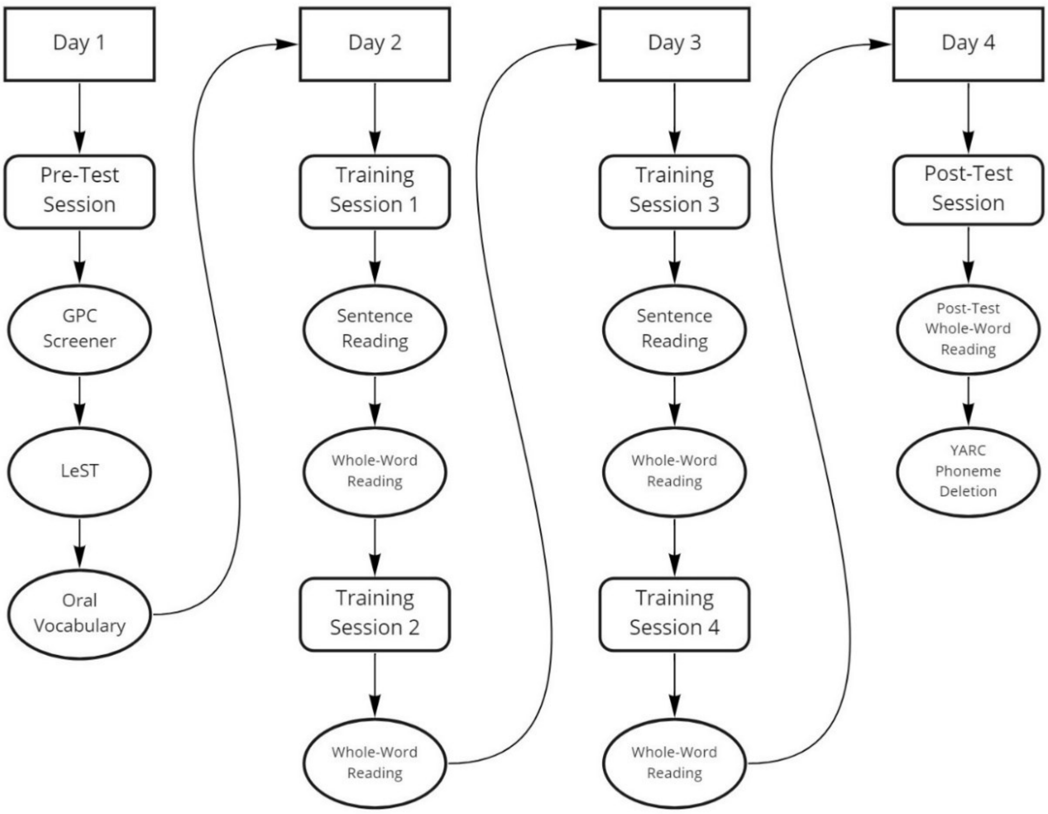
/EA/ Real Words	/EA/ Nonwords	/OU/ Real Words	/OU/ Nonwords
Gleam	Pream	Sound	Dound
Feast	Meast	Mouth	Pouth
Leach	Chean	Trout	Hount
Scream	Fleach	Crouch	Krouse
Preach	Keader	Around	Tounce

Appendix 5

Study 2 Procedure and Daily Test Schedule

Figure 1

Study Procedure and Task Order



Appendix 6

Study 2 Training Session Whole-Word Sentence Reading Stimuli

Table 7

Group 1 Real Word Training Sentences

1. The enchanted harp made a wonderful **sound.**
2. The horse had a carrot in its **mouth.**
3. To play hide and seek, first you have to **count.**
4. The pig had mud on its **snout.**
5. Kangaroos can carry babies in their **pouch.**
6. Sarah went fishing and caught a **trout.**
7. When you throw a ball, it may **bounce.**
8. Henry saw the fallen tree and had to walk **around.**
9. To hide from the goblin, they would have to **crouch.**
10. In the field, they saw a **grouse.**

Table 8

Group 1 Nonword Training Sentences

1. The treasure map led to a place called **pream.**
2. The alien lived on the planet called **meast.**
3. The name on the dog collar said **neach.**
4. George feared the ogre whose name was **chean.**
5. The comet flew past the spaceship called **jeach.**
6. Only cats can see the colour **spean.**
7. The dragon was eating a handful of **fleach.**

8. Princess Alice had shoes made of **streaf.**

9. The trees in the forest were covered in **keader.**

10. The weather was awfully **pleaky.**

Table 9

Group 2 Real Word Training Sentences

1. When you fall asleep, you often **dream.**

2. They went to the woods to look for the **beast.**

3. The book on the shelf was just out of **reach.**

4. Judy wanted to win, but she did not want to **cheat.**

5. The squirrel took a bite of the **peach.**

6. Upstairs Hannah heard a **creak.**

7. The bear was drinking from the **stream.**

8. John wanted to be the **leader.**

9. To clean the shirt, they would have to use **bleach.**

10. Everyone thought that the fox was **sneaky.**

Table 10

Group 2 Nonword Training Sentences

1. The elf lived in a house made of **nound.**

2. Through the telescope they saw the star called **routh.**

3. The name of the dragon was **sount.**

4. Dennis was building a boat for his friend **snoub.**

5. To get to the treasure, they would have to defeat the **mouch.**

6. The prince brought gifts of gold and **brout.**

7. The lemur was eating a fruit called a **trouch.**

8. Only Penny could speak to the mermaid called **prouse.**

9. The unicorn's name was **dounce.**

10. At the top of the beanstalk was the castle of **bround.**

Appendix 7

Word Stimuli Lists for the Phonological and Orthographic Choice Tasks within the ALP data collection.

Table 11

Phonological Choice Stimuli

Left	Right
nite	*kile*
beal	*bair*
glew	*plue*
slod	*steem*
poal	hoil
kape	dape
gass	hask
sharf	skore
gizz	duzz
grait	throut
blad	flud
chande	danse
cleen	vown
droom	creem
skait	plout
shoop	sleap
pellar	senter
strook	speek
sach	trax
teer	gair
cree	flie
sirst	bight
cloor	fleer
phan	shog
stee	floo
nule	rale
shurt	skart
fout	lait
fone	phote
phinny	munny
flate	bote
naim	soom

peneral	jenerous
reaton	seazon

Table 12

Orthographic Choice Stimuli

Left	Right
rume	*room*
young	*yung*
tertle	*turtle*
snoe	*snow*
between	betwean
lake	laik
sammon	salmon
fought	faught
grone	grown
perched	purched
wheet	wheat
trousers	trowsers
condence	condense
compliment	complimant
dignaty	dignity
pavement	pavemant
nusance	nuisance
resource	resourse
travle	travel
baisment	basement
assure	ashure
captain	captin
mysterey	mystery
several	sevral
distence	distance
backwords	backwards
explane	explain
demon	deamon
harth	hearth
wreath	reath
applause	aplause
sensitive	sensative
liberty	libberty

culpret	culprit