University of Reading

The vocabulary challenge of the English Language GCSE Exam and the implications for adolescent reading experience

A thesis submitted for the degree of Doctor of Philosophy in Education

Institute of Education

Beverley Judith Jennings

November 2024

Declaration of Original Authorship and Statement of Contribution

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Paper 1 (Chapter 4) – I sourced and collected the data from the exam boards to create the corpus. I formatted and uploaded all the texts to the corpus platform, *Sketch Engine*. I conducted the analysis within *Sketch Engine,* identifying keywords, searching for them in a reference corpus, and I also ran the corpora comparisons. I was first and corresponding author on the paper: *A Corpus Study of English Language Exam Texts: Vocabulary Difficulty and the Impact on Students' Wider Reading (or Should Students be Reading More Texts by Dead White Men?).*

Paper 2 (Chapter 5) – I designed and created the survey. I contacted schools and colleges and sent out information sheets. I circulated the survey, transferred the results to SPSS, and conducted the descriptive and statistical analysis. I was first and corresponding author on the paper: *Are some types of reading more equal than others? Adolescent reading experience and the requirements of high-stakes assessments.*

Paper 3 (Chapter 6) – I sent out information sheets to participants and collected the data from students and from the online school platform. I carried out the uploading and formatting, creating two more corpora in *Sketch Engine*. I conducted the frequency analysis and also the corpora comparisons. I was first author and will be the corresponding author on the paper: *Adolescent reading experience: independent choices and curriculum materials, redefining the deficit narrative.*

Beverley Jennings

Acknowledgements

I would like to thank my supervisors, Professor Holly Joseph and Associate Professor Daisy Powell, for all their support over the duration of my PhD. In every interaction, since my initial enquiry email back in March 2017 through to the final submission of this thesis, they have always been positive, encouraging, and constructive. They have been generous with their time, with opportunities and have also shown genuine interest and pleasure in my progress and achievements. I will miss their guidance and their company.

I would also like to thank Professor Sylvia Jaworska for her generous support and time as a consultant for the corpus linguistics parts of this research project.

This was not a funded PhD, so I have no official funders to thank or acknowledge. However, not everyone is fortunate enough to be able to study without funding and I wanted to acknowledge my privilege in being able to do so. I also want to thank my parents, Derek and Eileen Monds, for their generous contributions towards my fees. I would also like to thank my employers for facilitating the part-time employment that gave me the time and the money to complete this PhD.

I would like to thank the educational settings, teachers, and students who contributed to the data in this study. Everyone was very generous and supportive and none of what follows would have been possible without their input.

I have two PhD friendships to celebrate. In no particular order: Fatima Zhara Abbou for being the best study-buddy, holding me to account for my progress and supporting me throughout my journey; and, Cecilia Muldoon, who has watched, listened to and read absolutely everything I have done and was/is my biggest cheerleader.

Finally, a thank you to my family, Ian, Abigail and Kenton, for putting up with me spending so much of my time on this over the last six years, for never complaining and for always celebrating my progress.

Abstract

A new English language GCSE exam in England has put an increased focus on adolescents' comprehension of unseen literary texts, from all three of the 19th, 20th and 21st centuries. Since reading experience is a strong predictor of successful reading comprehension, it is important to know whether what adolescents read prepares them for this new exam. Any findings related to this particular exam could also have broader implications for other assessment jurisdictions and for theory. This PhD project focuses on: 1) the vocabulary challenge presented by the new exam texts; 2) the type of reading material that provides experience with this vocabulary; and 3) the actual reading experience of adolescents, including the vocabulary they encounter. First, a corpus of exam texts was created and analysed to examine typical vocabulary (Paper 1). Next, a survey reported respondents' attainment in the English language GCSE exam and their reading experience (Paper 2). Finally, two further corpora, created from samples of students' independent and curriculum reading materials, were analysed (Paper 3). The keywords in the exam texts were found to be typically low frequency and most likely to be encountered in older, literary texts. The reading survey showed that students who gained high grades also had more exposure to classic authors. The final corpus analysis showed that students' independent reading for pleasure was a better match for the exam texts than their curriculum reading, although the vocabulary in the curriculum reading was more challenging. This study contributes new primary data in the three new corpora and in the data on adolescent reading habits from the survey. It shows that reading experience can be analysed and explored through the use of corpus linguistics and through a genre-specific ART. Finally, a potential influence of assessment content on curriculum choices and on reading practices is identified.

Table of Contents

List of Tables i
List of Figures iii
List of Appendices iv
List of Acronyms v
Thesis Structure vi
Impact of COVID-19 vii
Prologue ix
Chapter 1. Introduction 1
1.1 Background 2
1.2 The Research Gap 5
1.3 Research Questions 6
Paper 1 6
Paper 2: 6
Paper 3: 7
1.4 Overview of the Study 7
Chapter 2. Literature Review 9
2.1 Components of Successful Reading 11
2.2 Intervention Studies for Improving Reading 21
2.3 Lexical Quality Hypothesis 30
2.4 Lexical Legacy Hypothesis 34
2.5 Reading Experience 37
2.6 Reading Habits and Practices 43
2.7 This Study 46
Chapter 3. Methodology 49
3.1 Corpus Linguistics 51

3.1.1 Corpus Studies of Academic Language 53
3.1.2 Corpus Studies of Non-Academic Language 55
3.1.3 Creation of New Corpora 55
3.1.4 Keywords 57
3.1.5 Corpora Comparisons 59
3.2 Reading Survey 59
3.2.1 Format and Structure 60
3.2.2. Self-reported Reading Measures 63
3.2.3 The ART 65
3.2.4 The Vocabulary Test 66
3.3 Summary 68
Chapter 4. Paper 1 A Corpus Study of English Language Exam Texts: Vocabulary Difficulty and the Impact on Students' Wider Reading (or Should Students be Reading More Texts by Dead White Men?) 70
Abstract 71
4.1 Introduction 71
4.1.1 Background 72
4.1.2 Vocabulary Knowledge and Reading Experience 74
4.1.3 Corpus Linguistics 75
4.1.4 This Study 76
4.2 Method 77
4.2.1 Data 77
4.2.2 Keywords 79
4.2.3 Comparisons with Other Corpora 80

4.2.4 Genre Sources for Keywords 81

4.3 Findings 82

4.3.1 Keywords 82

4.3.2 Comparison with Reference Corpora 85

4.3.3 Genre Sources for Keywords 86

4.4 Discussion 88

4.5 Further research and limitations 90

4.6 Conclusion 91

Chapter 5. Paper 2 Are some types of reading more equal than others? Adolescent reading experience and the requirements of high-stakes assessments 92

Abstract 93

Implications for Practice 94

5.1 Introduction 95

5.1.1 Research Aims of this Study 99

5.2 Method 100

5.2.1 Ethics 100

5.2.2 Participants 100

5.2.3 Impact of COVID-19 102

5.2.4 The Survey 104

5.3 Results 108

5.3.1 Descriptive Statistics 108

5.3.2 Analysis of Variance 111

5.4 Discussion 112

5.4.1 Reading Experience and Attainment 112

5.4.2 Implications for Schools and Curricula 114

5.5 Limitations and Future Research 116

5.6 Conclusion 117

Chapter 6. Paper 3 Adolescent Reading Experience, Independent Choices and Curriculum Materials: Redefining the Deficit Narrative 118

Abstract 119

6.1 Introduction 119

6.2 Literature review 120

6.2.1 Reading Experience 120

6.2.2 Curriculum reading 122

6.2.3 Independent Reading 123

6.2.4 The English Language GCSE 124

6.2.5 This Study 126

6.2.6 Research Questions: 127

6.3 Methodology 128

6.3.1 Ethical Approval 128

6.3.2 Independent Reading 128

6.3.3 Curriculum Reading 129

6.3.4 Creation of the Corpora 131

6.3.5 Frequencies of Parts of Speech and Lexical Diversity 132

6.3.6 Most Frequent Words 133

6.3.7 Corpora Comparisons 134

6.4 Findings 134

6.4.1 Parts of Speech in the Corpora 134
6.4.2 Lexical Diversity 136
6.4.3 Comparing the Most Frequent Words in the Different Parts of Speech in the Two Corpora 136
6.4.4 Corpora Comparisons 139
6.5 Discussion 140
6.6 Limitations and Further Study 145
6.7 Conclusion 146
Chapter 7. Discussion 148
7.1 Contribution 149
7.2 Overview of Results 150
7.3 Limitations 153
7.3.1 Methodological Limitations 153
7.3.2 Theoretical Limitations 158
7.4 Implications 159
7.4.1 Implications for Future Research 160
7.4.2 Implications for Theory 166
7.4.3 Implication for Practice 170
7.5 Conclusion 177
Reference List 180
Appendices 212

List of Tables

Table 1 Document Distribution in the Exam Text Corpus: 19th, 20th and 21st Century and Fiction and Non-fiction Sub-corpora 77

Table 2 Words, Unique Words and Documents in the Exam Text Corpus and Sub-corpora . 78

Table 3 Exam texts by David Lee Genre Categories 79

Table 4 Keywords with Scores (Ordered by Keyword Score) 83

Table 5 Comparisons of Reference Corpora to Exam Text Corpus: the Lower the Score the More Alike the Corpus is to the Exam Text Corpus 86

Table 6 Frequencies and Frequencies Per Million of the Exam Text Corpus Keywords in the British National Corpus 87

Table 7 Demographic Information for Respondents Who Started and Completed the Survey Alongside National Statistics 103

Table 8 Raw Numbers and Percentages of Students Who Obtain Each Possible Grade in English Language GCSE, for Those Who Completed the Survey, Those Who Started the Survey, and Those Who Took the Exam Nationally in 2019 (top) and 2020 (bottom) 104

Table 9 Spearman Rank Order Correlations Between Author Recognition Test Scores, Self-Reported Reading Measures and Vocabulary Scores 108

Table 10 Results from Self-Report Questions about Genres that Respondents Read 110

Table 11 Mean Scores, Standard Deviations (in brackets), and Number of Participants for the Author Recognition Tests and Vocabulary Test by Grade Group and Year 111

Table 12 Analysis of Variance Summary Table for Classic Author Recognition Test Score and Young Adult Author Recognition Test Score 111

Table 13 Analysis of Variance Summary Table for Vocabulary Score 112

Table 14 Texts used to create Independent Reading Corpus 129

Table 15 Curriculum Subjects in the Curriculum Reading Corpus with Word Counts 131

Table 16 Corpora Contents..... 132

Table 17 Raw Occurrences and Frequency per Million for Parts of Speech in the Corpora 135

Table 18 Comparison of Independent Reading Corpus and Curriculum Reading Corpus with Exam Text Corpus and Other Reference Corpora..... 140

List of Figures

Figure 1 Scatterplot of Keyword Scores and Frequency Per Million in the British National Corpus. Keywords are included for a small sample to illustrate the types of words found in different parts of the figure 84

List of Appendices

Appendix A Exam Text Corpus Keyword List

Appendix B Relative Frequencies of Keywords in the David Lee Categories

Appendix C The survey

Appendix D Individual Author Scores on the ART

Appendix E Keyword Vocabulary Test Results

Appendix F Independent Reading Corpus Top 100 nouns by frequency

Appendix G Curriculum Reading Corpus Top 100 nouns by frequency

Appendix H Independent Reading Corpus Top 100 verbs by frequency

Appendix I Curriculum Reading Corpus Top 100 verbs by frequency

Appendix J Independent Reading Corpus Top 100 adjectives by frequency

Appendix K Curriculum Reading Corpus Top 100 adjectives by frequency

Appendix L Independent Reading Corpus Top 100 adverbs by frequency

Appendix M Curriculum Reading Corpus Top 100 adverbs by frequency

Appendix N Ethical Approval Form

List of Acronyms

ART	Author recognition test
BNC	British National Corpus
CAG	Centre Assessed Grade
CRC	Curriculum Reading Corpus
DfE	Department for Education
ETC	Exam Text Corpus
FE	Further Education
FPM	Frequency Per Million
GCSE	General Certificate of Secondary Education
IRC	Independent Reading Corpus
LQH	Lexical Quality Hypothesis
LLH	Lexical Legacy Hypothesis
MATTR	Moving-average type-token ratio
MRT	Magazine recognition test
OECD	Organization for Economic Cooperation and Development
OFQUAL	The Office of Qualifications and Examinations Regulation
SVR	Simple View of Reading
TTR	Type to token ratio
YA	Young adult
YARC	York Assessment of Reading for Comprehension

Thesis Structure

This thesis is structured as a thesis by papers. Chapters 4, 5 and 6 were written as stand-alone papers and are reproduced as such in this thesis. The format, content and length of these papers has been influenced by the journals in which they have either been published or to which they have been or were planned to be submitted. Papers 1 and 2 (Chapters 4 and 5) have also been influenced by the peer review process. As they were written and submitted as independent papers, there is some unavoidable repetition of the literature cited. The methodology section of this thesis (Chapter 3) is an overview of the methods selected, with the specific details of the methodologies for each of the three separate parts of the study contained within the papers (Chapters 4, 5 and 6).

As is quite common with a thesis structured by papers, this thesis is shorter than it would have been if it had been written in a traditional format. Journal word limits influenced the length of the three papers and so the research projects reported in them have been written up concisely.

Impact of COVID-19

The research contained in this thesis was impacted in two main ways by the COVID-19 pandemic: 1) there was an impact on the attainment data collected in the student survey for the second paper (Chapter 5); and 2) there was impact on myself as a researcher and my workload as a teacher.

The primary data for the second part of this project was collected from 16–18-year-old students in 2021. Two General Certificate of Secondary Education (GCSE) cohorts were included in these respondents: those who took their GCSEs in 2019, prior to the COVID-19 pandemic; and those who were awarded their GCSE grades in 2020, when all public exams in England were cancelled. As the 2020 grades were awarded by schools (Department for Education, 2020), the grades attained by the two cohorts were not comparable. The 2019 cohort took two exams for their English language GCSE and 50% of the marks were awarded for questions based on the unseen reading texts in those exams. For the 2020 cohort, there was no consistent policy or method for schools to decide what grades to award to students, so there is no way of finding out how much of the grade, if any, depended on students' reading skills. There was also widespread grade inflation after norm referencing was abandoned due to a public outcry (Ofqual, 2020). Analysis of the results of the survey therefore split the respondents into the two cohorts, and cohort year was used as a factor in the analysis, which limited the power of the statistical tests.

As a part-time PhD candidate, I had also been working (up to April 2024) three days a week as a secondary school teacher. When schools closed in England for in-person teaching in March 2020, there was an enormous amount of pressure on all teachers, including myself, to quickly transfer our teaching online. There was a considerably higher workload in: creating suitable lessons and resources that students could access at home; learning new packages and techniques for online lessons; taking turns in school with the vulnerable students who were

attending in person; and dealing with any pastoral and safeguarding issues. Whilst the support from my supervisors and university activities moved seamlessly online, I found it very difficult to find enough time for PhD work alongside the increased demands of my school teaching. Therefore, in the autumn term of 2020, I applied for and was granted a three-month retrospective suspension to account for the time in the Spring and Summer terms where my teaching responsibilities had made it impossible to progress my PhD work as expected. I am grateful to the university for granting my suspension, which has allowed me the extra time I needed to finish. I am also proud of the service that I was able to offer my school, students and community during the school closures and in the challenging classroom environments that followed.

Prologue

Several years ago, alongside my work as an English teacher in a secondary comprehensive school, I worked as an examiner marking English exam papers. The extract, that the candidates had to read and answer questions about, described a backpacking young woman who was staying at a hostel. The dormitory and bathroom at the hostel were extremely dirty and poorly maintained and the manager was more interested in socialising with the residents than in undertaking his duties properly. At the end of the extract the mother of the backpacking young woman arrives for a surprise visit and makes a big fuss about the state of the hostel and the way it was being run. One of the final questions on the paper required candidates to show understanding of why the mother was angry. To me it seemed obvious: the dirt; the poor maintenance; and the socialising manager. Many candidates failed to identify these things and empathise with the mother's anger, and I struggled to understand why the text was not being comprehended.

Eventually I came across one candidate's paper that explained this seeming inability to understand the reason for the mother's anger and upset. The candidate wrote that the mother had no right to be angry and upset as she must have been at least partly to blame for her daughter having to live in a homeless hostel. The source of the problem, the apparent failure to understand the events described, was the word *hostel* and the nuances in its meaning. Whilst all the candidates seemed to understand that *hostel* described accommodation, many seemed not to have experience of the word as somewhere cheap to stay when travelling and only understood it as somewhere that homeless or other vulnerable people might stay. Candidates' experience with these different meanings of the word would have been impacted by their different life experiences. If they had not been on holiday or did not know anyone who had been backpacking, for example, then they may not have experienced *hostel* in the way it was used in the exam text. If they lived in an area that had

homeless or asylum or bail hostels, then this meaning would potentially be more dominant for them. This meaning of *hostel*, as housing for vulnerable or homeless people, is clearly completely valid but was not the meaning needed for this extract. This misunderstanding of an exam passage by candidates, set me on my journey exploring vocabulary and reading comprehension.

As a secondary English teacher, I had struggled to find ways, in the classroom, to ensure that students could independently comprehend what they were reading. If there were words that they did not know, or meanings that they had not come across before (as in the *hostel* example above), students' comprehension was compromised, and they were reliant on explanations from me or others to ensure their understanding. This seemed to be especially true for those who failed to reach the pass mark in the exam. This motivated me to make the focus of my PhD reading comprehension and vocabulary, with the aim of discovering the type of vocabulary that students needed to know and also to understand how vocabulary knowledge is built.

Chapter 1. Introduction

1.1 Background

In 2013 the Department for Education (DfE) started a reform of the General Certificate of Secondary Education (GCSE) qualifications. These qualifications are taken by students in England at the end of year 11 (age 16). The stated intention of the reform was to make the GCSEs "more engaging and worthwhile to teach and study, as well as more resilient and respected" and "to prepare young people better for the next steps in their education or employment in years to come" (Ofqual, 2013, p. 3). English language, English literature and maths were the first GCSEs to be reformed, with the new specifications beginning in 2015 and the first exams taken in 2017. This study focuses on the English language GCSE, success in which is critical as the grade achieved is used as a key indicator of literacy competence. A grade of 4 or above is a requirement for entry onto many post-16 courses, as well as for many jobs, and some university courses require at least a Grade 5. Any student who fails to achieve the pass grade (Grade 4) is required to continue to study English and retake the qualification until they either pass or leave education (usually two years later).

Prior to 2010 there were two specifications for English: English GCSE and English literature GCSE. The English GCSE was mandatory whilst only 75% of students sat the English literature GCSE (Isaacs, 2014). Controlled assessments (coursework) accounted for 40% of the English GCSE and there were two entry tiers for exams, foundation and higher. The foundation papers were less demanding but restricted candidates to grades C-G.

In 2010 a new set of three English qualifications was introduced, a general English GCSE (that combined language and literature), an English literature GCSE and a new English language GSCE. In the summer of 2012 around 70% of students took the English language GCSE and the English literature GCSE, the remaining being entered for the combined English (Isaacs, 2014). Analysis of the 2013 results from the largest exam board AQA suggested that lower attaining students were being entered for the combined English GCSE

(Isaacs, 2014). In this post-2010 qualification the controlled assessment (coursework) component had increased to 60% of the English language GSCE and tiered entry options also remained for the exams. Schools were therefore able to enter candidates for the combined English GCSE, who may have struggled to cover the content of the separate language and literature GCSEs, and were also able to select the foundation or higher tiered exam papers within all three of the qualifications (Isaacs, 2014). In addition, the qualification was modular, with exams offered in January and June each year, and students were able to retake each module once.

A study of the reading texts in the 2011-2013 exam papers of the largest awarding body, AQA, revealed a difference in the structures of texts chosen for foundation compared to higher tier papers. In foundation papers low-level text structures, such as description and collection/sequence, were used more often than they were in the higher tier papers, whereas in the higher tier papers high-level text structures such as causation and comparison were used more often (Gorelova et al., 2015). This suggests that foundation tier reading texts were being selected by those setting the exams to be more accessible to lower attaining readers.

The content of 2010 English language GCSE was required to have a weighting of 45%-55% on functional aspects of English language, these functional aspects were focused on communication skills in real life situations (Ofqual, 2011). Specification content alluded to 'cultural diversity, multimodal study and connections to the real world and daily life' (Isaacs, 2014) and newspaper articles were used to assess candidates' abilities to identify facts and opinion (Verhoeven, 2022).

When the new English language GCSE specification, the subject of this study, was introduced in 2015 five key changes were: 1) to move to exam only assessment, rather than having any controlled assessments, which were coursework assignments that were both produced and assessed in schools; 2) to remove the spoken language component from the

overall GCSE grade; 3) to make the course linear, with all exams taken in the summer at the end of the two year course; 4) to remove entry tiers, so all students would be entered for the same exams rather than having foundation and higher tier options; and 5) to change the grading structure to 9 - 1 rather than A*- G (Ofqual, 2013).

There were also changes to the content of the assessments. One of the implications of these changes was that reading and comprehending unseen texts in exam conditions moved from being worth only 20% of the total assessment in the old specification, to being worth 50% in the new specification. In the new English language GCSE students were presented with three unseen extracts and their ability to comprehend these successfully was key to answering the questions in the reading part of the exam. In addition, the removal of the tiers meant that there were no longer more accessible texts for those students entered for the foundation exam (aimed at students who would achieve the grades C-G), who were likely to have poorer reading skills. Instead, the same two papers would be taken by all students and the DfE's conditions and requirements for the new qualification specified that "high-quality, challenging texts from the 19th, 20th and 21st centuries" must be used and they "must include literature and extended literary non-fiction" (Department for Education, 2013, p. 4). This was a big shift from the previous specification where the functional English of daily life was more of a focus (Isaacs, 2014, p. 143). In the old specification less-skilled readers could be supported in the classroom, with the more challenging texts being studied for the controlled assessments, in addition to being selected for the foundation paper exam.

Overall, the new specification of the exam removed a number of ways in which students who found reading comprehension difficult had been supported in the past to pass their English language GCSE qualification. Teachers needed to find new ways to help them to prepare and succeed.

1.2 The Research Gap

This research project has grown out of a desire to understand how best to prepare students for the reading comprehension challenges that they face in these new exams. Without an understanding of the specific challenges that comprehending these higher-level texts present to students, it is very difficult to know how teaching and learning practices need to change to enable students to reach the required level for success. Whilst prior attainment is the main factor associated with success for most GCSE subjects (Baird et al., 2019), for English language GCSE, gender also has a high association with success, with female students performing better, even after accounting for prior attainment. In addition, attending a school location in an advantaged area and having a higher socioeconomic status increases the likelihood of getting a 'good grade' (A*- C) at English language GCSE (Baird et al., 2019). Beyond the English context, the issue of how the nature and content of examinations might have unforeseen consequences in terms of fairness, accessibility and impact on the curriculum is an important area of research in a wide range of international educational jurisdictions.

As noted above, reading comprehension skills assume a much higher level of importance in the new English Language GCSE exam. One crucial predictor of successful reading comprehension skill is reading experience (Acheson et al., 2008; Chateau & Jared, 2000; Mol & Bus, 2011). A reader's understanding of a text is highly dependent on them having encountered the vocabulary it contains previously, multiple times and in different contexts (Nation, 2017). These previous encounters with words in diverse contexts will allow the reader to build up, cumulatively, precise yet flexible knowledge of word meanings (Perfetti & Hart, 2002) which they can bring to bear on the text they are reading. Comprehension of the unseen reading texts in the English language GCSE exams therefore depends on the extent to which students' reading experience has provided sufficient exposure

to the type of vocabulary found in the exam texts. It is not only important, therefore, to have a detailed understanding of the reading comprehension challenges presented by the new English language GCSE specification, but it is also key to know what students are reading and whether or not their reading practices are adequately preparing them for the vocabulary in the exam texts.

1.3 Research Questions

This research project collected three sets of primary data: 1) a new corpus from publicly available English language GCSE exam texts; 2) survey data that measured adolescent reading experience, vocabulary knowledge and English language GCSE attainment; and 3) two new corpora of students' independent (out of school) reading and school curriculum materials. This PhD research project is structured into these three separate parts and has been written as three research papers. Each paper has its own research questions as follows:

Paper 1

RQ1: What type of vocabulary is typical of the exam texts?

RQ2: In what types of reading materials is the vocabulary that typifies the exam texts most likely to be found?

Paper 2:

RQ1: Do students who gain high grades (7-9) in the English language GCSE have higher levels of reading experience, compared to those with mid grades (4-6) with: a) classic authors; and b) Young adult (YA) authors?

RQ2: Do students who gain high grades (7-9) score more highly on a vocabulary test created from keywords from the exam texts, compared to those with mid grades (4-6)?

Paper 3:

RQ1: What is the linguistic make-up of students' independent and curriculum reading?

RQ2: What types of words typify the students' reading materials?

RQ3: How far does students' reading prepare students for the vocabulary in the English language GCSE exam?

1.4 Overview of the Study

This thesis consists of seven chapters. This chapter, Chapter 1, gives a brief background to the English language GCSE, the qualification that is the focus of this project, and then outlines the research gap and research questions. Chapter 2 starts with an overview of the literature on the components of successful reading and interventions for improving reading. Two theoretical hypotheses are then outlined: the lexical quality hypothesis (LQH) (Perfetti & Hart, 2002) and the lexical legacy hypothesis (LLH) (Nation, 2017). The ways in which reading experience is measured and the impact that it has on reading skills is then also explored. Finally, there is an overview of the literature on adolescents' reading habits and practices. Chapter 3 is an overview of the main methodologies selected for the study, corpus linguistics and an online survey. The rationale for using these methodologies and the key measures is given.

Chapter 4 is Paper 1, in this part of the study corpus linguistics methods were used to create a corpus of a sample of the exam texts and to identify and analyse the kinds of vocabulary that were representative of these texts. Further corpus study, using existing reference corpora, was then able to identify in which genres the exam vocabulary was most likely to be found. Chapter 5 is Paper 2, in this second part of the study an online student survey was used to explore the relationship between respondents' reading experience, vocabulary and attainment in the English language GCSE. Chapter 6 is Paper 3, this final part

of the study created two corpora of students' curriculum and independent reading. The linguistic contents of these corpora were analysed and then compared to the exam corpus, created for the first part of the research, to see how far students' actual reading experience matched what was in the exam.

Chapter 7, the Discussion, starts with a short overview of the results of the project. The limitations are then discussed, covering both methods and theory. Then there is a section on implications: for future research; for theory; and for practice. Conclusions are then drawn that cover the whole of the project.

Chapter 2. Literature Review

This thesis examines the reading comprehension challenge presented to students by the new specification of the English language GCSE exam, as outlined in the Introduction. There are many skills needed for successful reading comprehension, but the decision taken for this thesis was to focus on vocabulary knowledge and reading experience, in all three parts of the research.

Vocabulary knowledge is one of the strongest predictors of reading comprehension: at a very basic level, if the words in the text being read are not understood then the text cannot be comprehended. Perfetti and Hart's LQH (2002) and Nation's LLH (2017) explain how vocabulary knowledge is built through reading experience. The LQH characterises the ability to read a word efficiently as being able to access high quality representations of three components of a word: its orthography (written form); its phonology (sound); and its semantic information (meaning). Reading experience is seen as the "critical foundation" for skilled reading, because it is reading experience that provides repeated exposures to words and therefore repeated opportunities to build high-quality representations of the words that are encountered when reading (Perfetti & Hart, 2002, p. 212). High lexical quality is not possible without repeated encounters. The LLH explores how high-quality representations are built through "experiencing words in diverse and meaningful language environments" (Nation, 2017, p. 1) and makes a direct link between being a successful reader and "an individual's lexical experience", created through what they have read before (Nation, 2017, p. 3).

In the first part of this review, the components of successful reading will be considered, with particular focus on adolescent readers, who are the focus of this study. Next, in order to explore causal inferences, intervention studies aimed at improving reading comprehension will be reviewed. Then the LQH will be used as a theoretical framework to explore the ways in which reading experience and reading ability are linked and how vocabulary is learnt

through reading. Reading experience will then be considered within the context of the LLH, particularly the experiencing of words in diverse contexts. What we already know about the reading habits of adolescents will then be reviewed, due to the crucial role that reading experience plays in the acquisition of reading comprehension skill.

2.1 Components of Successful Reading

Successful reading is described by the *Simple View of Reading* (SVR) as being the product of two parts: the ability to decode written words (either by sounding them out or by recognising them immediately) and the linguistic comprehension of the words, what Gough and Tunmer (1986) describe as "the process by which, given lexical (i.e. word) information, sentences and discourses are interpreted" (p. 7). The simplicity of this framework, and the clarity with which it describes successful reading, means that the SVR has been incredibly influential in theories of and research into reading comprehension.

Kendeou et al.'s factor analysis study (2009) of young readers (aged between 4 and 6), showed that the two components of the SVR were distinct. Using listening comprehension as a measure of linguistic comprehension, they showed that a child might be able to decode well but have poor listening comprehension or they could score highly on listening comprehension measures but struggle to decode efficiently: in either scenario successful reading comprehension (as measured by Dynamic Indicators of Basic Early Literacy Skills, Story Retell) (Kaminski, 2002), was compromised, emphasising that each component is necessary but not sufficient for successful reading comprehension.

The relative importance of the two components of the SVR does not, however, remain consistent for readers across their reading development. As students get older, and their reading more proficient, the decoding element of the SVR declines in importance. This was illustrated in comparisons of decoding and passage comprehension scores from a longitudinal study of children from Grade 1 to Grade 9 (Francis et al., 2005). The highest correlation

between decoding and passage comprehension was in Grade 1 (0.89) and the lowest was in Grade 9 (0.63). Gough et al. (1996) argue that once a ceiling for decoding has been reached, the linguistic element must become more important for successful reading comprehension.

A study by Catts et al. (2006) found support for this change in the importance of the components of the SVR, by comparing the data they had collected on reading achievement for participants in the eighth grade with data that was available for the same participants from an earlier study (Tomblin et al., 1997). Children were divided into three groups on the basis of their eighth grade reading achievement: 1) poor reading comprehenders with typical word recognition skills; 2) poor decoders with typical reading comprehension; and 3) typical readers (with typical word recognition and comprehension). Whilst the poor decoders' reading comprehension scores were lower at 2nd grade than the poor comprehenders', this reversed at 4th grade, with the poor decoders gaining higher reading comprehension scores than the poor comprehenders. This suggests that as children progress through school it is linguistic comprehension that drives reading comprehension skill, rather than decoding ability.

James et al. (2020) studied three groups of children from ages six to thirteen, dividing them into three age groups (6-8 years, 9-11 years and 12-13 years). For the younger two groups reading comprehension was measured using the *York Assessment of Reading for Comprehension* (YARC) Passage Reading (Snowling et al., 2009). *YARC Secondary* (Stothard et al., 2010) was used for the older group. These reading comprehension measures were used as the dependent variables in hierarchical regression analysis, with age, non-verbal reasoning, vocabulary, phonological awareness, word reading and morphological awareness as predictor variables. For the youngest age group all variables were significant predictors. For the 9-11 years group, all the variables except word reading predicted reading comprehension, and for the 12-13 years group, all the variables except word reading and

phonological awareness predicted reading comprehension: clear support that decoding and word recognition, once mastered, decline in importance with age. The vocabulary measure, which used the *British Picture Vocabulary Scale Third Edition* receptive vocabulary test (Dunn et al., 2009), uniquely predicted reading comprehension in all three age groups in the study (from 6-13 years).

Tilstra et al. (2009) looked at slightly older children, testing participants from the fourth, seventh and nineth grades and also found that the variance in reading comprehension explained by decoding decreased as the students got older and that the variance explained by listening comprehension increased from the fourth to seventh grade but did not increase further at the nineth grade. They argued that components outside the SVR, for example depth of vocabulary knowledge (how well a word is known), prior knowledge and verbal working memory, were also contributing to reading comprehension in the oldest age group. This suggests that it is not only decoding that becomes less important as a predictor, but that listening comprehension also becomes less important as the increasingly difficult reading materials, encountered in the higher grades of school, deviate more from spoken language. That written language is seen as distinct, in structure and vocabulary, from spoken language is a point that will be returned to later in this section.

Braze et al. (2007) studied even older readers, specifically selecting participants aged 16 to 24 whose reading skills were "poorly developed" (Braze et al., 2007, p. 226). The *Peabody Picture Vocabulary Test-Revised* (Dunn & Dunn, 1997) and a subtest from *Wechsler Abbreviated Scales of Intelligence* (Psychological Corporation, 1999) were used to measure receptive and expressive vocabulary respectively. In many studies receptive and expressive tests of vocabulary are used together as a measure of listening or language comprehension (e.g. Catts et al., 2006), but Braze et al. (2007) used them as a separate vocabulary measure and used the *Peabody Individual Achievement Test–Revised* (Markwardt, 1998) to assess

speech sentence (listening) comprehension. Print sentence comprehension was also assessed using the *Peabody Individual Achievement Test–Revised* (Markwardt, 1998) and print passage comprehension was assessed using the *Gray Oral Reading Test* (Wiederholt & Bryant, 2001). Braze et al. (2007) found that the SVR was a reasonable fit for their data, with a model that included listening comprehension and decoding ability accounting for 76% of the variance in their composite reading comprehension measure (of sentence and passage comprehension). However, when the composite vocabulary measure (receptive and expressive) was added to the model it accounted for a further 6% of the variance, showing that vocabulary knowledge's contribution to reading comprehension "overlaps considerably with the contributions of decoding and listening comprehension, but, contrary to the predictions of the SVR, it is not wholly contained within them" (Braze et al., 2007, p. 234). Furthermore, they found that the vocabulary measure predicted unique variance in reading comprehension but not speech (listening) comprehension, which suggests that reading comprehension is more dependent on vocabulary knowledge. Braze et al. (2007) explained this difference by arguing that speech contains more linguistic information, for example prosody and contextual support, meaning that its comprehension does not have to rely so much on the recipient's vocabulary knowledge. Reading comprehension is more dependent on vocabulary knowledge in three keyways (1) it includes more low frequency vocabulary than speech, (2) it provides fewer linguistic cues (i.e. prosody and contextual support), and (3) it is a less practised modality than speech, for some at least (Braze et al., 2007).

Nation and Snowling (1998) studied two groups of children between 8 and 10 years old, who were matched for nonword reading and nonverbal ability but who differed in reading comprehension ability. Those in the typical comprehender group were at or above the expected age for reading comprehension, while those in the poor comprehender group were at least one year below the expected level for reading comprehension. Two expressive

(providing definitions and multiple meanings) and two receptive (synonyms and figurative expressions) measures of vocabulary were used from the *Test of Word Knowledge* (Wiig & Secord, 1992) and there was "a clear and consistent relationship between impoverished vocabulary knowledge and reading comprehension difficulties" (Nation & Snowling, 1998, p. 90). Further experiments tested performance on synonym and rhyme tasks and found that there was a greater difference between the poor comprehender group and the typical comprehender group on the synonym tasks than on the rhyme tasks, providing additional evidence that poor comprehenders have semantic rather than phonological processing weaknesses. In a final experiment it was found that the poor comprehender group performed more poorly than the typical comprehender group in reading low frequency exception words, that is words that do not occur frequently in language and have unusual spelling patterns (e.g. mould, dread). Nation and Snowling argued that this demonstrated a semantic deficiency in the poor comprehender group as the low frequency exception words depend more on semantic support for fluent recognition than high frequency regular words do. It is possible that this semantic deficit is due to less reading experience as, as will be outlined later in this review, vocabulary knowledge is built through repeated, diverse experiences with words (Nation, 2017; Perfetti & Hart, 2002).

Henderson et al. (2013) also studied poor comprehenders and tested how they well they read ambiguous words, that is words with multiple meanings (homonyms). The method the study used was semantic priming, which is when the target word is preceded by a prime word that is either related or unrelated to a meaning of the target word. Faster naming times would be expected for words that follow related primes than for those that follow unrelated primes. The first experiment in this study was a picture naming task. Participants heard each homonym three times, once with a picture of the dominant meaning, once with a picture of the subordinate meanings and once with an unrelated picture. One test gave participants 250

ms between hearing the prime words and the target picture. A second test gave participants 1,000 ms between prime and the target picture. Participants were asked to indicate, by saying yes or no, whether the word and picture were related. Comparison of reaction times showed that the poor comprehender group were slower in all conditions compared to a chronological age control group but not a vocabulary age control group. At 250 ms all groups showed dominant priming effects, however the poor comprehenders did not show an effect of subordinate priming, but both control groups did. This was despite the poor comprehenders group knowing the subordinate meaning of a word when tested in a vocabulary task (breadth of vocabulary). In the second experiment of the study participants listened to sentences with primes. An example of a prime in a sentence is, if the homonym *match* is preceded in a sentence by the prime *football* (as compared to lighting a fire), then the prediction is the appropriate meaning of match will be more quickly accessed. Participants listened to sentences that were biased, through the prime, to the subordinate meaning or control sentences. There was then 250 ms or 1000 ms intervals and then participants named either a picture depicting the subordinate meaning or dominant meaning of the homonym. At 250 ms all groups were faster naming subordinate pictures if they were preceded by a subordinate-biased sentences. Poor comprehenders and the vocabulary age control showed a reverse priming effect (faster to name dominant pictures when preceded by subordinate-biased sentences) but the chronological age control groups' inappropriate priming effect was not significant. At 1,000 ms, in contrast to the control groups, the poor comprehenders group showed no significant appropriate priming but did show significant inappropriate priming. Henderson et al. (2013) concluded that, even though poor comprehenders may know the subordinate meaning of a homonym (breadth of vocabulary), they struggled to reduce activation of inappropriate dominant meanings. This suggests that poor comprehenders' networks of semantic representations are poorly coded, particularly for less frequent

(subordinate) meanings. It is important, therefore, to consider a more nuanced view of vocabulary knowledge than just whether or not a word is known to a reader. As will be outlined below, the number of separate lexical items stored in the lexicon is considered the breadth of vocabulary knowledge whereas the detail of the semantic information about each item of vocabulary is the depth (Ouellette, 2006).

This idea, that depth rather than just breadth of vocabulary knowledge affects reading comprehension ability, had already been studied by Ouelette (2006). Ouelette used the model of the mental lexicon (Levelt et al., 1999) to describe how vocabulary is stored: phonological and semantic information are stored distinctly but are also connected. A word can be known, or stored, in the phonological part of the lexicon, for example when it is first heard or learnt, without it being fully understood or configured in either the orthographic or the semantic part of the lexicon. Just knowing the phonological form of a word would be considered breadth of vocabulary knowledge, whilst also having fully configured knowledge of a word's orthography and meanings would be considered depth of vocabulary knowledge. In Ouelette's study, vocabulary was measured using the *Test of Word Knowledge* (Wiig & Secord, 1992), with receptive and expressive subtests standing for breadth of vocabulary and word definitions and synonym subtests standing for depth of vocabulary. Ouelette was able to show that vocabulary breadth predicted decoding and visual word recognition whereas vocabulary depth predicted reading comprehension. Therefore, for reading comprehension skill, it is not just whether or not a reader knows a word that is crucial for reading comprehension but how much a reader knows about a word.

What has emerged is that the importance of the different components required for successful reading comprehension ability change over time. As readers become more proficient, the listening comprehension component of the SVR increases in importance over decoding ability and word reading. Subsequently, vocabulary knowledge, specifically the

semantic knowledge of words, particularly the depth of that knowledge in the reader's lexicon, increases in its contribution to reading comprehension skill.

If vocabulary is central to successful reading comprehension, particularly for older readers, who are the subject of this study, then it is important to consider how vocabulary is learnt. Whilst some vocabulary is taught to students in classrooms this is only estimated to be around 200-300 words each year (Nagy et al., 1987), whereas the overall estimate of the number of words that children learn each year is around 3000 (Nagy & Herman, 1984). Or, to put it another way: a twelve-year-old student is likely to learn 10-15 words a day with only one of those words being directly taught in the classroom (Landauer & Dumais, 1997). The gap between words learnt and words taught is closed, according to Nagy et al. (1987) by listening and reading. Children were found to be able to learn words from context, although not necessarily at a high rate (Nagy et al., 1985, 1987). For older children it becomes less likely that the acquisition of new words will come through listening as they will have already heard, by age 12, most words that predominantly feature in spoken language (Landauer & Dumais, 1997). For there to be "large scale vocabulary growth", there therefore needs to be a "sufficient volume of wide reading" (Nagy et al., 1987, p. 239).

Nagy et al. (1987) tested students' ability to learn words from written texts in an intervention with 418 students from between third and seventh grades. Participants were assigned texts to read, either two narrative or two expository, and were then tested on vocabulary found in these written texts 6 days after reading them. Scores were higher when the written texts were judged to be easier and when the texts were narrative rather than expository, a finding that will be returned to later in this section. Although the chance of learning vocabulary from context was small, Nagy et al. argued that this was still enough to explain vocabulary growth because of the large amount of text that children read independently. Landauer and Dumais (1997) simulated learning from reading by using latent

semantic analysis to analyse a large body of text and found that their model was able to increase its knowledge of vocabulary, at the same rate as children learn, without any other prior knowledge. This shows that words can be learnt from independent reading and this method of learning can account for much of the increase in vocabulary knowledge that occurs in children. The amount of reading that an individual reader does is crucial to this process as it gives ever increasing opportunities to encounter and re-encounter individual words in various contexts. It is hard to capture this incidental learning through reading experimentally as many encounters of a word are needed (Batterink & Neville, 2011; Godfroid et al., 2018; Hulme et al., 2019) but that does not mean it is not happening.

It is not just individual readers' prior reading experience that can impact on their reading ability, it is also the texts that are being read. At secondary or high school, reading experience becomes even more crucial as written text moves increasingly further away from the structures and vocabulary of spoken language. Academic, technical and subject specific language, connectives and words that indicate or create cohesion in non-narrative texts are all far more frequent in written than in spoken language and therefore reading experience is likely to be the chief source of exposure to this type of vocabulary (Tilstra et al., 2009). Secondary school students encounter more texts that inform rather than narrate, and techniques and strategies taught for the comprehension of narrative texts may not apply to expository texts and therefore will not aid comprehension or improve comprehension ability (Edmonds et al., 2009).

Denton et al. (2015) ran a study that considered different genres of texts using 325 participants between the ages of twelve and eighteen. Using results from the *Gates-MacGinitie reading comprehension* subtest (MacGinitie et al., 2000) students were separated into two groups called: adequate comprehenders (above the 25th percentile) and poor comprehenders (below the 25th percentile). Each participant was given two

texts to read: one that was accessible to their level of reading and one that was challenging. Half of the participants were given informational texts and half of them were given narrative texts. Participants' readings of the texts were assessed using a 'think aloud' methodology, where participants read aloud stopping at prompt lines to describe their thought processes. Thought process responses were coded as either: inference generation, monitoring (of understanding) or other responses (e.g. paraphrasing or repeating verbatim). Responses were then judged as either acceptable or unacceptable with reference to the meaning of the text. Poor comprehenders produced fewer acceptable paraphrasing and inference responses when reading informational texts, suggesting that this genre presents a higher level of difficulty even when texts were specifically selected to be accessible for the participants' level of reading. One limitation of the study was the availability of challenging narrative texts for the highest performing readers in the adequate comprehenders group. Adding 350 Lexiles, a scaled measurement of text difficulty, to the already high performing readers' scores meant that their challenging narrative texts frequently had historical settings. Denton et al. (2015) concluded that the fewer acceptable inference responses to the challenging narrative texts they read, compared to other texts, could have been due to the archaic vocabulary and unfamiliar syntax used. Whilst it seems that this study shows that expository texts can present challenges to comprehension and that this needs to be considered when studying adolescent readers, as non-fiction texts become a much larger part of their reading diet, narrative texts at the higher end of the difficulty level can also be problematic to comprehend, even for good comprehenders. This is especially pertinent to the current study as reading texts in the English language GCSE exam are required to be both fiction and non-fiction and to be taken from all three of the 19th, 20th and 21st centuries.

Therefore, there is likely to be additional challenge both from the non-fiction texts and from the older ones.

What we have seen so far is that the SVR describes two separate but necessary components of reading success: decoding and linguistic comprehension. In the early years of reading, decoding is a key component, but this recedes in importance as readers get older and more proficient. The second component of the SVR, linguistic comprehension, increases in importance, with listening comprehension becoming a stronger component of reading comprehension over time. Vocabulary also becomes important as reading materials increase in difficulty and move away from the familiarity of spoken language to include technical, academic, archaic and low frequency words. Vocabulary depth, rather than just breadth, also emerges as important to reading comprehension. With these components in mind, this review now turns to intervention studies and looks at the ways in which they have sought to improve readers and their reading ability.

2.2 Intervention Studies for Improving Reading

Whilst the studies outlined above have provided important information about some of the key elements needed for successful reading comprehension, intervention studies can provide insights into what can improve reading success, as they are designed to be able to determine causal relationships. In their synthesis of reading interventions for older struggling readers, Edmonds et al. (2009) note that one of the challenges in evaluating reading interventions is that they vary widely. Some intervention studies focus on improving word reading, some on fluency, some on comprehension of sentences and some on comprehension of larger pieces of texts, some are teacher-administered and some administered by the researcher, some use researcher-developed measures and some use standardised measures. Identifying which interventions improve which, if any, aspects of reading is not, therefore, straightforward. There is also an additional challenge, when looking at research into

adolescent reading, that, whilst interventions and reading instruction studies are common in primary or elementary schooling, they are far less common for secondary students (Edmonds et al., 2009).

Interventions that focus on strategies, for example: identifying the main idea in a paragraph (Jitendra et al., 2000); reciprocal teaching (Alfassi, 1998); morphemic analysis and context cues (Baumann et al., 2002); theme identification (Wilder & Williams, 2001; Williams et al., 1994); and metalinguistic ambiguity instruction (Zipke et al., 2009), have observed improvements in measures of that specific strategy but there is little evidence that such specific interventions improve general reading comprehension ability (Edmonds et al., 2009). This means, in interventions that focus on a particular strategy, participants can show improvements on measures of that strategy, but any transfer to improvements in reading comprehension remain inconsistent and elusive.

It might be expected that vocabulary interventions would have more success, considering the importance of vocabulary discussed in Section 2.1 of this review (Beech, 2002; Braze et al., 2007; Cunningham & Stanovich, 1990; Henderson et al., 2013; Nation & Snowling, 1998; Ouellette, 2006; Sénéchal et al., 1996; Stanovich & Cunningham, 1992; Stanovich & West, 1989; Tilstra et al., 2009). Vocabulary interventions that include instruction on specific word meanings have proved to be effective to support comprehension of a passage containing those same words (Brabham & Lynch-Brown, 2002; Gray et al., 2018; Hawkins et al., 2010; Kameenui et al., 1982). In vocabulary interventions that have used active processing, that is a combination of explanations of meaning and opportunities to encounter the words in different contexts, larger effects have been achieved (Apthorp et al., 2012; Beck et al., 1982; Lesaux et al., 2014). However, although effects have been shown in researcher-developed bespoke tests, that tested the vocabulary that had been taught, there have been no significant effects on standardised comprehension or vocabulary tests (e.g.

Apthorp et al., 2012; Beck et al., 1982; Nash & Snowling, 2006). Although these studies did not show that vocabulary instruction has improved general reading ability, that is the improvements in performance have not transferred to untaught words (Wright & Cervetti, 2017), other studies, to be discussed below, have shown that vocabulary can be the main driver of improvement (e.g. Clarke et al., 2010).

Two vocabulary intervention studies that have shown improvements on comprehension of passages are Lubliner and Smetana (2005) and Sampson et al. (1982). The first study, with fifth grade participants, was a twelve-week classroom-based intervention in a low-performing school that taught all three fifth grade teachers the techniques of the vocabulary intervention. Students in all three classes were then taught by the same individual teacher in their social studies lessons to both monitor their own understanding of vocabulary and to learn words from context. Although the intervention was only taught by one teacher, all three teachers referred to the techniques in other lessons too. The results, after the intervention, were compared to a control period before the intervention as well as to a control group in a high-performing school in the same state, who did not receive the intervention. Gain scores of the intervention group on their reading comprehension and vocabulary test scores showed "greater gains and larger effects during the experimental period" (Lubliner & Smetana, 2005, p. 182) and were not replicated by the control group in the high-performing school. The study suggested that the improvements on a standardised measure of comprehension were due to the multifaceted metacognitive nature of the intervention, that is: the involvement of students in how they were learning vocabulary; the transferring of these techniques to reading in other lessons; and the transferring of the techniques to their independent reading. It should be noted that this was not a randomised controlled trial and therefore caution should be exercised and the results not overinterpreted.

The second vocabulary intervention study (Sampson et al., 1982) was a fifteen-week intervention with third grade participants that used quasi-instructional cloze exercises. Traditional cloze procedures require the completion of a pattern by filling in a gap or gaps. For vocabulary tests or exercises this is usually the insertion of a word in a gap in a sentence, for example, *The boy wanted to* _____ *his book*. Sampson et al.'s intervention included a student discussion, after completing a cloze exercise, of whether the words selected by each of the students could or could not be correct in the context of the sentences. Using the comprehension subtest of the *Gates-MacGinitie Reading Tests* (MacGinitie & MacGinitie, 1992) the experimental group showed a significant increase in their results. There was not, however a significant improvement for the experimental group in vocabulary scores. Sampson et al. explain the rise in comprehension scores without a concurrent rise in vocabulary scores as attributable to the discussions that took place during the intervention, the suggestion being that these produced "vocabulary fine-tuning" (Sampson et al., 1982, p. 398) (i.e. depth of vocabulary knowledge), rather than knowledge of new words (breadth of vocabulary).

Kuhn and Stahl (1998) reviewed 14 studies that, similar to the Lubliner and Smetana (2005) study outlined above, aimed to improve the learning of words from context. As would be expected from the studies reviewed above, readers' ability in the taught strategy, that is learning words from context, improved. However, where a control group was asked to derive meanings from context, but without being taught specific strategies to do this, they performed as well as the intervention group who had been taught the learning from context strategies (Carnine et al., 1984; Sampson et al., 1982; R. M. Schwartz & Raphael, 1985; Sternberg, 1987). Kuhn and Stahl concluded that this could suggest "that students benefit as much from practice in deriving words from context as they would from instruction in either a specific set of strategies or a list of clues" (Kuhn & Stahl, 1998, p. 129). This suggests that additional

reading on its own could be enough to help adolescent readers build their vocabulary if they spontaneously employ their own strategies to derive word meanings from context.

Another vocabulary intervention that has shown improvements in generalized reading comprehension was a randomized controlled trial by Clarke et al. (2010). Their study compared four groups of 8–9-year-old students whose reading comprehension (as measured by *Neale Analysis of Reading Ability II* (Neale, 1997)) was lower than their reading fluency (as measured by the *Test of Word Reading Efficiency* (Torgesen et al., 1999)). The first group was taught text comprehension strategies (a combination of inference training, metacognition and reciprocal teaching) (TC); the second group had an oral language program that included vocabulary training (OL); the third group was taught an integrated program of both text comprehension strategies and oral language (COM); whilst a fourth group was the control/waiting list and did not receive any of the additional teaching. There were four data points: a pre-test, before the intervention started; a mid-test, after 10 weeks of intervention; a post-test after 20 weeks of intervention; and a delayed follow up approximately 11 months after the intervention finished. All the intervention groups, when compared to the control group, made significant gains on the *Wechsler Individual Achievement Test II* (Wechsler, 2005)of reading comprehension, which used a mix of genres and lengths of text to be read aloud and then open-ended comprehension questions to be answered orally, with the relative gain of the OL group increasing at the final time point. Relative gains on the *Neale Analysis of Reading Ability II* (Neale, 1997) test, which required children to read a passage of text aloud and then respond orally to open ended comprehension questions, were not significant until the final time point, 11 months after the intervention ended, and then only for the OL group, although the effect was close to statistical significance for the COM group. A mediation model showed that the variations in the vocabulary scores at Time 3 could account for the effects of the COM programme and partly account for effects of the OL programme

on the reading comprehension measure in the *Wechsler Individual Achievement Test II* (Clarke et al., 2010). Having a final data point so long after the intervention was completed was unusual in reading intervention studies at that time and it may be that it was this time lag that was crucial for showing improvements on standardised comprehension measures that other vocabulary interventions, discussed earlier, have not shown. These results point to vocabulary being a key focus for improvement in the long-term.

Based on the study above, Clarke et al. (2017) ran a second study that focused on struggling readers at transition to secondary school. These struggling readers were identified by drawing participants from students who achieved below the expected level 4 at KS2 and had a standard score of 91 or below on a single word reading test. There were three groups: 1) reading intervention (RI) – which targeted decoding; 2) RI and the oral language programme from Clarke et al. (2010) (RI +C); and 3) a wait list control group who began receiving an intervention after the third data point (20-week post-test). There were four testing points: a pre-test; a mid-test after 10 weeks of the intervention; a post-test after 20 weeks of the intervention; and a delayed post-test 40 weeks after first intervention started. It was found that the intervention RI+C, which had been run before in primary schools (Clarke et al., 2010), was much more complicated to run with secondary school students. Researchers experienced difficulties in both identifying enough suitable participants in all the schools and there were also high attrition rates. No fidelity of implementation data was collected so it was not possible to know whether the wait list control group had received any school-based intervention or whether they were business as usual (no intervention). The final difficultly was that the interventions were delivered by teaching assistants at the secondary schools who were judged by the researchers to be less well-equipped to benefit from the training, as they had less experience with both the teaching of phonics and with delivering interventions than their primary school colleagues from the previous study (Clarke et al., 2010). Neither of the

intervention groups, RI or RI+C, showed significant gains in word reading, compared to the control group in either of the post-tests. There was a small but significant gain for the RI+C group at the immediate post-test for reading comprehension, compared to the RI and control groups, but this was not maintained at delayed post-test. The RI+C did make gains, relative to the RI and the control, on a latent variable of vocabulary knowledge at both post-tests. The RI+C group also made gains on taught and non-taught vocabulary on a bespoke test, compared to the RI and control group, at the immediate post-test, though this was not maintained at the delayed post-test for the taught vocabulary, although this may be explained by the fact that the control group had received the intervention themselves by time four. Perhaps the most surprising result was that both intervention groups showed gains on a standardized vocabulary measure, compared to the control group, despite the RI group not receiving any vocabulary instruction. Clarke et al. suggest that this could be due to incidental word reading from the practice texts in the RI lessons. This study shows that the vocabulary component in the RI+C group seemed to bring about greater gains than were produced for the RI group but primarily it shows how hard it is to implement interventions in secondary schools and therefore the results must be treated with caution.

G. Elliot et al. (2024) conducted a randomised controlled trial with students in the first two years of their secondary education who, like in the Clarke et al. study (2017) above, had poor reading comprehension scores but adequate word reading ability (Snowling et al., 2009). Students received either: an oral language intervention (OL), where the components were vocabulary, reciprocal teaching, figurative language, spoken narrative; or a text-based intervention (TB), where components were metacognition strategies, reciprocal teaching, inferencing, and written narrative; or were the control/waiting list. A battery of reading tests was administered before the intervention started (pre-test) and at the end of the 8-week intervention. The intervention content for both the OL and TB groups was based on the book

The Last Soldier by Keith Grey and all intervention sessions started with re-attribution training (Berkeley et al., 2011). There were very large effect sizes of the intervention for both groups on the reading comprehension measure (OL 0.67 and TB 0.53). Both groups also showed improvements in a vocabulary standard score. This may be surprising as the TB intervention did not include a vocabulary component, however, as with Clarke et al. (2017) and Kuhn and Stahl (1998) this could be attributed to incidental learning from the intervention materials, that is, a practice effect. Unlike Clarke et al. (2017), this study was able to show gains in a secondary school context. A possible reason for this could be that the intervention was delivered by the lead researcher, who was a specialist trained teacher, and so avoided the issue that seemed to arise in the 2017 study with delivery by teaching assistants.

As has been suggested, for some of the intervention studies above, some of the gains in vocabulary and reading comprehension scores could have come from the practice of reading itself rather than any specific strategies or teaching (Clarke et al., 2017; G. Elliott et al., 2024; Kuhn & Stahl, 1998). Another study with secondary-aged students (Westbrook et al., 2019), provides some further evidence to support this possibility. In this study, conducted in 10 secondary schools, two whole texts (e.g. novels or playscripts) were read aloud in English lessons over 12 weeks. Usually when reading a text in class there would be frequent pauses in reading to complete activities or tasks linked to the text, for example a diary entry written in the voice of a character in the text, or some language analysis. However, in this study the lesson time was spent just focusing on reading, without stopping for any analytical or other activities to be completed. This was called 'fast reading' (FR). In one group in each school (FR +T) teachers received training about cognitive reading processes and pedagogical reading strategies (e.g. reading aloud, inference making, guided reading and graphic organisers) and in the other group (FR), the 'control', there was no additional training for teachers, but they were told to conduct 'fast reading' with their classes. It was expected that

the FR+T group would show some benefit of the teacher training but instead both groups showed an increase in performance on standardised tests of comprehension for all students. Of particular interest was that poorer readers in both groups (who had a reading age at least 12 months behind their chronological age) made significantly greater progress, an average of 16 months compared to an average of 9 months, than the other readers in both groups. Westbrook et al. argue that it was the uninterrupted reading that positioned the poorer readers as good readers, that is they were enabled to access the texts, through the shared reading aloud, in a way that good readers would usually access them. In this study the act of reading itself seemed to enable the progress.

Overall, findings from intervention studies are mixed. While interventions that focus on a particular reading component or strategy tend to show improvement for that reading component or strategy, improvements on standardised measures of reading comprehension are relatively rare. A focus on vocabulary or oral language has been shown to improve reading skills to some extent (Clarke et al., 2010; G. Elliott et al., 2024), but these effects have not always been replicated in shorter interventions or with adolescents. There was some suggestion that intervention methods increased depth of vocabulary knowledge, which led to improvements in reading comprehension (Sampson et al., 1982). There was also some evidence that the practice of reading itself could be the source of reading improvements, whether through 'fast reading' in a lesson (Westbrook et al., 2019), through the materials used in interventions (Clarke et al., 2017; G. Elliott et al., 2024; Kuhn & Stahl, 1998) or through the impact of independent reading on a delayed post-test (Clarke et al., 2010).

Whilst vocabulary emerged as an important component in successful reading, reading experience has emerged in the consideration of intervention studies as a key area. In order to further explore and understand the relationship between reading comprehension skills, vocabulary knowledge and reading experience, this review will now consider two hypotheses

that describe these components of interest. There will first be an overview of the LQH (Perfetti & Hart, 2002), which describes the different aspects of word knowledge and the links to skilled reading and reading experience. This will be followed by an overview of the LLH (Nation, 2017), which describes the type of reading experience that leads to skilled reading.

2.3 Lexical Quality Hypothesis

Perfetti and Hart's LQH (2002) proposes that a reader can efficiently process a word if they are able to access high quality representations of three components of that word: its orthography, phonology, and semantic information. If any of these separate components is not well specified in the reader's lexicon, then the representation of that word will be of a lower quality and the reader's access to it in the lexicon, and hence their comprehension of it, will be compromised (O'Connor et al., 2019). Readers differ in the quality of the representations in their lexicons and skilled readers have more high-quality representations than less skilled readers (Perfetti & Hart, 2002). Low-quality representations of words can cause a delay in processing. This delay was seen in the poor comprehenders in the semantic priming study (Henderson et al., 2013), outlined earlier in section 2.1, where poor comprehenders struggled with subordinate meanings. It was also seen in the semantic processing study (Nation & Snowling, 1998), also outlined in section 2.1, where poor comprehenders struggled with low frequency exception words. Delay in processing a word is important to comprehension because during the delay there are fewer processing resources available for comprehension of the message or meanings of the whole text. Perfetti and Hart (2002) argue that more highly skilled readers resolve confusion more rapidly than less skilled readers who experience longer delays, for example taking longer to access or inhibit subordinate meanings due to having less practised, less secure, less precise or less flexible semantic representations.

Participants in Perfetti and Hart's research (2002) were divided into one of three equally populated groups, less-skilled, average, or more skilled, based on their score on the Nelson-Denny comprehension test (M. J. Nelson & Denny, 1973). This comprehension test consisted of seven reading passages, taken from humanities, science and social science textbooks, with 5 multiple choice questions after each passage. A range of measures was then used to test phonological, orthographic and semantic knowledge. Phonological skills were measured using a phoneme elision task, a pseudoword reading task and a single word reading task. Orthographic skills were measured using a spelling discrimination task, a homophone choice task, the pseudoword decoding task and the single word reading task. Semantic skills were measured using the homophone choice task, the single word reading task and a vocabulary test. Separate factor analysis for each of the groups, for both speed and accuracy (speed was number of items answered on the vocabulary test within two minutes, accuracy was the number of correct vocabulary test answers from items attempted), indicated that readers who scored in the lowest third for reading comprehension were found, across the range of measures designed to test phonological, orthographic and semantic knowledge, to depend more on phonological and semantic knowledge (than orthographic) and to have less well integrated orthographic knowledge, than the more skilled readers. One explanation for this deficit of integrated representations of words in the lower skilled group, using the LQH, could be that the less skilled group had less reading experience than the more skilled groups because, without as much exposure to the written form of a word (orthography), the less skilled readers had fewer high-quality orthographic representations of words, leading to an impact in the overall lexical quality of their word knowledge.

A cross-sectional study of 247 primary school children, from first to fourth grades, in Germany (Richter et al., 2013) sought to test how the three lexical representations of LQH acted together in developing readers. A series of tasks was designed around the idea of an

alien, named Reli, who wanted to learn about earthlings' language. Phonological ability was assessed through a pseudoword comparison task; orthographic knowledge through a lexical decision task; and semantic knowledge through a categorisation task. Reading comprehension ability was measured through a standardized test, *ELFE 1–6* computerized version (Lenhard & Schneider, 2006). All the components of lexical quality, phonological, orthographic and semantic, correlated with reading comprehension skill. Using structural equation modelling, Richter et al. suggested that the causal order in their model, of the three types of lexical representations, in developing readers, revealed that phonological and orthographic representations influenced reading comprehension skill via semantic representations. This means that the relationship between the three representations is "asymmetric" (Richter et al., 2013, p. 430). High quality representations of meanings seemed to depend on phonological and orthographic representations, but the reverse was not true. This suggests that, as would be predicted by the SVR, even if phonological and orthographic representations are good, a deficit in semantic representations could cause problems in reading comprehension. Therefore, whilst, as shown earlier in section 2.1, decoding ability (phonological and orthographic representations) is essential to the ability to read, once decoding is mastered, if semantic knowledge does not continue to build, through reading experience, then according to the LQH, reading comprehension ability will not progress.

Another study tested whether the relationships between the different representations specified in the LQH, with reading comprehension ability, would be the same in English as a first language and English as an additional language learners (O'Connor et al., 2019). Grade 5 English as a first language and English as an additional language students were tested for: phonological awareness, using the *Auditory Analysis Skills Task* (Rosner & Simon, 1971); orthographic processing, using the Spelling subtest of the *Wide Range Achievement Test–Revised* (Wilkinson, 1993); semantic knowledge, using an expressive vocabulary measure

(Biemiller & Slonim, 2001); listening comprehension, using The Listening to Paragraphs subtest from the *Clinical Evaluation of Language Fundamentals–Third Edition* (Semel et al., 1995); and reading comprehension using both expository and narrative reading passages ability (MacGinitie & MacGinitie, 1992). Those students who scored highly on the reading comprehension measure also scored highly on the phonological, orthographic, semantic and listening measures, as would be expected according to the LQH and the SVR. Students who scored poorly on the reading comprehension measures tended to have difficulties in all component measures. This pattern was followed by both first language and additional language learners.

But it is not just different readers, with differing levels of representations, that can have an impact on the varying levels of lexical quality: the LQH also acknowledges that "it is the functional identifiability of a word (not the reader's skill in absolute terms) that is critical, and this can vary for the same word across readers and for the same reader across words" (Perfetti & Hart, 2002, p. 194). A word's functional identifiability can be affected by its raw frequency (a measure of how often a word appears) as well as any multiple mapping in any part of the three constituents that make up lexical quality. Multiple mappings could include: multiple orthographical representations for a single phonological one (e.g. seed, cede); multiple phonological representations for a single orthographic one (e.g. bass); or multiple semantic representations for words that have identical phonological and orthographic forms (e.g. count) (Hoffman et al., 2013). Any of these divergencies from singular relationships between the three constituents causes what Perfetti and Hart refer to as "confusion" (Perfetti & Hart, 2002, p. 195) as there is a delay, even if it is momentary, in matching or mapping the word across the constituents (Hoffman & Woollams, 2015).

The LQH describes how all three aspects of lexical quality are important to the efficient processing of words and therefore to reading comprehension skill. Perfetti defines

the building of efficient processing as depending on practice (2007), emphasising the importance of reading experience to the creation of high-quality representations. He extends the components of the LQH to include morpho-syntactic knowledge (how words are structured in different meaningful parts and how they function in grammar) as well as the original three components of orthography, phonology and meaning. In addition, there is a fifth feature, the "binding" together of the four components securely and coherently. This coherence is built through the constituent parts becoming strongly associated with each other, something that depends on practice (Perfetti, 2007; Reichle & Perfetti, 2003).

Whilst the LQH makes it clear that reading practice is essential to the improvement of the quality of representations in a reader's lexicon, some studies, that will be reviewed later in Section 2.5, suggest that certain types of reading experience are more strongly correlated with reading ability than others (Mar & Rain, 2015; Martin-Chang et al., 2020; McGeown et al., 2015; Torppa et al., 2020). To understand why different reading experiences may have different impacts, this review will turn to consider the LLH (Nation, 2017).

2.4 Lexical Legacy Hypothesis

Nation's LLH (2017) offers an explanation for the way that novice readers, after they have mastered decoding, become skilled readers. As described by the LQH (Perfetti & Hart, 2002), repeated encounters with words build higher quality lexical representations. The LLH argues that it is repeated experiences of words in meaningful and diverse contexts, rather than just repeated experiences per se, that build the higher quality representations needed for skilled reading. Frequency has long been used as a strong indicator of whether or not a word is likely to be known (Anderson & Freebody, 1981) and therefore how difficult a word is to process (Anderson & Freebody, 1982; Balota et al., 2004; Kintsch et al., 2005; Perfetti & Hart, 2002), since words that have been encountered many times will be more familiar and hence recognised more quickly. Traditionally, word frequency has been measured by taking

the number of times a word appears in a corpus and then calculating the normalised frequency - frequency per million (fpm) (number of occurrences of word / number of words in the corpus x 1,000,000 = fpm). For many years, word frequency was used across a wide range of psycholinguistic studies as a proxy for word difficulty, familiarity and presumed processing speed (Morrison & Ellis, 1995; Starr & Rayner, 2001). However, in 2006, Adelman's seminal study questioned whether word frequency was the best predictor of word processing time and used a measure of contextual diversity, operationalised as the number of different documents in which words appeared in a corpus (Adelman et al., 2006). Adelman et al.'s study (2006) used mean response times in word naming and lexical decisions from six pre-existing data sets (Balota et al., 2007; Balota & Spieler, 1998, 1999; Spieler & Balota, 1997) and used three corpora (Brown, 12th grade portion of LSA/TASA, and the written portion of the British National Corpus (BNC)) to calculate word frequency and contextual diversity for the words. Regression analyses showed that both high word frequency and high contextual diversity were associated with faster response times. However, while contextual diversity accounted for unique variance across all analyses, word frequency, only accounted for unique variance across six analyses, and furthermore these were negative effects, with high word frequency leading to slow response times. Adelman's paper had a significant impact on the field as researchers realised that it is diversity of reading experience (represented by the contextual diversity measure) that confers a processing advantage, rather than simply the number of times a word is encountered (represented by the frequency figure) (Adelman et al., 2006). Since 2006, there has been a large number of studies investigating contextual diversity and related concepts (e.g. semantic diversity) to try to understand exactly what it is about this measure that impacts human language processing and why (Johns et al., 2016; Joseph & Nation, 2018; Norman et al., 2023; Pagán et al., 2019; Pagán & Nation, 2019; Rosa et al., 2017).

The importance of contextual diversity to word processing can be explained by seeing, via the LQH (Perfetti & Hart, 2002), that each experience with a word is an opportunity to increase lexical quality, especially semantic (Share, 1995). Experiencing a word in a variety of meaningful contexts, through its co-occurrence with different words and different situations, can create a more precise, flexible and nuanced understanding of that word's possible meanings or, as Nation describes it in the LLH, there "might be a measurable legacy that follows from reading experience, where instances with words in meaningful text brings about differences in lexical quality" (Nation, 2017, p. 3). For example, if a reader only ever experiences the word 'trifle' in texts that are about cooking or food then the word will probably only be known as a cold dessert. If, however, the reader encounters 'trifle' in other texts, where it is used to mean something of little value or importance, then a second meaning of the word is experienced and a more nuanced understanding, or a higher quality semantic representation, is built of the same orthographic and phonographic representations. The same can be said of encounters with 'crest', whether it is known as the top of a mountain or wave, a part of the head on a bird or animal or in the heraldic sense as an emblem, all meanings that would be likely to occur in quite different contexts. These examples show that diverse experiences with words are needed to build high quality nuanced representations of different possible meanings, and this can only be achieved through encountering the words in diverse, rather than similar, contexts. This can explain why reading experience emerges as an important component in skilled reading comprehension and why reading experience is an important area of research.

The next step in the LLH, from understanding that diverse reading experience with a word seems to lead to a greater increase in lexical quality and therefore potentially improved processing of that word, is to consider the relationship in reverse. Successful processing of a word depends on "the learning opportunities afforded by an individual's lexical experience"

(Nation, 2017, p. 3). That is to say, the processing or reading ability of any reader depends on that reader's past reading experience of the words currently in front of them. While the word "sea" may be fairly high in frequency, if a particular reader has never encountered it before then they will be slow to process it. In contrast, a word such as "episiotomy" may be rarely encountered by many but very familiar to someone who is pregnant or who has recently given birth. Therefore, it is critical to consider whether a reader's prior language experience allowed them to build a sufficiently high-quality representation of the word currently being read. This leads to a consideration of what types of reading experiences are optimal and what readers' reading experience needs to contain, in order for them to be skilful readers of certain texts. Whether a particular genre or mix of genres provides reading experience that is superior for improving reading skills will be covered in the next section of this review.

2.5 Reading Experience

Reading experience emerged as an important component of reading comprehension skill in section 2.2 above. This section will consider studies that have focused on the relationship between reading experience and reading skills. The use of an author recognition tests (ART) and other reading experience measures will then be explored. Finally, an overview of the relationship between reading skill and different genres of reading will be given.

Stanovich's seminal paper "Matthew effects in reading" (1986), proposed that the domain of reading has "rich-get-richer" and "poor-get-poorer" effects (Stanovich, 1986, p. 382) through a reciprocal relationship between reading ability and reading experience. That is, more proficient readers are likely to read more and learn more from their reading and, by reading more and learning more, they become even more proficient readers. For less proficient readers the reverse is true. Not all studies find that the relationship between reading experience and reading skills is in the same direction. Using data from a twin study, which

allows for direction of causation modelling of cross-trait cross-twin correlations between observable characteristics, by testing competing models (reading experience and reading ability), van Bergen et al. (2018) found that in children aged 7 the causal relationship ran from reading ability to print exposure, suggesting that, in this age group, it is children's reading ability that determines how much they choose to read, rather than vice versa. Torppa et al. (2020) also showed an association between leisure reading and reading skills from ages 7 to 16. Whilst for younger readers it was comprehension ability that predicted leisure reading, for the older readers it was leisure reading that predicted reading comprehension and it was the reading of books, rather than other reading materials (e.g. newspapers, magazines, emails and blogs), that promoted reading comprehension most strongly. Torppa et al. concluded that "for comprehension of continuous printed text, book reading was a superior predictor" (Torppa et al., 2020, p. 14). The study did not suggest why book reading was superior, but it could be that it provides the more diverse contexts, outlined as important by the LLH (Nation, 2017), within which to encounter and re-encounter vocabulary. Fiction can have a greater diversity of places, times, and actions than other genres and therefore the diversity of contexts within which words appear is greater, or maybe the reading of books leads to reading a greater quantity of text.

Measuring reading experience accurately can, however, be difficult. Self-report questionnaires, that ask participants to report on their own reading habits and experiences are one means of assessing reading experience. This method can, however, be affected by the perceived social desirability of reading which can cause respondents to inflate or deflate the amount of time they report that they spend reading (Stanovich & West, 1989; West et al., 1993) and it also depends on accurate memories of time spent reading. Another method is to ask participants to keep reading logs or diaries to record reading practices but these can also be affected by inflated reporting due to social desirability (Moore & Gordon, 2014) and can

also be an inaccurate reflection of actual reading activities due to the high level of participant cooperation required and the need for retrospective estimations of reading time, which can be difficult to accurately recall (Bisson et al., 2012; West et al., 1993). Checklist measures of print exposure such as ARTs, magazine recognition tests (MRT) and title recognition tests, that work on the assumption that knowledge of authors, magazines and book titles reflect reading experience, are thought to be a more accurate and objective measure (Cunningham & Stanovich, 1997) as the inflation of reading time due to social desirability is avoided and there is not the same requirement to accurately recall time spent reading. Respondents cannot inflate their answers because any incorrect identification of authors or titles are deducted from their total score.

ARTs can be difficult to create because there needs to be a balance between authors being well known enough for participants to have realistically had a chance of reading something written by them, but not too famous to elicit recognition without having read any of their works. There are also issues with lists having to be updated, as author popularity rises and falls over time, especially for contemporary authors. Some lists have also attempted a balance between authors likely to have been read in school or college, as part of the curriculum, and authors that are more likely to represent reading for pleasure outside of school or college (Acheson et al., 2008).

Stanovich and West (1989) conducted two studies with undergraduate students in order to explore the links between phonological processing, orthographic processing and reading experience using different measures of reading experience. Their first study used an ART, an MRT, and a reading and media habits questionnaire as measures of reading experience. Three spelling measures were used to measure orthographic processing skill. The ART was the only measure of reading experience that correlated with the spelling measures and there was a greater association between reading experience, as measured by the ART, and

the spelling of exception words, that is words that do not follow common phonological patterns. This supports the finding in the Nation and Snowling (1998) study discussed above, that found a link between comprehension skill and knowledge of exception words, that is words that do not follow common phonological patterns. The Stanovich and West (1989) study showed that the ART was a superior predictor of reading ability compared to the MRT and the questionnaire. Their second study used a much wider range of measures: standardised reading measures (word identification and passage comprehension); spelling tasks; phonological processing tasks; and orthographic processing tasks. Again, as in the first study, the ART was highly correlated with other variables and was a "unique predictor, consistently accounting for additional variance in word recognition skills" even after phonological ability had been accounted for (Stanovich & West, 1989, p. 410). Both the word reading and spelling measures contained words that had very low frequency, and therefore errors were inevitable if they were unknown by some participants. Stanovich and West concluded that this was because the ART, as a measure of reading experience, represented an overall word (vocabulary) knowledge that was needed for these very low frequency words. This argument, about the importance of vocabulary knowledge, has been supported by the findings of several subsequent studies (Beech, 2002; Cunningham & Stanovich, 1990; Sénéchal et al., 1996; Stanovich & Cunningham, 1992).

It could be argued that ARTs do not actually measure reading experience but rather more general knowledge about books and authors. To investigate this, Martin-Chang and Gould (2008) tested undergraduate students using an ART with an additional question to differentiate between primary and secondary reading experience. When participants indicated that they recognised an author's name, they were asked whether they had actually read anything by that author (primary print exposure) or whether they just recognised the name as being an author (secondary print exposure). The results indicated that it was primary print

exposure, that is the act of reading itself, rather than secondary print exposure, knowledge of information about reading, that was the stronger predictor for performance on reading measures. Many other studies have used an ART as a measure of reading experience and shown a link to reading skill using a range of reading proficiency measures (Choi et al., 2015; Cipielewski & Stanovich, 1992; Mano & Guerin, 2018; Martin-Chang & Gould, 2008).

A further development, in the study of the relationship between reading experience and reading skill, has been in studies that differentiated between different genres of reading. McGeown et al. (2015), in a study with older children and adolescents, found a positive correlation between reading fiction and all their reading skill measures (word reading, comprehension, summarisation and reading speed) but correlations were not as strong between the reading skill measures and reading non-fiction, text books or digital reading (such as social media posts), with no positive correlations between reading speed and the non-fiction categories of reading. Another study by Mar and Rain (2015) also found that it was scores on their fiction ART, but not their non-fiction ART, that correlated with higher scores when testing components of reading ability (e.g. synonyms, sentence completion and reading comprehension).

A study by Pfost et al. (2013) expanded the genres of reading experience and, using results from a student questionnaire, put participants into five different profiles of readers: 1) those who read a broad variety of texts; 2) those who read exclusively online; 3) those who read online but also read some print; 4) those who read traditional print only; and 5) those who avoided reading print. Those in profiles 1, 3 and 4, whose reading included narrative texts and books, were more likely to score highly for vocabulary and reading comprehension. In contrast. those in profiles 2 and 5, who read mostly online or avoided print, had lower scores for vocabulary and reading comprehension.

While questions remain about the direction of causality, reading experience, especially fiction or book reading, emerges as having an important relationship with reading skills. Reading experience's role as an important component can be understood through the LQH (Perfetti & Hart, 2002) and LLH (Nation, 2017), because it provides practice with words and the opportunity to encounter words in diverse contexts, possibly offered to a higher degree by fiction texts, that allows the quality of their different lexical representations to build within a reader's lexicon.

One concern or possible limitation to this relationship, between fiction reading experience and measures of reading skill, is whether the types of reading texts that have been used in the reading measures themselves could have impacted the findings. In Torppa et al. (2020) the standardised comprehension test used for Grades 1-6 was a fiction story, for grades 7 and 9 it was stated that a similar test was used, although the genre was not specified. McGeown et al. (2015) used the *YARC* (*Secondary)* standardised test (Stothard et al., 2010) and participants received at least one factual and one fiction passage. The study by Mar & Rain (2015) used comprehension questions take from the Student Achievement Test, used by many American universities for admission. Whilst the original measure used one fiction and two non-fiction passages, the fiction passage questions were removed from the results in order to improve the test reliability. Pfost et al. (2013) designed different reading comprehension tests for the three different grades 5, 6 and 7 but the genre of the texts was only specified for the Grade 7 test (two expository and one narrative). It is difficult to ascertain therefore, whether or not the genre of the text in the test had an impact on any results. Careful consideration of the genre or register of the text used in the measure of reading comprehension skill should form part of any research that looks at the relationship between reading experience and reading comprehension.

The relationship between reading experience and reading skills is complex. In younger readers it seems that the direction of the relationship is from reading skill to reading experience (Torppa et al., 2020; van Bergen et al., 2018), but as readers get older and grow in proficiency the causal direction changes so that more reading experience leads to better reading skills (Torppa et al., 2020). ARTs have been found to be the reading experience measure that correlates most closely to reading skill levels (Beech, 2002; Choi et al., 2015; Cipielewski & Stanovich, 1992; Cunningham & Stanovich, 1990; Mano & Guerin, 2018; Martin-Chang & Gould, 2008; Sénéchal et al., 1996; Stanovich & Cunningham, 1992; Stanovich & West, 1989), probably due to its more objective nature as a measure. Studies that have differentiated between genres of reading experience, have found that fiction book reading has the closest relationship to reading skill (Mar & Rain, 2015; McGeown et al., 2015; Pfost et al., 2013). However, the genre of the texts used in the measures of reading skill are often not accounted for or even specified, so it could be that the texts used in the test of reading skills themselves are impacting the findings.

Moving on from this overview of the relationship between reading experience and the ways that reading experience is measured, this review will now turn to research on reading habits and practices. Reading experience, including the genres that make-up that reading experience, has been shown to have a strong relationship with reading skill, the next step then is to understand what the research says about children's and adolescents' current reading habits and practices.

2.6 Reading Habits and Practices

The research into the reading habits and practices of adolescents is mainly focused on independent reading, also known as reading for pleasure. Independent reading's potential link to educational outcomes has been the subject of much research (Clark & Rumbold, 2006; Cremin & Scholes, 2024; Mullis et al., 2017; Sullivan & Brown, 2015; Torppa et al., 2020).

The *Reading for Change* report by the Organization for Economic Cooperation and Development (OECD) (Kirsch et al., 2002) linked free time reading engagement positively to reading proficiency levels, although this is an associative relationship rather than one of causality. There is also evidence that reading for pleasure is linked to progress in maths (Sullivan & Brown, 2015) and some reports argue that it can be a lever for social change (Cremin & Scholes, 2024; OECD, 2021). There are however complexities noted, with results possibly being impacted by the reading measures used, the directionality of the analysis and the geographical locations of the studies (Cremin & Scholes, 2024). Some studies also showed reciprocal relationships with reading ability, with prior reading ability having an impact on reading amount, as also shown by the van Bergen et al. study (2018) outlined above, as well as reading quantity having a relationship to later reading ability (Cremin & Scholes, 2024; van Bergen et al., 2021).

Reading has been recommended, encouraged and listed as a requirement for English school curricula in government publications and in independent reviews commissioned by the Government (Department for Education, 2012, 2013b, 2023a; Rose, 2006). There are also campaigns and projects that promote and encourage reading for pleasure, for example World Book Day (https://www.worldbookday.com/) and the Open University's Reading for Pleasure website (https://ourfp.org/). However, reading for pleasure can be problematic as a concept, as it is open to many different interpretations and can be confusing to educators, who may feel that all reading, including reading that is part of the curriculum, should be somehow made pleasurable (Cremin et al., 2022).

Another important consideration is the content of reading experience. A UK and Republic of Ireland annual report *What and How Kids are Reading* (Topping et al., 2023), that is based on data collected by the school reading programme, Accelerated Reader, showed that for secondary school students in years 9-11 (ages 13-16), the age group that is the subject

of this study, the titles that were read most often were either fiction books that were likely to have featured on the school curriculum (e.g. *Of Mice and Men*, *Animal Farm* and *An Inspector Calls*), or were titles by children's and YA authors (e.g. *Heartstopper Volume 1*, *The Hunger Games*, *One of Us is Lying*, *It Ends with Us*, *Harry Potter and the Chamber of Secrets* and *They Both Die at the End*). In the 2023 National Literacy Trust survey, which annually collects data on the literacy habits of children and adolescents in the UK, fiction dominated as the most popular genre for 8-18-year-olds, with 50.7% reading fiction on paper and 25.8% reading it on screens (Clark et al., 2023). There is evidence that reading on paper is better for reading comprehension than reading on screens (Delgado et al., 2018; Kong et al., 2018; Salmeron et al., 2024), so this split between paper and screen is worth noting. The same 2023 National Literacy Trust survey found that enjoyment of and regular engagement in independent reading seems to be declining, with only 43% of 8-18-year-olds saying that they enjoyed reading, the lowest level since the survey began (Clark et al., 2023). However, there was actually a rise in enjoyment amongst 14-16 year olds, with 40.8% saying they enjoyed reading in 2023, compared to 32% in 2005 (Clark et al., 2023). A similar pattern was seen in levels of daily reading, with a small decrease overall in 8-18-year-olds, from 40.5% in 2005 to 36.9% in 2023 but in 14-16-year olds the level increased from 21.4% in 2005 to 25.7% in 2023 (Clark et al., 2023).

The picture for reading experience and reading habits is therefore complex. Much reading research finds a positive link between reading experience and reading skills (Acheson et al., 2008; Chateau & Jared, 2000; Davidse et al., 2011; Mol & Bus, 2011). However this relationship is not always consistent across socioeconomic groups and prior reading attainment groups (Cremin & Scholes, 2024; van Bergen et al., 2021). Whilst free choice and motivation are seen as key components to how much is read, the research reveals a complex picture where, again, socioeconomic background, prior reading ability and the content of

what is being read can have impact (De Naeghel et al., 2012; Miyamoto et al., 2018; Schaffner et al., 2016; Stutz et al., 2016; Troyer et al., 2019). In a context where time spent reading appears to be dropping (Clark et al., 2023), it seems, for adolescents that do choose to read, fiction and particularly YA fiction dominates. It is therefore difficult to have an accurate picture of whether or not adolescents are reading, whether the amount they are reading is enough and also whether the content of what they are reading is providing enough experience of words in diverse contexts in order to build the high quality lexical representations that they will need.

2.7 This Study

The focus of this study is the successful reading of examination texts in the English language GCSE exam. The texts have changed since the introduction of a new specification of the GCSE in 2015. The exams are now required to include literature and literary non-fiction and the texts are to be from each of the 19th, 20th and 21st centuries (Ofqual, 2013). To successfully comprehend the examination texts, students will have had to have been on a long reading journey. Beginning, as young readers, as described by the SVR, with successful decoding and word recognition abilities, combined with listening comprehension, they then move, as older readers, to a time when knowledge of the vocabulary of written language becomes more important than their listening comprehension. Reading comprehension skills increase through reading experience and the concurrent increase in depth of vocabulary knowledge. Lexical quality, of all the components of a word, is built through repeated encounters with that word, in diverse contexts, to enable fluent reading and free cognitive effort for comprehension.

Readers' abilities also depend on the text they are reading. If, as understood via the LQH and LLH, successful reading depends on high quality representations of the words in the text that is being read, then these will have had to have been built through prior reading

experience with those same words. Therefore, it is important to identify the kinds of words that are present in the exam texts, in order to identify what types of reading experience are optimal for providing repeated exposures to them. Or, in other words, to comprehend a piece of exam text students must have a legacy of reading experience that will match the vocabulary demands of that exam text. What the vocabulary demands of the exam texts are likely to be, what legacy of reading experience is required and how closely students' current reading experience matches this legacy are the subject of this research project.

The first set of data, presented in Paper 1 (Chapter 4), was collected from the available sample of past English language GCSE exam papers. A new corpus was created, called the exam text corpus (ETC) and representative vocabulary was identified and analysed. Reference corpora, large collections of text created to be representative of registers of language, were then used as comparisons to the newly created ETC, to identify which genres of texts were most likely to provide experience with the vocabulary found in the exams. The aim of this innovative use of corpus methods was to provide valuable insights into the nature of the vocabulary found in the exam texts, as well as to identify the genres in which the vocabulary was most likely to be found in students' previous reading experience.

The second set of data, presented in Paper 2 (Chapter 5) was collected from a large online survey, sent to students aged 16-18. Scores on a reading habits questionnaire, a reading experience measure (a uniquely designed ART), that differentiated between classic authors and YA authors, and a vocabulary test, were compared to the participants' English language GCSE grade. The aim with this set of data was to provide insights into adolescent reading habits and the particular genre of reading experience that had the closest relationship to gaining the higher grades in the exam.

The third set of data, presented in Paper 3 (Chapter 6), collected text from two types of student reading materials: independent reading and curriculum reading. These two small, specialist corpora were created to allow for exploratory analysis of the vocabulary in students' actual reading materials. The independent reading corpus (IRC) represented the reading carried out as part of a free choice reading homework for one year 10 class. The curriculum reading corpus (CRC) represented the materials given to year 10 students (aged 14-15) in one week of lessons. The most frequent words in each of the four parts of speech from the two corpora were analysed and comparisons conducted with the ETC from Paper 1 and other reference corpora. This data provided exploratory findings about the vocabulary in students' actual reading materials, rather than relying on reference corpora or textbooks to represent it and allowed analysis of the vocabulary found in these two different collections.

Presenting this thesis in a by papers structure, whilst having benefits in that it enabled early publication of results and early receipt of independent feedback through the peer review process, has meant that the three parts of this study have been written up more independently of each other than they would have been in a traditional thesis structure. However, the focus throughout all three papers on the English language GCSE as a real-life high-stakes context, and the emphasis, again throughout all three papers, on vocabulary and reading experience, due to the theoretical underpinnings of the LQH and LLH, also tie them together. This thesis, therefore, represents a coherent and sustained body of work on adolescent reading comprehension within the English educational context and also contributes to the wider literature on vocabulary, reading experience and assessment.

Chapter 3. Methodology

This methodology chapter will provide an overview of the methodological approach and design of the thesis as a whole. The details of the particular methods, within the three discrete parts of the project, are contained in the methodology sections of the individual papers, presented in Chapters 4, 5 and 6.

As outlined in the prologue to this study, the motivation for this research project came from an observed problem in practice, students struggling with the comprehension of vocabulary in exam texts. Using methodological pragmatism as the research paradigm for this study (Foster, 2024), means that the research methods were determined by the needs of the research questions. Whilst the three separate papers in this study have their own research questions, the overarching questions in this thesis were: what is the vocabulary challenge of the exam texts; and what reading experience provides the best match for encounters with the vocabulary. Methodologies were therefore required that could analyse vocabulary, measure vocabulary knowledge and measure reading experience. More specifically, the research questions required methods that would be able to: analyse vocabulary in the English language GCSE exam texts; identify likely genre sources for vocabulary from the exams; measure students' vocabulary knowledge and reading experience and compare them to attainment in the exam; and analyse vocabulary in real examples of adolescent reading experience.

The example outlined in the prologue, of different meanings of *hostel* causing confusion in an exam, illustrated the problem of vocabulary knowledge and context well, but this specific item of vocabulary was unlikely to feature again in exam texts. Close qualitative analysis of other exam texts, used in different past exam papers, may have revealed additional examples of vocabulary, like *hostel*, that could have potentially caused confusion or were particularly difficult, but each would, again, have been quite unlikely to reoccur. In the English literature GCSE, where the same texts are studied each year, a more qualitative close reading would probably have been helpful, identifying particular items of vocabulary that

would be useful to teach students to aid their comprehension of the literature texts. In English language however, with the texts being different in each exam each year, close study of past examples was not as appropriate. To address the observed problem, a more theoretical and generalisable understanding of the vocabulary challenges in the exam texts was needed. Individual students' gaps in vocabulary knowledge, or confusion around meanings, are also very specific and whilst examples of students' reading experiences are interesting to explore, they do not always repeat across multiple readers. Therefore, in order for the findings and conclusions of this study to be applicable to many students, teachers, classrooms, and exam texts, it would be beneficial if they were generalisable.

Two main methods were therefore selected for the research presented in this thesis. The first was corpus linguistics and the second was an online survey, that included measures of reading experience. The rationale for using these methods, and the tools and measures adopted within them, will be outlined below.

3.1 Corpus Linguistics

Corpus linguistics methods allow for the empirical and systematic analysis of a large body (corpus) or bodies (corpora) of texts. The methods used are both quantitative and qualitative (Biber et al., 1998; McEnery & Hardie, 2012). Corpus linguistics software is used to analyse language frequencies and collocations, through the use of frequency lists and concordances, identifying and exploring patterns of language use that it would not be possible for researchers to identify or explore manually in large bodies of texts (Evison, 2015). Corpus methods can be used to create, manage and study very large collections of text. For example, the English Web 2021 corpus, contains 52 billion words downloaded from the internet between October 2021 and January 2022.

Most corpus software automatically tags parts of speech for uploaded text, that is they will label each word in a corpus with its part of speech in context (e.g. noun, verb, adjective

or adverb). There are also tools available that will tag word senses, structural grammatical features (such as predicate or subject), as well as phrases, clauses and sentences (Anthony, 2013). Corpus linguistic tools can therefore be used to calculate total frequencies of individual words and, through the automatic tagging, total frequencies of types of words (e.g. adjectives) and other linguistic features. The frequency lists can then be used to identify items of interest, for example a frequent word, part of speech or multiword unit. These items of interest can then be accessed and explored within the contexts of the texts that make up the corpus through a concordance tool (Anthony, 2013). The ability to view examples of words or items of interest together in a concordance list then allows for the testing or generation of hypotheses (Evison, 2015). Corpus linguistics tools can, therefore, accurately produce large counts and identify all occurrences of any particular item. Bias can also be minimised, when selecting examples for further analysis, through the use of random sampling features. Qualitative analysis can then be conducted with the words or patterns that have been identified by the quantitative methods (Biber et al., 1998).

As a method, corpus linguistics can be used for many different purposes and by different types of users. For example, lexicographers can study empirical data to help create entries for dictionaries, applied linguists can create and study different registers and contexts of language use, and teachers can explore language use with language students, both as a learning and pedagogical tool. Large reference corpora can be accessed via corpus software and can be used to study language use in collections created to represent different registers. New corpora can also be created to study and analyse particular contexts or uses of language.

For the current study, corpus methods were used to create and analyse three corpora. The first corpus (the ETC) was made of all the available exam texts (at the point of the creation of the corpus), the second (the IRC) from a sample of students' independent reading, and the third (the CRC) from a sample of curriculum reading. These three collections were

the focus corpora of this study. Full details of the corpus methods used are given in the two papers in Chapter 4 and Chapter 6. Two key corpus linguistic methodologies, the generation of keywords to identify typical vocabulary and comparisons with other corpora, will be discussed in more detail below. First overviews of previous corpus research into both academic and non-academic language will be given.

3.1.1 Corpus Studies of Academic Language

Corpus methods have been used in previous studies to examine vocabulary in academic texts. These studies have been more prevalent in the higher education context. A multi-disciplinary academic word list was created by Coxhead (2000) who used corpus methods to provide a frequency-based list of core academic vocabulary for teachers, learners and researchers. This was followed by the academic vocabulary list (Gardner & Davies, 2014) that sought to provide a list of vocabulary that was common across academic disciplines, excluding commonly used and everyday language. Other corpus linguistic studies have looked at the differences in word use between disciplines, specific language within disciplines, and the use of multi-word bundles across disciplines (Hyland, 2008, 2017; Hyland & Tse, 2007).

There has also been some corpus research with language and vocabulary in schools, similar to the higher education studies outlined above. Secondary school vocabulary and phrase lists have been created for different subjects by Green and Lambert (2018, 2019), these were predominately aimed at English language learners. There have also been studies of language use in school resources, for example Monaghan (1999) analysed subject specific vocabulary in a set of secondary school maths resources. Language use in science has also been the subject of corpus studies, with the vocabulary from science textbooks analysed and found to be very demanding (Coxhead et al., 2010; Deignan & Love, 2019). The level of vocabulary demand was also assessed, in a separate study, across a range of different subjects

being taught in English to students at an international school in Germany (Coxhead & Boutorwick, 2018), where again the challenge was found to be high.

Whilst these studies have produced valuable data, providing lists of vocabulary for study and teaching, creating lists of words would not work for the English language GCSE because, as outlined above, each exam has different unseen texts, and so vocabulary will be very different from year to year. Instead, for the first part of this study, a new corpus of exam texts was created and used to identify a sample of typical vocabulary. This list of typical vocabulary then provided informative descriptive data. Using a large reference corpus, the frequency of these typical words in different genres was then identified, which provided likely genre sources for reading experience with these words. This is important in a study of reading and vocabulary knowledge, as it is through repeated encounters with words in diverse contexts, according to the LQH and LLH that the high lexical quality, needed for skilled reading, is built (Nation, 2017; Perfetti & Hart, 2002).

Most corpus studies of classroom materials have focused on, or sourced their data from, textbooks or commercial schemes of work (Coxhead et al., 2010; Green & Lambert, 2018, 2019; Monaghan, 1999). In the past, textbooks would have provided a fairly accurate source of data for reading in the classroom as they were the predominate classroom resource for reading. Textbooks are also a convenient and expedient way of sourcing data as they do not require any visits to different classrooms or individual participants. However, textbooks are no longer the main source of reading in classrooms, slides, online resources and online learning platforms are now as common, if not more so. Therefore, for the third part of this study, curriculum reading materials were sourced from an online platform used to share a variety of forms of classroom resources. While capturing the reading experience from this variety of modes and materials posed formatting and uploading issues, it nevertheless

provided useful data about the vocabulary that forms students' curriculum reading in today's classrooms.

3.1.2 Corpus Studies of Non-Academic Language

Corpus studies of non-academic texts have been used to study the language in texts that are more likely to be read independently outside the classroom, reading that is sometimes called reading for pleasure. For example, there have been studies of the language in children's picture books (Cameron-Faulkner & Noble, 2013; Dawson et al., 2021; Montag, 2019; Montag et al., 2015), which have examined the ways in which picture book language includes more complex vocabulary and syntax than general spoken language or child directed speech. There have also been studies of reading books for older children (e.g. Hsiao et al., 2022; Montag & MacDonald, 2015) which showed that more complex grammatical structures (e.g. passive relative clauses) were more frequent in the written language of books than in child-directed speech. Whilst there are reference corpora of books written for children (e.g. the Oxford Children's Corpus, and the children's reading subset of the Corpus of Contemporary American English), there are fewer corpora that cover adolescent reading specifically, especially older adolescents. Not only is this because this age group is understudied, but it is also more difficult to identify their typical reading materials because it can be hard to differentiate which genres may form adolescents' independent reading as they transition to more general adult fiction. The collection of examples of adolescent students' actual reading materials in the third part of the current study, addressed this gap in the research and provided valuable analysis into the content of independent reading experience for this age group.

3.1.3 Creation of New Corpora

As outlined above three new corpora were created for this study. Corpora which are created to represent a specific context can provide data and findings that are particularly

relevant to their purpose (Koester, 2015). Large reference corpora are representative through their size, specialised corpora can be representative of their specific situations (Koester, 2015; Lee, 2015; McEnery & Hardie, 2012; M. Nelson, 2015). It can also be a strength of small corpora that, as was the case with this project, the compiler was also the analyst, which enabled a close and detailed knowledge of the corpora and therefore detailed insights and findings (Koester, 2015). Clear compilation inclusion criteria and detailed records of the texts used to create the corpora were kept as they could be essential to later analysis (Reppen, 2015).

3.1.3.1 ETC. The ETC, the first focus corpus in this current study, was created in the corpus tool *Sketch Engine* (Kilgarriff et al., 2014). All publicly available English language GCSE exam reading files were downloaded from the exam board websites and the reading extracts copied onto plain text files. Plain text files were used because they do not include formatting codes and so this avoids any extraneous material being included in the corpus. For the same reason, question wording, introductory information, line numbers and glossaries were all removed so that their content would not be included in any analysis or calculations (M. Nelson, 2015; Scott, 2015). A detailed record of each text was kept so that the original source of any item of vocabulary (that is, the exam board, the specific exam paper, the date of publication, genre and author) could be identified if needed (Reppen, 2015). Sub-corpora were created within the ETC with texts placed as either fiction (15,000 words) or non-fiction (21,500 words) and in addition each text was also placed in either a 19th, 20th or 21st century sub-corpora for their date of production (11,000, 15,000 and 10,500 words respectively). Further details about this corpus are included in Paper 1 (Chapter 4).

3.1.3.2 Student Reading Records. Initially the intention was to track the reading of a number of individual students. The plan was for them to keep a reading record, of their reading in lessons and at home, by taking photographs of texts with their school tablet device.

These photographs would then have been used to create a student reading corpus. However, it was very difficult to get agreement from any students to take part in the study. The level of commitment, for an activity (reading) that it is not socially desirable for many in this age group, was seemingly too high. Even when students agreed to keep the photographic record, the effort required to both remember and actually take the photo seemed to be too difficult on top of other lesson activities and commitments. Alternative data collection methods therefore had to be found.

The alternative method used to collect samples of independent reading was to take the reading texts used for an independent reading homework task for one year 10 class and create the IRC. The alternative method used to collect curriculum materials was to download one week's worth of classroom materials, for a range of year 10 classes, from a school's online classroom platform. These curriculum materials were used to form the CRC. Similar formatting procedures were followed, to those used in the creation of the ETC, with texts converted to plain text files. Further details about the data and methodology for the IRC and CRC are included in Chapter 6 (Paper 3).

3.1.4 Keywords

Keyword lists were used for the first paper in this study (Chapter 4). Keywords are identified in *Sketch Engine* by a tool that compares the focus corpus (the one that is being studied) to a reference corpus and calculates which words feature more frequently in the focus compared to the reference (Evison, 2015; McIntyre & Walker, 2015). Keyword lists can be useful as a way of identifying, quantitatively, how a smaller, more specialist corpus differs from a general corpus because words will be identified that are significantly more frequent in the focus corpus than in a larger reference one (Kilgarriff et al., 2014). Therefore, keywords represent the specificity of the focus corpus (Evison, 2015).

The reference corpus used to create the keyword list for the ETC was the BNC. The BNC was chosen as the reference corpus as it included the widest variety of registers and genres of written and spoken language. After the keywords had been identified by the tool in *Sketch Engine*, searches for these same keywords, in the BNC, were then able to identify the source genres in which they most frequently appeared. These genres would then suggest the type of reading experience that would be most likely to provide encounters with these words. The BNC was used as the reference corpus for these genre searches because not only were source texts identified as from either written or spoken registers, but the written register is also further subdivided into imaginative and informative written text types.

In addition to the registers, the BNC is also subdivided into the much more specific David Lee Categories. After the publication of the BNC, Lee (Lee, 2001) created these categories because he felt that more specific genre categories would be of benefit to researchers. Lee took 'genre' to mean a culturally constructed type of text, as compared to 'register' which described linguistic patterns. Lee carried out the classification himself, which gave consistency but also means that it should be noted that the categorisations are subjective. The final 70 categories, of which 24 are different genres of spoken language and 46 are different genres of written texts, were decided upon to represent as far as possible the widest range of all the different types of texts in the English language represented in the BNC, but also with reference to categories used in other corpora to enable comparisons (Lee, 2001). Aston (2001), in his review of Lee's categories, notes that researchers should be cautious about treating the categories as representative as some have very small numbers of texts within them, but sees them as a useful method for generating hypotheses about genres of texts.

Further details about the methods used for this part of the research are given in Chapter 4 (Paper 1).

3.1.5 Corpora Comparisons

Whole corpora can be compared in *Sketch Engine*, to give a quantitative measure of similarity or difference between two different corpora (Kilgarriff, 2001). This was used to compare the ETC to a range of reference corpora in the first part of the study, reported in Chapter 4 (Paper 1) and to compare the IRC and CRC to the ETC and to a range of reference corpora in Chapter 6 (Paper 3). Comparing corpora in this way gives a numerical value for how alike two corpora are, based on the attribute chosen for the comparison. A score of one indicates the corpora are the same. The higher the score is above one the larger the difference between the two corpora for the chosen attribute. For this study the attribute selected was words, as the focus of the analysis was vocabulary knowledge. The numerical scores that indicate difference do not explain what the nature of the difference is or why there is a difference, but they do suggest which genres or registers of reading are most like the focus corpora. These were valuable findings in this study as they could give an indication of the types of reading experience that could provide encounters with the words in the exam texts.

Full details of the reference corpora used and how the data for the comparisons was calculated are in Chapter 4 (Paper 1) and Chapter 6 (Paper 3).

3.2 Reading Survey

An online survey (Appendix C), created using the survey tool *REDCap*, was used to collect information about respondents' English language GCSE exam grades, reading experience, and vocabulary knowledge for the second paper in this study (Chapter 5). The survey was granted ethical approval by the University of Reading's Institute of Education (Appendix N). Respondents consented to participate after reading the information at the beginning of the online survey.

3.2.1 Format and Structure

3.2.1.1 Online Survey. A digital survey was selected as the most appropriate data collection tool to gather data on student reading habits for several reasons. The first is that where a large number of responses is sought then a digital or online survey can cheaply and quickly reach large numbers of possible respondents (Heiervang & Goodman, 2011; Kılınç & Fırat, 2017). An additional advantage is that digital surveys are well placed to make use of expedient sampling (Kılınç & Fırat, 2017), particularly important in this study as the survey could then be distributed to whole cohorts of students at educational institutions that agreed to take part. Another advantage in choosing a digital survey are the benefits for data analysis. Responses are automatically available digitally for analysis, reducing the time needed for data entry and checking, and also reducing any loss of data that might occur during the entry process (Kılınç & Fırat, 2017). It is also possible to monitor responses in real time and to correct or resolve any technical issues if required (Evans & Mathur, 2018). Filters and skip patterns can be used, which can, again, reduce data loss and also streamline the respondents' experience (Evans & Mathur, 2018). Questions on a digital survey can be designed in a way that requires completion, although there are important caveats to this that must be followed to ensure that accuracy is not affected (to be discussed below). The digital format also lends itself to user-friendly design, all of which can promote participation and completion of responses (Evans & Mathur, 2018).

Another consideration in making the choice to use a digital survey is the perceived social desirability of reading, which means that respondents may be motivated to give false answers, in a face-to-face survey, by the desire to make a good impression on the interviewer. This is of particular concern in this survey because in educational institutions, reading is commonly promoted and seen as a particularly desirable activity and so, for students, this could be a sensitive topic to answer honestly. Where the topic of the survey is known to have

a social desirability or sensitive aspect, then a self-completion method, like a digital survey, is considered to be advantageous (Kılınç & Fırat, 2017; Saris & Galhofer, 2014).

3.2.1.2 Voluntary Participation. Voluntary participation is obviously essential for ethical reasons, and it has been shown that answers are more reliable when participation is voluntary (Chesney & Penny, 2013; Kılınç & Fırat, 2017). Whilst there are strategies that can be followed to encourage participation, which will be discussed below, it is still important to the accuracy of the data collected that all responses are voluntary. A digital survey, where the invitation can be issued and accepted or declined remotely, removes some of the possible compulsion that may be felt by students, due to the survey being sent to them via their school or college email address.

Whilst it is important that those who are invited to respond do so without pressure, it is also important to encourage as much voluntary participation as possible to avoid high levels of skew in the data through self-selecting participation (Kılınç & Fırat, 2017). It is also worth noting that there has been research showing that nonresponse bias or drop-out in terms of demographic variables are important to bear in mind when assessing results (Fry et al., 2017; Sax et al., 2003; Wolke et al., 2009).

It is also crucial for completion rates that the survey does not take too much time to complete. Evans & Mathur (2018) specify that surveys should not take longer than a few minutes and they also emphasise that, especially with mobile phone surveys, the formatting must fit on the device screens so that they are easy for the respondents to complete (Evans & Mathur, 2018). Although a post-payment lottery incentive was considered, to increase participation, there is inconsistent evidence about the effectiveness of offering monetary incentives (Frick et al., 2001; Porter & Whitcomb, 2003) and so this was rejected.

Kılınç & Fırat (2017) found, in their research on voluntary participation, that short data collection tools, good communication of purpose and importance, the option to be

informed about the results, and the assurance of safety were all key to increasing voluntary participation rates. These are all areas that were emphasised in the invitation and information sent to potential participants. Safety of personal information and the confidentiality of data can be of particular concern with digital surveys and arrangements need to be clear before the survey is started (Evans & Mathur, 2018).

3.2.1.3 The Question Order. The order in which the questions are presented to the participants required careful consideration as it can have an impact on the answers given, especially where there are obvious relationships between the questions (Saris & Galhofer, 2014). For example, if a question about something similar has been asked before, then it may affect a response through the "evenhandedness effect" (Saris & Galhofer, 2014, p. 157), that is the desire by respondents to appear consistent. There could also be an "anchoring effect" (Saris & Galhofer, 2014, p. 157) if the questions move from something familiar to something unfamiliar, that is the response to the familiar may affect the response to the unfamiliar. There could also be a "carryover effect" (p. 157) where questions about specific examples can affect later responses (Saris & Galhofer, 2014). However, attempting to separate questions about similar things and asking for responses in a way that tries to avoid any order effects, can then make a survey appear disjointed and confusing to respondents and increase the cognitive burden as they try to make sense of disparate requests (Saris & Galhofer, 2014). Therefore, whilst it is important to guard against order effects it is also important to present questions to respondents in a logical way that makes easy sense and therefore encourages completion (Iarossi, 2006; Saris & Galhofer, 2014).

The introduction to the survey is important because it can motivate participants to respond to the survey. In the introduction to the survey in the current study the authority of the survey was highlighted through making it clear it came from a university. There was also

a clear indication of how long the survey will take, *12 minutes*, designed to meet the timescales specified by Evans and Mathur explained above (2018).

After the introduction there was a short statement of ethics with the relevant contact details and then the consent agreement. Whilst it was important to motivation to proceed quite quickly to a question that directly related to the main topic of the survey, it was also important to realise that the first responses may not be as good as later responses as it takes a little time for respondents to become familiar and confident with the survey. Therefore, the survey in this project followed the recommendation to start with general questions (Saris & Galhofer, 2014) by starting with the demographic ones. After the demographic information, there was a section on respondents' attainment in the English language GCSE, including the year it was taken, to be able to separate the COVID cohort from the pre-COVID cohort. The next three sections, covering self-reported reading habits, the ART, and the vocabulary test will be covered in more detail below.

3.2.2. Self-reported Reading Measures

The self-reported reading habits questions, which can be found in the copy of the full survey (Appendix C), and covered time spent reading and genres of reading. Iarossi (2006) notes that respondents use different methods of recall when asked about frequency or time questions. To encourage respondents to directly recall and count-up time spent reading (episode enumeration), rather than using a rule, available or automatic estimation, then they need to be given a time frame that makes it possible to recall individual moments of reading and add them up. Blair and Burton (1987) showed that there needed to be fewer than ten individual episodes to be recalled during the time period to encourage this method of answering. In order to prompt responders, it can be useful to use question wording to aid recall or to construct a series of questions that create a landmark or reference point (Iarossi, 2006). One day was chosen as the time period for the initial question, as there were likely to

be less than 10 episodes of reading that needed to be recalled and counted. A series of questions was asked to help the responder create a reference point, and then to estimate an average amount of time spent reading (Saris & Galhofer, 2014). In the survey for the current study this started with:

> Minutes spent reading (at home and at school/college) yesterday: e.g.
>
> 1 hour = 60 minutes
>
> 2 hours = 120 minutes
>
> 3 hours = 180 minutes
>
> 4 hours = 240 minutes
>
> [_____] (minutes spent reading yesterday)

Followed by:

> Is [minutes spent reading yesterday] more, less or the same as the amount of time you spend reading on a typical day?
>
> More than usual [ ]
>
> The same as usual [ ]
>
> Less than usual [ ]

And finally:

> On average, how much time do you think you usually spend reading each day?
>
> This should include both reading at school/college and at home.
>
> [_____] (average minutes reading each day)

Minutes were chosen as the unit of time as it suggests the need for a more accurate rather than a general amount of time. An open format was given for the minutes as a closed form could provide "a guide" (Saris & Galhofer, 2014) and might, therefore, impact on the answer given. Even if the time period is just one day, the range of possibilities for time spent reading is large and so there would either have to be a very long list of possible categories,

which would be difficult to fit on one screen, or a few categories which would not provide very much nuance in the answers. N. Schwartz et al. (1985) found that if respondents were given a low range of time categories to choose from then they were likely to underreport the time they had spent watching television, whereas if they were given a high range of time categories then they were likely to overreport. Decisions about how to present any closed answer categories for this question could have led to inaccuracies in the estimates that responders gave. Using the open format meant that responders had to rely on their recall of episodes of reading, promoted by the question about reading time yesterday, and their assessment of whether or not this was typical, followed by an estimate of an average time spent reading each day. Whilst processing times for an open answer are not as efficient as for a closed answer, the higher quality of the answers for the open version meant that it was decided that it was the best way to format this section (Saris & Galhofer, 2014).

Acheson et al. (2008), in a study with college students, found self-reported estimates of time spent reading were not as reliable as an ART nor were they as reliable as a comparative question that asked students to compare their reading habits to their peers. Two comparative questions were therefore included in the survey: one asked, using a five-point Likert scale, the respondent to compare the amount of time they spend reading to other students; the second, again using a five-point Likert scale, asked the respondent to compare their enjoyment of reading to other students. Time was chosen as one comparative as it linked coherently to the previous questions, enjoyment was chosen as the second comparative as this question had the highest correlation with the standardised reading test and ART score in Acheson et al.'s study (2008).

3.2.3 The ART

The self-reported reading measures were followed by the ART, a copy of the ART can be found in the survey (Appendix C). An overview of ARTs as a reading measure has been

given in Chapter 2, the Literature Review, but will be summarised here briefly. ART scores have been found by many studies to correlate with reading skills (Beech, 2002; Cunningham & Stanovich, 1990; Sénéchal et al., 1996; Stanovich & Cunningham, 1992). ARTs present a list of authors and foils, in equal numbers, and ask participants to select any names that they know to be authors. A score is then calculated by taking the total number of authors selected by a participant, minus any foils that they have also selected (Davidse et al., 2011; Martin-Chang & Gould, 2008). In order to differentiate between two different types of reading experience, the author list for this survey's ART covered two genre areas: YA fiction authors and classic fiction authors. Full details of the ways in which the author names were selected and the foils created are given in Chapter 5 (Paper 2).

A list of author names for non-fiction as a genre was also considered, but because the range of the non-fiction genre is so great it was difficult to get a reliable shortlist of best sellers and, therefore, a reliable list of authors' names that could feasibly be expected to be recognised. There was also the problem that some bestselling non-fiction authors, particularly of autobiographies, would be well known generally and recognition of their name on a list may be due to their fame outside of being an author and therefore not a good indication of reading experience.

3.2.4 The Vocabulary Test

A vocabulary test was included in the survey (Appendix C) as vocabulary was expected to be a mediator between attainment in the English language GCSE and reading experience. A bespoke test was designed, using vocabulary from the ETC, as scores on this test would then more closely indicate knowledge of the vocabulary that was typical of the English language GCSE exam, rather than just general vocabulary knowledge. Words for the test were selected from the keyword list (see Appendix A), which represented vocabulary that was typical of the exam texts. The original context in which the keywords appeared in the

exam texts was also used in the test, as this would replicate most closely the vocabulary knowledge needed.

The concordance function of *Sketch Engine* was used to access the context of the keywords and appropriate examples were selected by the researcher. Taking each word in turn, from the highest scoring on the keyword list downwards, suitable choices were selected using the following criteria: 1) they represented a range of different parts of speech; 2) they appeared in a context that gave enough information to select a correct answer; 3) that they had a suitable synonym or short phrase that would work as a correct answer. For example, the second keyword, *nasally,* was not selected, as the contexts in which it appeared in the exam text were quite simple (e.g. *they talk nasally*). The third keyword, *thrill*, was also not selected as it only appeared in a short photo caption and as part of a proper noun in a reference to a Blue's song (*The Thrill is Gone*). The noun, *swimmer*, was not selected due to lack of synonyms for a correct answer. It was difficult to select verbs as there were fewer examples at the top of the keyword list. In total fifteen words were selected for the test (four adverbs, two verbs, four nouns, five adjectives), as this number had the potential to give a large enough range of scores, but would not take too much time to answer, an important consideration for survey completion rates.

Multiple choice options were selected that were the correct part of speech for the context and therefore could work grammatically. Options were also chosen that could make sense semantically. For example, the test presented the adverb *incessantly* in a context from the exam texts and then gave four possible adverbs as answers:

Select the word nearest in meaning to the word in italics

Having once got hold they never let go but struggled and wrestled and rolled *incessantly*.

continually [ ]

suddenly [ ]

freely [ ]

separately [ ]

The answers selected as correct by the researcher, for all fifteen items in the test, were checked in a short validation survey, circulated on social media (see Appendix E). The final vocabulary test is printed in Appendix C as part of the online survey and further details about its validity are given in Chapter 5 (Paper 2).

3.3 Summary

The focus of this study is the comprehension challenge presented by the English language GCSE. Using the LQH (Perfetti & Hart, 2002) and the LLH (Nation, 2017), vocabulary and reading experience were selected as the key components of interest, as vocabulary has been shown to be critical to comprehension skill, particularly for older readers (Braze et al., 2007; Ouellette, 2006; Tilstra et al., 2009), and reading experience the principal source of the building of high lexical quality (vocabulary knowledge). Methodologies were therefore required that could measure vocabulary knowledge and reading experience. More specifically, the research questions required methods that would be able to: analyse vocabulary in the English language GCSE exam texts; identify likely genre sources for vocabulary from the exams; measure students' reading experience and compare it to vocabulary knowledge and attainment in the exam; and explore vocabulary in real examples of adolescent reading experience.

The two main methods used to answer the research questions in this project were: corpus linguistics and an online survey, including an ART and a bespoke vocabulary test. As has been outlined above, corpus linguistics methods were used to identify and analyse vocabulary, a key component in reading, in three newly created corpora. Reference corpora were also used to identify the genres or registers of texts that were most like the newly

created corpora, and that were good sources for the keywords from the ETC. The online survey was used to collect new primary data from adolescents about their English language GCSE attainment, their reading habits, their reading experience and their vocabulary knowledge. Full details of the methodologies used for each of the three separate parts of the studies are contained in the papers that form Chapters 4, 5 and 6.

Chapter 4. Paper 1

A Corpus Study of English Language Exam Texts: Vocabulary Difficulty and the Impact on Students' Wider Reading (or Should Students be Reading More Texts by Dead White Men?)

Authors: Beverley Jennings, Daisy Powell, Sylvia Jaworska, Holly Joseph

The paper reproduced here was published in the *Journal of Adolescent and Adult Literacy* in January 2024
https://ila.onlinelibrary.wiley.com/doi/full/10.1002/jaal.1331#

Abstract

Students in England sit an important gateway examination in English at age 16. Major changes were made to this exam in 2017 resulting in more emphasis on the comprehension of unseen literary texts. This paper uses corpus linguistics methods to identify the kind of vocabulary encountered in these exam texts and compares it to vocabulary encountered in other sources of written language (classic literary fiction, biographies, poetry etc.). Results showed vocabulary in the exam texts was typically low in frequency and that older literary fiction texts contained similar types of vocabulary. This suggests that students and teachers should rely more on older literary fiction to best prepare for the exam. However, this raises ethical questions about whether an exam should dictate students' reading experience, especially when older literary fiction is likely to be less diverse and dominated by Dead White men.

Keywords: vocabulary, reading, assessment, teaching, English, corpus linguistics

4.1 Introduction

Nationally set external exams have been a feature of the educational system in England for more than a century. These high stakes exams inevitably have an impact on the curriculum that is taught in schools, as the grades achieved by students' affect their education and work choices post-16. This study uses corpus linguistics to analyse the type of vocabulary that is found in a new format of one of the most important of these exams, the English language GCSE. The type of vocabulary that features in these exams is identified and the likely genres of reading that could help students build their knowledge of this vocabulary is also found through comparisons with reference corpora.

4.1.1 Background

Education in England is divided into four key stages: Key Stages 1 and 2 (ages 4-11) are taught in primary schools; Key Stages 3 and 4 (ages 12-16) are taught in secondary schools. At the end of Key Stage 4 there are national examinations in each different subject called the General Certificate of Secondary Education (GCSEs). These qualifications are administered by independent exam boards who are regulated by a government department called the Office of Qualifications and Examinations Regulation (Ofqual). Results in these GCSE exams are then used to gain admittance to post-16 education options, which include: qualifications in traditional academic subjects at a school or college; vocational qualifications at a college; and apprenticeships or traineeships. Most of these options require students to have a minimum of a Grade 4 (previously Grade C) in both English language and maths, as these are taken to indicate a competent level of literacy and numeracy. Any students without the minimum pass grades are required to retake the qualifications as part of their post-16 option. This makes the English language and maths qualifications very high stakes for students. GCSE results, in these two subjects especially, are also very high stakes for schools as pass rates are published by the government and are used to judge school performance and effectiveness.

Externally set exams have been a feature of the education system in England since the middle of the 19th century, with the first national qualification for 16-year-olds introduced in 1918. The content of curricula in England has therefore been influenced by externally set exams for over a century. Schools and teachers in England are therefore used to having a Key Stage 4 curriculum that is focused on high-stake external exams and teaching a curriculum that is heavily influenced by content that is set by the Government DfE, regulated by Ofqual and administered by independent exam boards. Whilst there may be a general belief that testing has the ability to raise standards (P. D. Hart & Teeter, 2001; Mitchell, 1997), high

stakes testing and exams have also been found to have a detrimental effect on teaching practices and curriculum decisions (Brown, 2015; J. L. Jennings & Bearak, 2014; Volante, 2004). The design and content of any test, has the potential to skew classroom practices in favour of drills and practice testing (Sacks, 2000) and to eliminate any curriculum content that is not predicted to be on the test (Volante, 2004). The content of national exams, and their potential impact on teaching practices, is therefore an important area of research.

The current study focused on the reading part of a new specification of the English language GCSE. The format of this exam changed in 2017, as part of wider government reforms intended to raise standards, from being partly assessed through coursework, an oral assessment and exams that could be taken at several different points during the course, to being linear and solely assessed through two written exams at the end of the course. The unseen reading part of the exam is now worth 50% rather than 20% of the total, with the remaining 50% testing students' writing ability. The form and age of the reading texts also changed from the previous specification where "cultural diversity, multimodal study and connections to the real world and daily life" were more of a focus (Isaacs, 2014). Now the Government's Department for Education (DfE) specified that the texts must provide a high challenge and be in a traditional form such as an essay, review or print journalism and explicitly excluded forms of writing found online. They also specified that the exam texts must be literature or literary non-fiction and be drawn from each of the last three centuries (19^{th}, 20^{th} and 21^{st}) (Department for Education, 2013a). An additional difference to the previous qualification format was that there are no longer two exams at different levels, one with more accessible texts for students working within the lower half of the grade range (C-G) and one with more challenging texts for the higher grades (A*-C). In the new specification all students sit the same exams and read the same challenging texts (grades awarded: 9-1).

Preparing students to successfully comprehend previously unseen literary texts that have been drawn from the previous two centuries, has therefore become much more of a focus in Key Stage 4 English classrooms and the potential impact of the new English language GCSE is relatively new. Text comprehension involves many different levels of processing, from decoding and understanding word meanings to working out the structure of the text and constructing a situation model (Kintsch et al., 2005). It also involves comprehension skills such as inference-making and comprehension monitoring (Perfetti et al., 2005). Whilst these are all worthy of study, at the heart of many of these processes is having access to a wide vocabulary which enables a reader to efficiently process texts, thus freeing up resources needed for high level comprehension processes. Vocabulary is therefore the focus of the current study.

4.1.2 Vocabulary Knowledge and Reading Experience

One of the strongest predictors of successful reading comprehension is vocabulary knowledge: at a very simple level, if you don't know what words mean (or have only basic knowledge of their meaning) in a text then you can't understand the text. This is especially the case for older students as reading materials increase in difficulty (Braze et al., 2007; Henderson et al., 2013; Lervåg et al., 2018; Nation & Snowling, 1998; Tilstra et al., 2009). Perfetti and Hart's lexical quality hypothesis (LQH) (Perfetti & Hart, 2002) describes high-quality vocabulary knowledge, needed for successful comprehension, as depending on repeated exposures to words through reading experience. Building on this, Nation's lexical legacy hypothesis (LLH) (Nation, 2017) suggests that exposures to words need to be multiple and diverse, so that readers gradually build lexical quality as they encounter words in different contexts over time. For example, the word 'crest' appears three times in the exam texts collected for the current study. One time it refers to a heraldic emblem on a tin, the other two times it is referring to the top of a wave. These are two quite distinct meanings, as is a

third possible meaning, part of the head of a bird or animal. Knowledge of these distinct meanings and more nuanced understandings within them, for example that 'crest' can also refer to the top of a mountain as well as a wave, would need to be built through repeated diverse experiences with the word.

Previous studies of the relationship between reading experience and reading ability, have shown that it is fiction book reading that improves reading comprehension performance, rather than the reading of non-fiction, magazines, newspapers or digital reading (McGeown et al., 2015; Pfost et al., 2013; Torppa et al., 2020). It is not clear from these studies why fiction book reading was a superior predictor, but it seems likely that it provides more diverse contexts within which to encounter and re-encounter the kind of vocabulary that is found in the standardised and researcher developed reading measures that were used. What this study examines, using corpus linguistics methods, is the nature of the vocabulary in the reading tests themselves (in this case the English language GCSE exam) and then the genres in which that vocabulary is most likely to be found. If, as previous studies suggest, source genres for the vocabulary in the English language GCSE exams are predominantly fiction genres, then this could provide important information for practitioners. Whether or not practitioners should recommend particular genres of reading to students, on the basis that they may provide multiple exposures to the type of vocabulary that will be in the exam, or whether this could instead be regarded as "teaching to the test" is a matter for teachers, policy makers and test developers to discuss.

4.1.3 Corpus Linguistics

Corpus linguistics is the study of lexical and grammatical patterns in a body (corpus) or bodies (corpora) of texts using both quantitative and qualitative methods (Biber et al., 1998). It uses computer software, for example corpus packages like *Sketch Engine*, *AntConc*, *#Lancsbox* and *Wmatrix*, to automatically retrieve and analyse language use (Anthony, 2013).

Corpora can be used for a variety of research purposes, for example: by lexicographers as empirical frequency data for dictionary entries; by applied linguists to study language use in specific contexts and registers; and by language teachers and learners to explore language use with the view to inform pedagogical practice. Corpus software, like *Sketch Engine* (Kilgarriff et al., 2014), give access to reference corpora, large collections of texts created to be representative of certain registers or genres, which can be compared to other purpose-built or smaller specialist corpora. New corpora can be created by uploading texts into the corpus software, which then allows the language in any new corpus to be interrogated using analytical tools such as frequencies, concordances, collocations and keywords. These tools allow for an empirical and more systematic, and consistent analysis of words and their uses in larger data sets and for discovering patterns that might simply escape the attention of an analyst performing a sole qualitative analysis based on 'manual' reading of texts. They also reduce the possibility of human error when counting words and minimise certain biases such as primacy bias that might inadvertently influence qualitative research of texts and vocabulary therein.

4.1.4 This Study

Previous research has highlighted the importance of vocabulary knowledge to reading comprehension and also the importance of reading experience to building this vocabulary knowledge. The key aims of the current study were to create a corpus of a sample of the texts used in the new exams to: (1) identify vocabulary that is typical of the exam texts; and (2) identify in which types of reading this vocabulary is most likely to be found. This could then suggest which genres of reading would provide the best reading experience for the types of vocabulary found in the exams.

The research questions for this study were:

RQ1 – What type of vocabulary is typical of the exam texts?

RQ2 – In what types of reading material is the vocabulary that typifies the exam texts most likely to be found?

4.2 Method

4.2.1 Data

A small, specialised corpus was created in the corpus tool *Sketch Engine* (Kilgarriff et al., 2014), to be referred to here as the 'Exam Text Corpus' (ETC), to represent the reading extracts from the English language GCSE exams. The extracts were sourced from the sample assessment materials and the three sets of past papers that were publicly available when the data was collected (June 2017, November 2017 and June 2018) from the four awarding exam boards in England: AQA, Edexcel, Educas and OCR. In total there were 59 extracts available from the exam board websites. The ETC contains 36,585 words, of which 6,854 are unique. The documents were categorised as fiction or non-fiction and by their century of publication (see Table 1 and Table 2).

Table 1

Document Distribution in the Exam Text Corpus: 19th, 20th and 21st Century and Fiction and Non-fiction Sub-corpora

	19th century	20th century	21st century
Fiction	6	11	2
Non-fiction	14	10	16

Table 2

Words, Unique Words and Documents in the Exam Text Corpus and Sub-corpora

	Words	**Unique words**	**Documents**
Whole corpus	36,585	6,854	59
Sub-corpora			
Fiction	14,946	3,328	19
Non-fiction	21,639	5,154	40
19th century	11,060	2,856	20
20th century	15,008	3,455	21
21st century	10,517	3,264	18

The 59 exam texts were then divided into the David Lee Categories (Lee, 2001), which are genre categories devised by Lee from the contents of the British National Corpus (BNC) (see Table 3). Lee created these categories, after the publication of the BNC, because he argued that the existing classification of texts within the corpus was too broad and that researchers would benefit from being able to identify specific genre categories. Lee took 'genre' to mean a culturally constructed type, as compared to 'register' which described linguistic patterns. Lee carried out the classification himself, which gives consistency but also means that it should be remembered that the categorisations are subjective. The final 70 categories, of which 24 are different genres of spoken language (e.g. broadcast news, conversations, courtroom speech, and meetings), and 46 are different genres of written texts (e.g. biography, prose fiction, letters, and newspapers) were decided upon to represent as far as possible the widest range of all the different types of texts in the BNC, but also with reference to categories used in other corpora to enable comparisons (Lee, 2001). Using the David Lee Categories also allows for distinction, within BNC's wide genre categories of 'imaginative' and 'informative' written texts, of sub-genres that may blur the boundaries of the genres. For example, biographical writing (informative writing) may have more in

common, as a register, with fictional prose (imaginative writing), than it does with other informative writing (e.g. commercial writing).

Table 3
Exam texts by David Lee Genre Categories

David Lee Genre Category	Number of exam texts in category
Scripted speech	1
Biography	17
Essay	1
Prose fiction	19
Personal letters	3
Miscellaneous	3
Newspaper articles	13
Popular magazines	2

4.2.2 Keywords

Keyword lists are generated in *Sketch Engine* by comparing the focus corpus, the ETC, to a reference corpus. The keywords are individual words that appear more frequently in the focus corpus than in the reference corpus. This is calculated by dividing the frequency per million (fpm) of each word in the focus corpus by the fpm of the same word in the reference corpus and by adding the simple maths parameter to account for the zero problem in divisions (Kilgarriff, 2005). The bigger the difference between the two fpm values, the higher the keyness score of the word. Keywords can therefore show what is specific or different about the language in the focus corpus compared to general language, as represented by the reference corpus (Evison, 2015; Kilgarriff et al., 2014). The keywords generated from the ETC were able to provide the answer to RQ1, what type of vocabulary is typical of the exam texts. The BNC was chosen as the reference corpus as it was designed to represent a

cross section of both spoken and written British English and so best represents language that is generally used or experienced by students.

According to Koester (2015), there can be problems of local density in small corpora, like the ETC, where a word appears more frequently in the corpus due to just one document. Keywords were, therefore, selected from the 1,000 generated by *Sketch Engine*, using the following criteria: (1), they appeared at least twice in the corpus; and (2), they appeared in more than one exam; and (3), they appeared as only one part of speech. *Louder* was removed as a keyword as its selection was due to a tagging error in *Sketch Engine*

After applying these criteria, 146 keywords remained from the initial list of 1,000. Fpm was recorded for each keyword from the BNC to give an indication of the frequency in general language (see Appendix A).

4.2.3 Comparisons with Other Corpora

By selecting a range of register-specific corpora it was possible to see which genres of texts were most like the ETC and were therefore most likely to contain the vocabulary in the exam texts (RQ2). The corpora were selected to represent the different types of texts that students may encounter and so came from a range of sources and dates of publication (see Table 5).

The corpus comparison tool in *Sketch Engine* compares two corpora at a time by taking the 5000 most frequent words from each corpus and calculating keyword scores for words that are in both corpora. The mean of the highest 500 keyword scores becomes the overall score for the comparison. The lower the overall score is, then the closer the match between the reference corpus and the focus corpus. Comparing the ETC, which is the focus corpus in this study, to a range of reference corpora demonstrates the kind of registers and genres with which it aligns and from which it differs.

4.2.4 Genre Sources for Keywords

Whilst the whole corpora comparison in 4.2.3 identified likely genre sources for the vocabulary in the ETC by comparing it as a whole to other whole corpora, this was explored further by searching the BNC for the 146 keywords from the ETC and identifying the specific genres sources for them. The BNC was selected as the reference corpus for these searches as it has the most diverse and specifically labelled range of genres. The raw frequencies and fpm for the ETC keywords were found for the general registers of spoken texts and written texts and then the raw frequencies and fpm were found for two subcategories within the written texts: written imaginative and written informative.

Sketch Engine uses relative frequencies to calculate how likely it is for a word to appear within one of the David Lee genre categories, compared to the whole corpus. This is calculated by taking the number of occurrences in a genre category divided by the total occurrences in the whole corpus and then dividing by the size of the genre category within the corpus. A score of 100 would mean there was an equal likelihood of finding the word in the David Lee genre category as in the corpus as a whole. As scores rise above 100 the relative frequency of the word in the genre increases, if the score is below 100 then the relative frequency was lower in the genre than the corpus as a whole. For example, the noun *tea* has a score of 5 in the Hansard category (the record of debates in the UK Parliament), a score of 102 in the arts sections of regional newspapers and a score of 395 in spoken conversations. These relative frequencies show that *tea* is very infrequent in parliamentary debates, appears in the arts sections of regional newspapers with about the same frequency as the whole corpus and is a more frequently occurring word in spoken conversations. Those genres with higher relative frequencies are likely to be the categories in which the vocabulary that typifies the exam texts is most likely to be found (RQ2).

4.3 Findings

4.3.1 Keywords

The 146 keywords, which are the words that typify the ETC (RQ1) are displayed in Table 4 (see Appendix A for a more detailed list). It would be expected that a corpus of written text would contain more low frequency words than spoken language (Korochkina et al., 2024). Words with a fpm of less than five are considered to be low frequency in general language (Brysbaert et al., 2018) and 71% (104/146) of the ETC keywords have a fpm of less than five in the BNC, with 96% (48/50) of the top 50 keywords having a fpm of less than one and thus represent very unusual words. There was a small negative correlation between the two variables of keyness and fpm in the BNC, $r = -.21$, $n = 146$, $p = .01$. This is shown in Figure 1 where the keywords cluster at the lower end of the frequency measure with the highest keyness scores also having low frequencies.

Table 4

Keywords with Scores (Ordered by Keyword Score)

breaker (74.49)	Iceland (18.69)	humiliate (12.58)	vessel (9.77)
nasally (69.44)	wade (18.06)	hasty (12.49)	shore (9.74)
thrill (42.59)	stocky (17.82)	horribly (12.46)	boat (9.73)
boulder (36.3)	trifle (17.74)	envelop (12.43)	tow (9.63)
uncontrollably (35.23)	wistfully (17.7)	scorch (11.96)	consonant (9.61)
clang (32.93)	burnt (17.29)	wardrobe (11.89)	wrestle (9.61)
napkin (32.63)	nightdress (17.25)	ooze (11.73)	rotten (9.59)
majestically (32.41)	motionless (17.18)	tweed (11.68)	amusing (9.59)
cravat (31.47)	crumpled (16.72)	cork (11.63)	Nelson (9.58)
balloon (30.56)	sickening (16.67)	crest (11.59)	ocean (9.5)
slosh (29.58)	horrid (16.57)	fiercely (11.49)	prison (9.43)
dispirit (27.34)	gust (15.8)	prisoner (11.42)	eyelid (9.34)
incessantly (26.41)	quicken (15.77)	wail (11.36)	sofa (9.29)
swimmer (25.84)	hoarse (15.75)	soup (11.27)	rejoin (9.28)
Lucy (25.37)	bedside (15.69)	float (11.27)	pizza (9.28)
handshake (25.14)	Pat (15.66)	arrogant (11.17)	moonlight (9.25)
giddy (24.72)	hearty (15.22)	landing (11.0)	dwindle (9.25)
molten (23.74)	agony (14.72)	fro (10.9)	packed (9.2)
solitary (23.03)	drip (14.49)	sock (10.72)	blanket (9.11)
rut (22.93)	fragrant (14.37)	meaningless (10.72)	tub (9.11)
crockery (21.93)	dangle (14.19)	ghastly (10.69)	dreadful (9.09)
rekindle (21.72)	neglected (14.0)	kitten (10.67)	woollen (9.06)
divest (21.16)	amiable (13.86)	siren (10.61)	stocking (9.05)
idleness (20.99)	rosy (13.86)	frantically (10.61)	dart (9.02)
spoonful (20.91)	upside (13.75)	jersey (10.42)	utmost (8.84)
plank (20.75)	housekeeping (13.72)	hopelessly (10.34)	nostril (8.8)
nightfall (20.67)	mantelpiece (13.72)	shabby (10.3)	expedition (8.8)
homework (20.39)	ledge (13.5)	instant (10.21)	defiance (8.8)
endurance (20.28)	sane (13.45)	muffle (10.21)	midday (8.77)
gobble (19.91)	hue (13.38)	hideous (10.17)	ice (8.66)
shriek (19.68)	thrilling (13.12)	radiate (10.02)	deck (8.64)
ox (19.64)	tea (13.09)	mighty (9.99)	creak (8.59)
speck (19.21)	throb (13.02)	seep (9.98)	foam (8.57)
yank (19.21)	weary (13.02)	fury (9.87)	vigour (8.5)
savoury (19.14)	miraculous (12.99)	hillside (9.87)	distressed (8.48)
spiky (19.14)	online (12.64)	feeble (9.84)	
smelt (18.74)	spoon (12.59)	ache (9.82)	

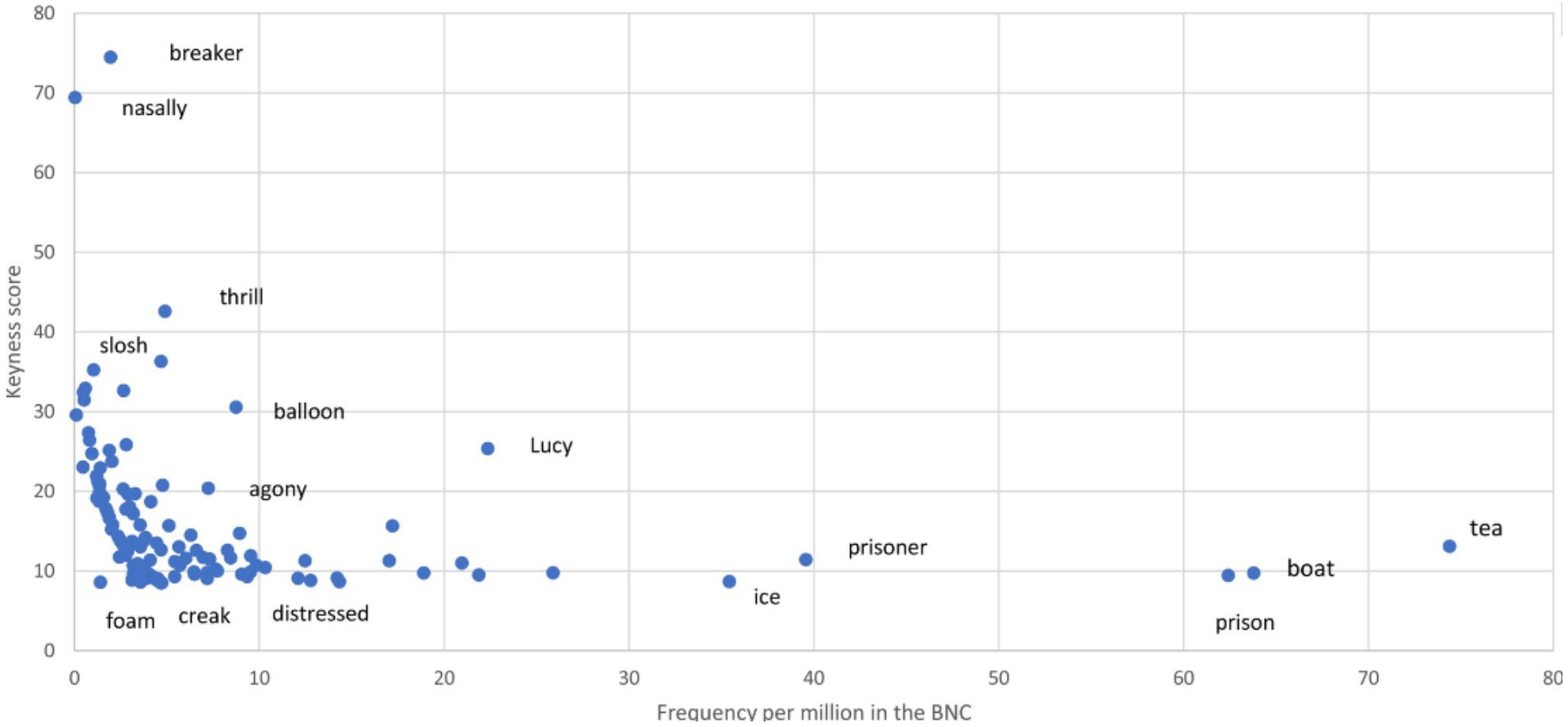

Figure 1 Scatterplot of Keyword Scores and Frequency Per Million in the British National Corpus. Keywords are included for a small sample to illustrate the types of words found in different parts of the figure

A small number of the keywords could be described as archaic, like the adverb ‘fro’, the adjective ‘woollen’, the verb ‘envelop’, and the noun ‘tweed’. Clothes and material feature, with ‘cravat’, ‘nightdress’, ‘tweed’, ‘sock’, ‘jersey’, and ‘stocking’ in the noun list and ‘woollen’ as an adjective. Food appears too with ‘trifle’, ‘tea’, ‘soup’, and ‘pizza’ appearing as nouns as well as items to do with food such as ‘napkin’, ‘crockery’, ‘spoonful’, and ‘spoon’. Words to do with the sea, such as ‘breaker’, ‘crest’, ‘vessel’, ‘shore’, ‘ocean’, and ‘deck’, also feature in the nouns. These patterns suggest that extracts selected for the exams tend to describe people (and their clothes), and social gatherings (food), as well as travel or exciting events (represented by the sea). Verbs, adjectives and adverbs add to this focus on description with the trend seeming to be towards extremes: the verbs indicating dramatic or negative events with ‘shriek’, ‘yank’, ‘wail’, and ‘wrestle’; the adjectives either being pejorative, for example ‘sickening’, ‘horrid’, ‘ghastly’, ‘hideous’, and ‘dreadful’, or the more positive ‘fragrant’, ‘rosy’, ‘thrilling’, and ‘miraculous’; and the adverbs cover a range from ‘frantically’ and ‘fiercely’ to ‘wistfully’ and ‘majestically’. Overall, these keywords seem to be centred on people or characters, be highly descriptive and tending towards either domestic affairs like meals or extreme or dramatic events.

4.3.2 Comparison with Reference Corpora

The corpus comparison tool in *Sketch Engine* takes the mean of the highest 500 keyword scores between two corpora as an overall score for similarity between them. The ETC was compared to a range of reference corpora (see Table 5). The corpus that had the lowest overall score when compared to the ETC, and therefore was the most like it, was Project Gutenberg English (2.16) (a free digital library of mostly out of copyright literary texts); the corpus with the highest overall score and therefore the most different from the ETC, was British National Corpus Spoken (3.69). The more modern corpora, English Web

2015 and English Broadsheets 1993-2013, also had high scores and so, along with the spoken corpus, were furthest from the ETC.

Table 5

Comparisons of Reference Corpora to Exam Text Corpus: the Lower the Score the More Alike the Corpus is to the Exam Text Corpus

Corpus compared to ETC	Score
Project Gutenberg English	2.16
Brown Family (written American and British English)	2.29
British National Corpus	2.41
English Broadsheets 1993-2013	2.65
English Web 2015	2.91
British National Corpus 2014, Spoken	3.69

4.3.3 Genre Sources for Keywords

4.3.3.1 BNC Genres. To explore the most likely source genres for the vocabulary from the exam, a search was run for each of the 146 keywords from the ETC in the BNC to identify the frequency counts in different genres of texts. Where the ETC keywords are most frequently found gives an indication of where the students might be most likely to encounter the words in their reading. Raw frequencies and normalised frequencies (fpm) are shown in Table 6 below. Fpm allows for comparison between different sizes of corpora. Frequencies are given for the two general registers of spoken and written texts and then for two sub-registers within the written text register: written imaginative and written informative.

Overall, the fpm scores, which were used for the comparison rather than the raw frequencies due to the different sizes of the sub-corpora, were higher for written than spoken texts. There was also a much higher fpm score for written imaginative texts compared to written informative texts. This indicates that, as with the corpora comparison result in 4.3.2

where the ETC was most like a written corpus of older literary texts, the ETC keywords are most likely to be found in written imaginative texts.

Table 6
Frequencies and Frequencies Per Million of the Exam Text Corpus Keywords in the British National Corpus

Registers in BNC	Raw frequency of keywords	Fpm of keywords
Whole corpus	109,841	977
Spoken texts	8,133	689
Written texts	101,771	1,012
Written imaginative	36,611	1,851
Written informative	65,159	806

4.3.3.2 Relative Frequencies in the David Lee Categories. For a more detailed breakdown of types of texts that contained the keywords, a relative frequency was calculated for the David Lee Categories (see Appendix B). Scores for the spoken genre categories in the classification were generally well below 100, meaning that the keywords were much less likely to be found in the spoken genres than in the BNC as a whole. There were higher scores in the written genre categories, with by far the highest averages in poetry (371.66) and prose fiction (227.17). Considering that only 19 (32%) of the exam texts were prose fiction, and none were poetry these are surprisingly high relative frequency scores. There were also some high relative frequencies amongst the non-fiction categories, for example biography (134.22), the arts sections of broadsheet newspapers (145.21) and tabloid newspapers (125.12), mirroring the largest non-fiction David Lee Categories of the exam texts (biography and newspaper articles) and suggesting that the boundary between the general categories of imaginative and informational texts is not always distinct linguistically.

4.4 Discussion

The background to this paper is the introduction in 2017 of a new specification of a high-stakes national exam in England, the English language GCSE. The reading part of the exam is now worth a higher percentage of the marks and also now has to include literary texts from the 19th, 20th and 21st century, thereby introducing an increased focus on the types of vocabulary found in these texts.

The majority of the keywords from the ETC were low in frequency presenting a challenge to comprehension as it is more difficult for readers to experience them multiple times in diverse contexts to build high quality representations. Whilst the aim of the new English language GCSE was to create an improved qualification that was "more engaging and worthwhile to teach and study, as well as more resilient and respected" and "to prepare young people better for the next steps in their education or employment in years to come" (Ofqual, 2013, p. 3), it is hard to see how the obscure nature of the vocabulary that typifies the exam texts is helping to fulfil these objectives.

Pressure to prepare students for these exams could lead teachers to feel they should include more older texts in the curriculum, to increase the exposure to archaic words. Or design lessons that focus on the rote learning of low-frequency vocabulary, when there is little evidence that teaching word meanings directly improves comprehension (Wright & Cervetti, 2017). If past exam papers are used for preparation lessons, with their focus noted above on characters, social and dramatic events, then students' curriculum reading experience could be narrowed. Such exam focused activities, as Volante (2004) points out, are not always in the best interests of the students or necessarily effective activities for learning.

The genre in the BNC that contained the exam keywords most frequently was imaginative texts. This is in line with the literature that identifies fiction as the type of reading experience that best predicts reading ability (McGeown et al., 2015; Pfost et al.,

2013; Torppa et al., 2020), supporting the hypothesis that fiction reading is superior because it provides the best source for vocabulary encounters due to its diversity. Fiction takes us to places we've never been, to times we could never travel back or forwards to and puts us into action that we might never normally experience. This diversity of place, time and action is described through a diversity of vocabulary that we might not otherwise encounter. But the closeness of the exam vocabulary to older literary fiction, through the match with Project Gutenberg, calls into question whether this is the type of fiction that adolescents are, or even should be, predominantly reading.

Whilst fiction texts are obviously important to read, they are not the only type of genre that students will need in their future education and employment. Some non-fiction source genres featured in the findings, but these were limited to biography and newspaper articles. Familiarity with, for example, instructional texts, academic texts and new media could also be considered essential or at least useful to students' future literacy and employment. However, the reading of this range of genres is not the best preparation, according to the findings in this study, for the vocabulary in the current exam texts. This places schools and teachers in a difficult position when selecting classroom texts or activities. Gaining a good grade in this exam is essential for young people to access their next steps in education and employment. However, preparing students to obtain this grade is not necessarily going to prepare them for the literacy demands of their future. Given that curriculum time is not unlimited, teachers may have to choose between preparing students for the exam, by choosing reading that will exposure them to the types of vocabulary that is likely to be in the exam (using older, literary fiction), or choosing curriculum materials that they feel will prepare students for their further studies (e.g. academic texts), employment (e.g. instructional and commercial texts), and successful societal relations (e.g. new media and

online texts). These external pressures on curriculum time could leave little, if any space, for reading which teachers might choose that is inspirational or enjoyable or thought provoking.

The English curriculum, both before, during and after the GCSE qualification, already receives criticism for its lack of diversity and representation (V. Elliott et al., 2021). Much good work has been done in schools to promote reading for pleasure that includes diverse voices, contemporary concerns and spaces where students see themselves represented (Clark & Rumbold, 2006; Department for Education, 2012). If preparation for the English language exams is better served by reading traditional literary fiction, a canon of texts that is already covered by a separate English literature GCSE, then wider representation could be jeopardised and the dominance of the writings of dead white men could become further entrenched, as schools may feel that they should be recommending students read older literary texts instead of encouraging freedom of choice.

4.5 Further research and limitations

While the main finding of the current study is clear, there are some limitations which should be acknowledged. First, the focus was of this study was solely on vocabulary as an important component of comprehension. Further research on other aspects of comprehension such as collocations, syntactic and morphological structures, and the need for high-level processing such as inferences and comprehension monitoring with regard to the exam texts would be advantageous but were beyond the scope of this study.

A second limitation was the small size of the ETC. It would be useful to continue to grow the ETC, as more exam texts become publicly available, to monitor whether or not later extracts change the typical vocabulary found, and to broaden the kinds of analysis that are possible.

This study makes some assumptions about students' reading experiences. Further research into the actual reading habits of students, for example through a reading survey,

would be valuable. The extent to which the exam text vocabulary is already found in curriculum materials and the choices that students are currently making about what to read for pleasure would be a valuable addition to the current literature.

4.6 Conclusion

Due to the long history of externally set exams, teachers in England have been used to teaching a curriculum, at Key Stage 4 especially, that is heavily influenced by the content in the exams. This study has found that the vocabulary in the new English language GCSE to be typically low frequency and that found predominantly in older, literary fiction. This calls into question whether the qualification is achieving its stated aim of preparing students for future study and work. It could also potentially skew what teachers feel they ought to choose for students to read in lessons and recommend that they read at home, as students' access to post-16 education and work opportunities depends on success in this exam. However, the promotion of this type of reading could limit students' reading experience to the literary canon of mostly dead white men, undermining efforts to increase diversity and representation in the curriculum and encourage freedom of choice in reading for pleasure. As the impact of this new qualification begins to become clear, teachers may feel that this is a step too far and move to question the influence of the test on the curriculum.

Chapter 5. Paper 2

Are some types of reading more equal than others? Adolescent reading experience and the requirements of high-stakes assessments

Authors: Beverley Jennings, Daisy Powell, Holly Joseph

The paper reproduced here was submitted to the *Journal of Adolescent and Adult Literacy* in October 2024

Abstract

Reading experience is a key predictor of reading comprehension skill. However, we do not know whether reading experience predicts performance in examinations and whether the type of reading experience matters. This study sought to investigate whether reading experience and vocabulary knowledge had a relationship with attainment in a high-stakes examination taken in England at the end of secondary school (age 16). This exam requires students to comprehend literary texts from all three of the 19th, 20th and 21st centuries. 330 16–18-year-old students completed a survey which asked about their attainment, their reading habits, had an author recognition test (ART) (which measured classic and young adult (YA) fiction authors separately) and a vocabulary test. Responses to the survey came predominantly from high attaining, female students, from high socioeconomic backgrounds. They did show that, for these respondents, there was no significant difference between the vocabulary scores of high-grade and mid-grade students. However, the high-grade group obtained significantly higher scores compared to the mid-grade group on the ART for classic but not YA authors. It is therefore likely that, for these respondents, there was a relationship between their reading experience and attainment. This could have implications for current trends, certainly in the UK, of encouraging reading by promoting student choice and the diversity and representation of reading texts. This would be more difficult if, with an aim of improving attainment, reading experience with classic authors was promoted instead.

Keywords: reading comprehension, adolescents, reading for pleasure, reading experience, author recognition test (ART)

Implications for Practice

What is already known about this topic:

- Prior reading experience is an important predictor of reading comprehension skill.
- The reading of fiction has been found to have a closer relationship to reading skill than other genres of reading.
- Assessment of English language, at age 16 in England, has changed to now include older literary reading texts in the exams.

What this paper adds:

- Students choose to read modern fiction for their independent reading.
- Students who obtain high grades in their English Language GCSE exam have better knowledge of classic authors than students who gain mid grades.

Implications for theory, policy or practice:

- Measures of print exposure need to be nuanced enough to capture the type of reading experience, as well as how much.
- The implications for student choices in their independent reading should be considered when selecting texts for the curriculum and assessments.
- The impact of the new criteria for exam texts could lead to the promotion of classic texts over modern or YA fiction, which could compromise representation, diversity and free choice in reading for pleasure.

5.1 Introduction

Prior reading experience is one of the strongest predictors of reading comprehension proficiency (Acheson et al., 2008; Chateau & Jared, 2000; Mol & Bus, 2011). Nation's lexical legacy hypothesis (LLH) (Nation, 2017) explains this predictive relationship by viewing reading experience as a means for providing a gradual accumulation of knowledge about a word through multiple, diverse encounters with it. With each encounter a reader infers something about its meaning through the words that surround it (i.e. its context). This means that, over time, readers gradually build more precise and nuanced lexical representations of the words they read (Perfetti & Hart, 2002). Reading experience is especially important for the acquisition of high-quality representations of low frequency vocabulary as it is through reading, rather than through spoken language, that low frequency vocabulary is more often encountered and re-encountered (Cunningham, 2005; Landauer & Dumais, 1997).

However, not all types of reading experience are associated with reading comprehension skill to the same degree (Allen et al., 1992; Anderson et al., 1988; Spear-Swerling et al., 2010; Walberg & Tsai, 1984). It is fiction book reading that appears to be superior, for predicting pathways to reading comprehension (Torppa et al., 2020), or pathways from reading skill (McGeown et al., 2016), for predicting reading ability measures (word reading, comprehension, summarisation and reading speed) (McGeown et al., 2015), and for positive relationships to measures of information retrieval, inferential understanding and vocabulary (Pfost et al., 2013). Conversely the reading of digital texts (e.g. emails and blogs) has a negative path to reading comprehension (Torppa et al., 2020), an inverse relationship with reading skills (McGeown et al., 2016) and lower scores for information retrieval, inferential understanding and vocabulary (Pfost et al., 2013). Non-fiction reading, like newspapers and magazines, did not predict reading comprehension (Torppa et al., 2020)

and also had weaker correlations with reading ability measures and no correlation with reading speed (McGeown et al., 2015). Reading speed is an important indicator of reading competency because, if reading is too effortful (slow), fewer cognitive resources are available for comprehension (Martin-Chang et al., 2020).

The evidence is strong, therefore, of a positive relationship between book or fiction reading and reading proficiency. As shown above there are predictive or correlational relationships between measures such as word reading, reading speed, sentence completion, summarisation, and passage comprehension with experience of reading fiction. There is less evidence however, on the causal relationship between reading ability and reading experience. Studies of young readers have found that the direction of effects are from reading skills to print exposure (Harlaar et al., 2011; Leppänen et al., 2005; Torppa et al., 2020; van Bergen et al., 2018). Studies that were with or included older children found that the relationship was either reciprocal (Torppa et al., 2020) or that reading from mid childhood onwards affects later reading skills (van Bergen et al., 2021).

It would therefore be expected that with adolescents, the subject of this study, there would be a relationship between reading experience and reading comprehension skills and that the reading of fiction would be an especially important variable to explore as it is likely to be a predictor of comprehension skill and attainment. Adolescents are an understudied group as they can be a difficult to include in research as, at this age, the focus in schools is on final school exams or tests for graduation. One study that did include this age group, Spencer (2017), looked at the relationship between measures of language ability at age 13-14 years old, and subsequent attainment in English language, English literature, and mathematics General Certificates of Secondary Education (GCSE) exams. GCSE exams are national exams that are sat by students in England at the end of their secondary education (age 16). The data from Spencer's study showed that all the language measures, which included spoken

language and vocabulary, were associated with a higher grade in English language, English literature, and maths. As language measures have been shown before to be associated with reading experience, it might be expected that reading experience would therefore have a relationship to exam attainment.

As shown above, fiction reading is associated with higher reading comprehension proficiency (McGeown et al., 2015; Pfost et al., 2013; Torppa et al., 2020) and language ability has been associated with better attainment in public exams (Spencer et al., 2017). There has also long been general recognition, in reports from charities and research organisations, that there is a link between leisure reading, sometimes called reading for pleasure, and general attainment. The National Literacy Trust's review into reading for pleasure (Clark & Rumbold, 2006) cited benefits for writing, comprehension, and breadth of vocabulary. The OECD's 'Reading for Change' report (Kirsch et al., 2002) positioned engagement in reading in free time as important for reading proficiency. Sullivan & Brown (2015) linked reading for pleasure to vocabulary and progress in maths. This has led to Government reports, as well as independent reviews commissioned by Government, recommending, encouraging, and even requiring schools to promote reading for pleasure (Department for Education, 2012, 2023b; Rose, 2006). Charities and non-profit organisations also promote reading for pleasure, for example through World Book Day (https://www.worldbookday.com) and projects like Reading for Pleasure by the Open University (https://ourfp.org).

Several studies have shown that allowing students to choose what they read promoted their reading for pleasure (Casey, 2010; Gambrell et al., 1996; Moss & McDonald, 2004) and the 'What and How Kids are Reading' report (Topping et al., 2023), showed that the books read most often by students in years 9-11 (age 13-16), outside of texts that are on the English literature curriculum, were all by children's or young adult (YA) fiction authors (e.g. Suzanne

Collins, J K Rowling, and David Walliams). Reading for pleasure as a concept can, however, be problematic (Cremin et al., 2022). It cannot be guaranteed that giving readers agency and free choice will automatically lead to intrinsic motivation to read and it can also be true that extrinsically motivated reading tasks, such as a regular school-mandated reading activity, can be found to be pleasurable by some. There is also a danger that allowing young people to choose what they read could lead them to choose books which are not challenging for them, thereby reducing opportunities for vocabulary growth. A recent report (Picton & Clark, 2022) has also identified the difficulty that certain groups of children and young people have in finding themselves represented in books, with the suggestion that increased diversity and representation in books could potentially support increased reading engagement.

The National Literacy Trust's Annual Literacy Survey, which asks children and young people about their literacy attitudes and habits, reported that only 43% of 8–18-year-olds in 2023 said that they enjoyed reading and only 51% of 8–18-year-olds said that they read fiction on paper (Clark et al., 2023). This is of concern as reading on paper has been shown to be better for reading comprehension than reading on screens (Kong et al., 2018; Salmeron et al., 2024). The headline figure on reading enjoyment (43%) marks a decline from previous years' results, which the National Literacy Trust described as a crisis in children's reading in the UK (Clark et al., 2023).

In summary, we know that reading experience is an important predictor of reading comprehension skill, at least in part, because it builds vocabulary knowledge. There is also evidence that book (fiction) reading in particular is beneficial to comprehension. However, there are few studies that focus on adolescents and this age group can be challenging to include in studies, due to their involvement in high-stakes assessments. It is important to include this age group in research, as reading experience has been linked with improved

attainment, both generally (Clark & Rumbold, 2006; Kirsch et al., 2002; Sullivan & Brown, 2015) and with specific qualifications (Spencer et al., 2017).

This study, therefore, focuses on a new specification of the English language GCSE which was reformed in 2015 with the first new exam taken in 2017. In the previous specification of the exam the reading texts included diverse topics, modern modes, and real and everyday content (Isaacs, 2014). In the new specification, the three reading texts in the exam must be literature or literary non-fiction and from each of the 19th, 20th and 21st centuries. 50% of the marks are now awarded for the reading part of the exam, whereas in the previous version exam reading was 20% of the marks, and all students now sit the same paper, rather than there being a more accessible foundation paper for the lower half of the grade range. The new specification therefore places more emphasis on comprehension of previously unseen older literary texts.

There is an opportunity and a need therefore to study attainment in this new exam and the relationship that it has with student reading habits. It would be predicted, from the research outlined above, that attainment in this exam would be positively linked to reading experience with fiction, something that has been encouraged by government and schools. However, the new specification of the English language GCSE exam includes older literature and literary non-fiction and the type of fiction that adolescents predominantly choose to read is written by contemporary children's and YA authors. It is important therefore, to consider whether there is a difference in students' reading experience – both with fiction that matches the type used in the exam (classic) and fiction which students typically choose (YA) - between those who do, and do not, attain highly.

5.1.1 Research Aims of this Study

Research questions

1. Do students who gain high grades (7-9) in the English language GCSE have higher levels of reading experience, compared to those with mid grades (4-6), with a) classic authors and b) YA authors?
2. Do students who gain high grades (7-9) score more highly on a vocabulary test created from keywords from the exam texts compared to those with mid grades (4-6)?

It is predicted that students with higher English language GCSE grades will have higher levels of reading experience with YA authors, as the literature is strong on the relationship between the reading of fiction and comprehension. It is also possible that students with higher grades will have higher levels of reading experience with classic authors, due to inclusion of texts from the 19th and 20th centuries in the examination. In addition, it is predicted that students in the high-grade group will score more highly on the vocabulary test of keywords taken from the exam texts than students with mid grades, as we expect vocabulary to be a mediator of the relationship between reading experience and English language GCSE attainment.

5.2 Method

5.2.1 Ethics

This study was granted ethical approval by the University of Reading's Institute of Education. Respondents consented to participate after reading the information at the beginning of the online survey.

5.2.2 Participants

The online survey was circulated to students (aged 16-19) at:

1. a sixth-form college in the South-East of England
2. a further education (FE) college in the South-East of England

In England FE colleges typically attract a more diverse and representative student body than sixth form colleges, as the latter tend to only offer more academic qualifications (Advanced Levels) while the former offer a broader range of courses including vocational qualifications. Requests to circulate the survey were sent to many FE colleges in England. Initially the focus was on colleges in the South-East, to match the location of the sixth-form college but none agreed, many citing too high a workload due to COVID. Attempts were widened to other areas of England, including using the personal contacts of the authors, but still no FE colleges agreed to circulate the survey. Eventually a small college in the South-East did agree to send out the link but this only generated a very small number of responses. The vast majority of respondents were therefore from the more academic sixth form college and therefore, as will be outlined below, this was not a representative sample.

The survey was started by 692 respondents. There was a gradual attrition as the survey progressed with 330 completing the entire survey (see Tables 7 and 8). The data were analysed using only the 330 fully completed surveys, of which 316 gave their English language GCSE grade for 2019 or 2020.

It is evident, from Tables 7 and 8, that respondents were not representative of students nationally: respondents were disproportionately female, from a high socio-economic group (indexed by both eligibility for free school meals and postcode), White (although the percentage, 78%, is roughly in line with the national figure of 73%) and extremely high attaining in English language GCSE (73% obtained grades 7-9 compared to the national average 24% grades 7-9). In addition, slightly more students had English as an additional language than is seen nationally. Whilst a more representative sample would have been valuable, it is perhaps unavoidable that an online survey about reading will be completed by groups for whom reading is viewed as a desirable and valued activity and who already excel at and enjoy reading. Interpretation of the results of the survey therefore take into account the

respondents' narrow demographic group and the results can only be generalised to this group of privileged, successful and competent readers.

5.2.3 Impact of COVID-19

The aim of this study was to examine the relationship between the reading experiences of students and the grade they attained in their English language GCSE. Two cohorts of students completed the survey: the first took their GCSE exams in 2019 under normal conditions; the second were awarded their GCSE grades in 2020, which, due to the COVID-19 pandemic, were awarded by the submission of a centre (school) assessed grade (CAG) (Department for Education, 2020). This meant that whilst 50% of the marks awarded for the 2019 grades were for the reading comprehension of unseen texts in exams, there was no uniform method for the CAGs given in 2020, so it is not possible to tell how much, if any, of the grade reflects reading comprehension ability. In addition, after a public outcry, no standardisation or norm referencing was applied to the 2020 CAGs so there was a high degree of grade inflation (Ofqual, 2020). Descriptive data for the two-year groups' measures are therefore reported separately, and cohort year was included as a factor in the analyses.

Table 7

Demographic Information for Respondents Who Started and Completed the Survey Alongside National Statistics

	Completed surveys		All surveys		National for all pupils in England (2019-20)	
Demographic	%	n (330)	%	n (692)	%	n (8,890,357)
Gender						
Male	23	77	27	142	51	4,534,082
Female	72	238	70	365	49	4,356,275
Non-binary	2	5	1	7		
no data	3	10	1	7		
FSM/PP[a]						
Yes	9	30	9	46	17.3	1,538,032
No	85	280	86	438	82.7	7,352,325
No data	5	20	5	26		
IDACI decile[b]						
1-5 (most deprived)	8	26	5.7	40	10	889,036
6-10 (least deprived)	52	170	34	237	10	889,036
no data	40.60	134	60	415		
EAL						
Yes	26	84	24	123	18	1,584,600
No	71	234	75	376	82	7,305,400
no data	4	12	1	5		
Ethnicity						
White	78	257	78	375	73	6,489,961
Mixed	7	22	9	41	6	533,421
Asian	10	32	10	47	11	977,939
Black	2	5	2	7	6	533,421
Other	1	3	1	5	2	177,807
no data	3	11	1	7	1	88,904

Note. FSM = free school meals; PP = pupil premium recipient; IDACI = Income Deprivation Affecting Children Index; EAL = English as an Additional Language.

[a]FSM/PP are used as indicators of disadvantage or low socioeconomic status. Pupil premium is an additional amount of funding given to schools for students who are: eligible for free school meals or have been recorded as eligible in the past 6 years; and/or are currently looked after by the local authorities or have been previously looked after by the local authority or other state care. Participants were asked as a combined question if they received free school meals or pupil premium.

[b]IDACI decile is a government measure of deprivation based on small local areas. Areas are ranked and then divided into deciles with 10 being the least deprived and 1 being the most deprived. It is possible to identify participant deciles from their home postal code, where they were willing to enter it.

Table 8

Raw Numbers and Percentages of Students Who Obtain Each Possible Grade in English Language GCSE, for Those Who Completed the Survey, Those Who Started the Survey, and Those Who Took the Exam Nationally in 2019 (top) and 2020 (bottom)

English Language GCSE grade	Completed surveys		All surveys		National	
	%	n	%	n	%	n
2019	n = 120		n = 165		n = 546,607	
1-3	0	0	0	0	29	158,516
4	3	4	4	7	17	92,923
5	10	12	12	19	20	109,321
6	25	30	28	46	16	87,457
7	24	29	24	40	9	49,195
8	20	25	18	30	6	32,796
9	17	20	13	22	3	16,398
2020	n = 196		n = 293		n = 564,701	
1-3	0	0	0.3	1	20	112,940
4	3	6	3	8	18	101,646
5	7	13	8	24	20	112,940
6	17	34	22	64	18	101,646
7	30	58	24	69	13	73,411
8	19	38	20	59	7	39,529
9	24	47	22	65	4	22,588

5.2.4 The Survey

5.2.4.1 Demographic Questions and Self-reported Reading Habits. The survey (Appendix C) was created using an online survey tool (REDCap) and distributed to students via their educational institution email addresses, to be completed in their own time. The survey started with five questions asking for demographic information, then there was a question about attainment in the English language GCSE, and then respondents were asked about their reading habits. There were self-report indications of the types of genres that they read. They were asked to indicate which genres they read at all, which they read most often, which they enjoyed, which they read at school/college and which they read for pleasure.

Respondents were then asked to estimate the time they spent reading each day and then to compare their own reading and enjoyment of reading to others on a five-point Likert scale. The survey then moved onto an author recognition test (ART) and finished with a vocabulary test.

5.2.4.2 Author Recognition Test. Accurately measuring reading experience is complex and ARTs are often used as an objective measure of print exposure (Beech, 2002; Cunningham & Stanovich, 1990; Sénéchal et al., 1996; Spear-Swerling et al., 2010; Stainthorp, 1997; Stanovich & Cunningham, 1992; Stanovich & West, 1989). Research has shown that such print exposure measures are associated with reading proficiency across the school years and into adulthood (Beech, 2002; Cunningham & Stanovich, 1990; Sénéchal et al., 1996; Spear-Swerling et al., 2010; Stanovich & Cunningham, 1992; Stanovich & West, 1989). ARTs are seen as more accurate than self-reported reading frequency measures because of the perceived social desirability of reading, which can lead respondents to inflate or deflate the reading habits they report (West et al., 1993). Self-reported measures alone can also be an inaccurate reflection of actual reading actitivies due to the high level of participant cooperation required and the difficulty for respondents in making retrospective estimations of reading time (Bisson et al., 2012; West et al., 1993). An ART presents participants with a list of real author names and carefully constructed foils. A score is calculated of all correctly identified authors minus any foils selected. The design is based on the assumption that those respondents who read more will correctly identify more authors' names and thus obtain a higher score.

The ART created for this study was divided into two subsets in order to refine the measure of print exposure, as was suggested by Pfost et al's work (2013). The two subsets were young adult (YA) fiction authors and classic fiction authors. A long list of YA author names was sourced from: best young adult books on goodreads.com; UK public libraries

most borrowed list June 2018-2019; Scholastic Book Club author list; bestselling teenage/young adult books on Amazon; and bestselling teenage book authors in Waterstones. These lists were then merged, and 25 authors selected for the YA list. As whom is considered a 'classic' author can be contentious, the 'classic authors' were drawn from a list of authors included in the English literature GCSE specifications. The literature qualification specifies that students are taught: 1) a play by Shakespeare; 2) a nineteenth century novel; 3) a modern (post-1914) fiction or drama text; and 4) an anthology of poetry since 1789 (including Romantic poetry). Students would therefore not be familiar with a large number of authors on the list just from lessons but would be expected to have read texts by a small number in this concurrently taught GCSE. The foils were created to match the form of the real authors' names (e.g. use of middle names, double barrelled names, initials rather than first names). None of the foils were real authors. The whole list of 100 names, which consisted of 25 YA authors, 25 classic authors and 50 foils, were presented in the survey in blocks of ten in alphabetical order by first name. Respondents were asked to tick any names that they knew were real authors. ART scores were calculated by taking the total number of correctly identified authors minus any foils selected. Two scores were used for the analysis: 1) a classic author score; and 2) a YA author score.

Four Cronbach alpha coefficients were calculated: 1) for the authors correctly identified for the whole ART, it was .95; 2) for authors correctly identified for the classic ART, it was .85; 3) for authors correctly identified for the YA ART, it was .80; and 4) for foils it was .70. The relationship between the results from the classic author subset of the ART and the YA ART, investigated using Pearson correlation coefficient, was a strong positive correlation, $r = .76$, $n = 330$, $p < .001$, with high scores on the classic author subset associated with high scores on the YA author subset. A list of scores for the authors on the ART is in Appendix D.

5.2.4.3 Vocabulary Test. The vocabulary test was a multiple-choice test that required participants to select one of four choices of synonym or short phrase which most closely matched the meaning of the target word. As part of a long survey, it was only feasible for this test, due to the time demands on respondents, to have a small number of items. Therefore, the test was created from 15 keywords (see Appendix E), extracted from a corpus of exam texts created in a separate study (B. Jennings et al., 2024). Keywords are identified, using corpus linguistics techniques (Evison, 2015; Kilgarriff et al., 2014), as being typical of the corpus. The keywords were presented in the same sentences in which they appeared in the original exam texts. In order to validate that the researcher-selected synonym could be considered the correct answer, the test was circulated on social media. 205 adults completed the test and these external adult results, which validated the researcher selected answers, are presented in the final column of Appendix E. Each test was scored (one mark per correct response) out of a total of 15. The Cronbach alpha coefficient for the vocabulary test was .55, meaning that the results of this test should be treated with caution as it was not internally consistent.

5.2.4.4 Data Analysis. To check that the two new ARTs, created for this study, were capturing reading experience, the relationships between the self-reported reading measures, the ART scores and the vocabulary test scores were investigated using Spearman rank order correlations (Table 9). Further analysis with the GCSE grade groups was then only conducted using the ART scores as the literature shows that they correlate more strongly with reading proficiency measures than self-reported measures (Cunningham & Stanovich, 1997; Sénéchal et al., 1996; Spear-Swerling et al., 2010; Stanovich & Cunningham, 1992). Correlations between the ART scores and the self-reported reading measures indicate that the two ARTs, created for this project, were capturing reading experience. Also consistent with previous research, analysis showed that there was a moderately strong correlation between print exposure, both to classic and young adult fiction, and vocabulary scores.

Table 9

Spearman Rank Order Correlations Between Author Recognition Test Scores, Self-Reported Reading Measures and Vocabulary Scores

Measure	1	2	3	4	5	6
1. YA ART	-					
2. Classic ART	.76**	-				
3. Comparative time reading	.35**	.29**	-			
4. Comparative enjoyment reading	.41**	.33**	.69**	-		
5. Average minutes spent reading	.27**	.23**	.41**	.45**	-	
6. Vocabulary score	.48**	.45**	.12*	.18**	.19**	-

** p < .01 (2-tailed). * p < .05 (2-tailed).

The GCSE Grade 1-9 variable was ordinal rather than continuous and because no respondents obtained a grade lower than 4, it was treated as categorical, with respondents divided into two groups: high-grade (grades 7-9) and mid-grade (grades 4-6). These two grade groups were also divided by cohort year because of the different assessment methods used in 2019 and 2020 due to COVID-19. Three 2 (Attainment group: high; mid) x 2 (Cohort: 2019; 2020) between groups Analyses of Variance were conducted to explore the impact of GCSE grade group and the year the GCSE was taken (2019 or 2020) on: 1) the vocabulary test scores; 2) the classic author subset of the ART; and 3) the YA authors subset of the ART.

5.3 Results

5.3.1 Descriptive Statistics

Descriptive statistics for the self-reported genre reading questions are reported in Table 10. Overall modern fiction was read most often (43%), most enjoyed (62%) and read most for pleasure (61%). It was only in the 'most read at school' that modern fiction was not the single highest choice, with non-fiction books (60%), online articles (13%), newspaper articles (8%), and classic fiction (6%) getting more responses.

Table 11 shows the mean ART and vocabulary test scores. Mean scores were higher on the classic ART than the YA ART. Overall respondents performed very highly on the

keyword vocabulary test with a mean score of 12.13 (*SD* = 2.02) out of a maximum of 15 (see Appendix E for full list of items and scores). This suggests that there could have been a ceiling effect in the data (75% of respondents scored more than 11). The high attaining groups scored, on average, a little less than one point higher in the vocabulary test than the mid-grade groups in both cohorts. Scores were about half a point higher in 2019 compared to 2020 for both high and mid-grade groups.

Table 10

Results from Self-Report Questions about Genres that Respondents Read

Genre	Read at all (tick all that apply)	Read most often (tick only one)	Most enjoyed (tick only one)	Most at school (tick only one)	Most for pleasure (tick only one)
Modern fiction	245 (74%)	141 (43%)	203 (62%)	13 (4%)	200 (61%)
Classic fiction	72 (21%)	9 (3%)	12 (4%)	20 (6%)	10 (3%)
Poetry	45 (14%)	2 (1%)	10 (3%)	5 (2%)	3 (1%)
Drama/scripts	43 (13%)	3 (1%)	2 (1%)	7 (2%)	3 (1%)
(Auto)biographies	62 (19%)	3 (1%)	7 (2%)	5 (2%)	8 (2%)
Non-fiction books	161 (49%)	67 (20%)	38 (12%)	197 (60%)	30 (9%)
Newspaper articles	133 (40%)	22 (7%)	10 (3%)	25 (8%)	12 (4%)
Online articles	213 (65%)	61 (19%)	26 (8%)	42 (13%)	34 (10%)
Magazines	55 (17%)	3 (1%)	3 (1%)	0 (0%)	7 (2%)
Other	35 (11%)	14 (4%)	12 (4%)	8 (2%)	16 (5%)

Note. Numbers for each genre are given as raw numbers and then as a percentage of responses

Table 11

Mean Scores, Standard Deviations (in brackets), and Number of Participants for the Author Recognition Tests and Vocabulary Test by Grade Group and Year

Measure	Year	High-grade	Mid-grade
Classic ART Score	2019[a]	11.86 (4.97)	9.64 (4.41)
	2020[b]	11.45 (4.21)	7.50 (4.14)
YA ART Score	2019[a]	7.76 (4.17)	6.47 (4.13)
	2020[b]	7.97 (3.77)	4.71 (2.80)
Vocabulary Score	2019[c]	12.63 (1.89)	11.84 (2.09)
	2020[d]	12.35 (1.77)	11.04 (2.41)

Note. Only 302 students fully completed the vocabulary test and gave grades for 2019 or 2020

[a]*N* = 120. [b]*N* = 196. [c]*N* = 116. [d]*N* = 186.

5.3.2 Analysis of Variance

A two-way between groups Analysis of Variance was conducted to explore the impact of GCSE grade group (high-grade 7-9; mid-grade 4-6) and the year the GCSE was taken (2019; 2020) on the two different ART scores (Table 12).

Table 12

Analysis of Variance Summary Table for Classic Author Recognition Test Score and Young Adult Author Recognition Test Score

Source	*df*	*F*	*P*	Effect size
Classic ART Score				
Year GCSE taken	2	2.71	.069	.017
GCSE attainment group	1	8.62	.004*	.027
Year x Group Interaction	2	1.43	.240	.009
YA ART Score				
Year GCSE taken	2	1.47	.233	.009
GCSE attainment group	1	3.31	.070	.010
Year x Group Interaction	2	2.24	.108	.014

Note. Effect size = partial $\eta 2$. * $p < .005$

The effect of the year GCSE was taken was not significant for either the classic or YA ARTs. There was a significant effect of attainment groups on the classic ART score but no significant effect of attainment group for the YA score, with students gaining higher English

language GCSE grades achieving higher scores on the classic ART than those gaining mid GCSE grades.

A two-way between-group Analysis of Variance was also conducted to explore the impact of GCSE attainment group and year on vocabulary scores (see Table 13). There were no significant effects of either the year the GCSE was taken or the GCSE attainment group on vocabulary score and no interaction between them.

Table 13
Analysis of Variance Summary Table for Vocabulary Score

Source	*df*	*F*	*p*	Effect size
Year GCSE taken	2	2.40	.092	.016
GCSE attainment group	1	.811	.368	.003
Year x Group Interaction	2	.992	.372	.007

Note. Effect size = partial *η2*.

5.4 Discussion

5.4.1 Reading Experience and Attainment

While we know that leisure (fiction) reading is crucial for developing vocabulary knowledge and comprehension skills in general, we know much less about how adolescent reading experience relates to attainment. The aim of this study was to explore whether high attaining English language GCSE students, in England, differed from mid attaining students in their vocabulary scores and in the quantity and type of their reading experience. In particular, because the new English Language GCSE exam specification requires the inclusion of literature and literary non-fiction texts from all of the 19th, 20th and 21st centuries, the current study differentiated between students' reading experience with classic authors and YA authors to see if there was a difference between types of reading experience and performance in the exam.

As noted above, the responses to the survey were not representative. The respondents to the survey were disproportionately female, from a high socio-economic group and extremely high attaining in English language GCSE. The results, therefore, can only be generalised to relationships between attainment and reading experience for this particular group. Whilst the high-grade group had higher mean scores on the vocabulary test than the mid-grade group, the difference between the groups' scores was not significant. The low internal consistency of the test or a ceiling effect, possibly caused by the relatively small number of items, could be a simple explanation for the lack of significance. It is possible that had the test been more sensitive a relationship may have been found with the GCSE grade groups, as was predicted. It could also be that, because the English language GCSE grade represents much more than just vocabulary knowledge, reading experience is a better measure.

As measures of print exposure, ARTs have traditionally been constructed to measure reading outside of school or college, so have avoided including classic authors from the literary canon and have focused on creating and using contemporary and popular authors (Martin-Chang & Gould, 2008; Stanovich & West, 1989). However, for this study, recognition of classic and older authors was important to the measure as they represented the types of authors that are chosen for the exam texts. Recognition of classic author names could perhaps be because of their high profile culturally, rather indicating that students have actual reading experience with them. For example, the high recognition of Shakespeare, Dickens and Austen likely points to their high cultural profiles, but the YA authors J K Rowling and Jacqueline Wilson had similarly high recognition. Participants are likely to have read some of the classic authors for their English literature GCSE, but this would only account for a small number of authors on the list as the literature exam only requires the study of: 1) a Shakespeare play; 2) a nineteenth century novel; 3) a modern (post 1914) text and 4) a short

poetry anthology. The significant relationship between the classic author ART scores and the YA ART scores, as well as the significant relationships between the self-reported reading measures (average minutes spent reading and comparative time and enjoyment to peers) indicates that the classic author ART score is likely to be representing reading experience. It is worth noting that both the YA and classic ART scores were strongly correlated with the vocabulary scores, the YA ART slightly more so. This supports the literature on the close relationship between reading experience and vocabulary knowledge (Nation, 2017; Perfetti & Hart, 2002).

The significant effect of GCSE grade group on performance in the classic ART, but not the YA ART, when it was predicted that there would be a relationship with both, suggests that it is the higher levels of reading experience with classic authors that relates to attainment in this particular exam. Whilst this intuitively fits with the inclusion in the exam of older literary texts, the need to match reading experience to assessment texts is not something generally acknowledged in the literature. Work that promotes reading for pleasure suggests promoting student choice in their independent reading rather than matching reading to assessments (Casey, 2010; Gambrell et al., 1996; Moss & McDonald, 2004). This study, in line with previous surveys (Topping et al., 2023), finds that this age group has a preference for modern fiction when reading for enjoyment or pleasure. If it is the work of classic authors, however, that has the stronger relationship with the attainment of higher grades in this new qualification, then we have a misalignment that it is important to acknowledge.

5.4.2 Implications for Schools and Curricula

If, as the findings above suggest, students in the high-grade group have more reading experience with classic authors, then this could lead to these authors being promoted in schools and students encouraged to read more of them, in order to emulate the success of this high attaining group. To recommend that students read more classic literary texts is not

though without controversy. Whilst a good grounding in the literary canon might be seen as unapologetically aspirational by some, there is already criticism of literature curricula's lack of diversity and representation (V. Elliott et al., 2021), as the canon is dominated by the writings of traditionally privileged groups – sometimes labelled *dead White men* or as *stale, male and pale*. The recognition that the construction of the canon is sociocultural and ideological (Ervin, 2022; Löffler, 2017), rather than being aesthetically objective, has led to efforts to diversify curricula choices and teaching as well as integrating critical engagement with the ways in which texts are selected (Centre for Literacy in Primary Education, 2023; Dyches, 2018; Ervin, 2022; L. Johnson, 2018; Picton & Clark, 2022). It is difficult, therefore, not to see a turning back to the traditional authors of the canon as a retrograde step. Analysis of the 2020 National Literacy Trust annual survey (Picton & Clark, 2022) showed that children reported that reading books that had characters like them could increase their confidence, especially if they were from an ethnic minority background. Many children also felt it was important to be able to read about characters that were different from themselves. The Centre for Literacy in Primary Education's 'Reflecting Realities' report (2023), showed that, whilst there have been improvements, there is still a need for more multi-dimensional characterisation in order to offer full diversity of representation in children's books.

The English language GCSE purports to be focused on students' reading and writing abilities across a range of texts (Department for Education, 2013a). If the qualification was designed to demonstrate the level of students' general reading ability and indicate their reading competence for future study and employment, then perhaps the choice of texts specified for use in the exam needs to widen to include non-literary genres. Other genres of texts will be important to students' later prospects, achievements and successful adult lives, for example: academic texts for higher education courses; instructional texts for employment; advice texts for health and family; news for civic engagement; and, of course, the ability to

critically engage with a variety of forms of new (social) media. The requirement to solely select literature and literary texts for the exams, will likely push other genres of reading to the boundaries of the curriculum and class time.

Self-reported measures in this study indicated that students, from the narrow demographic group of respondents at least, mostly enjoy reading modern fiction and this is the genre that they choose to read for pleasure outside of their studies. If schools feel duty bound to recommend a narrow and dated diet of texts, in order for students to be prepared for the high-stakes English language GCSE exam, then there is a danger that the numbers of young people who say they enjoy reading may continue into a greater crisis of decline.

5.5 Limitations and Future Research

While the current findings clearly show that reading experience, of the respondents to this survey, differs between those who gained high and mid grades in their English language GCSE, there are some limitations to this study when interpreting these results. The responses to the survey were only from a narrow demographic group and were not representative of all levels of attainment, genders and socioeconomic groups, which did not allow for a full comparison between a range of attainment groups. In addition, the impact of COVID-19 and the use of CAGs or calculated grades to award the final grades for the 2020 cohort also limited the extent to which the data represented reading comprehension of the new types of texts in the exam and limited the power of the statistical tests. It would also have been useful to have used a more detailed measure of reading comprehension to compare to the GCSE attainment groups, although lengthening the online test could have impacted the representativeness of the sample further. The vocabulary test, based on the keywords from previous exams, also needs further development to increase its reliability.

Future research should try and reach respondents for whom reading is not as socially desirable, who come from less privileged backgrounds, and who have lower attainment in

their English language GCSE. Future research should also further explore the relationship between reading experience and a range of attainment measures, including but not limited to public exams. The choices that students make about what to read outside of the school curriculum is also an area that is understudied and could provide valuable data about types of reading experience and their impact. The nature of curriculum reading, highlighted by the difference between what was reportedly read in school or college and what was read outside for pleasure, also warrants further investigation. The impact of reading experience with classic authors rather than YA authors also warrants further research. Maybe there is something inherent to classic literature, possibly it provides exposure to better vocabulary, that improves comprehension or attainment.

5.6 Conclusion

This study sought to explore, through the comparison of attainment groups, the difference in the reading experience between groups of students attaining high grades and those obtaining mid grades in the new English language GCSE. Although responses to the survey were not representative, and there was an impact of COVID-19 on the 2020 cohort, useful results were obtained. Even though the literature shows a strong relationship between reading for pleasure and reading comprehension proficiency, it was reading experience with classic authors, rather than YA authors, that had a significant difference between the two attainment groups. The impact of these findings are that: 1) they increase the evidence for the importance of measuring reading experience in a nuanced way; and 2) they highlight the impact that the choices of texts used in assessments can have. Whilst this may suggest that students should be encouraged to read more work by classic authors, questions of enjoyment, independent choice, diversity, and representation should not be ignored.

Chapter 6. Paper 3

Adolescent Reading Experience, Independent Choices and Curriculum Materials: Redefining the Deficit Narrative

Authors: Beverley Jennings, Daisy Powell, Sylvia Jaworska, Holly Joseph

The paper reproduced here is intended for submission to *Applied Corpus Linguistics* in November 2024

Abstract

Reading comprehension ability is assessed in England within the English language GCSE exam. This is a high stakes exam, taken by all 16-year-olds, and a pass grade is needed to progress onto the next stage of education and employment. Since reading experience is an important predictor of reading comprehension ability, two different types of reading materials were explored to see how well they matched the reading required in the exam: 1) curriculum reading; and 2) independent reading. Two corpora of texts representing the two types of reading were created and explored using the methods of Corpus Linguistics. The independent reading corpus (IRC) had higher linguistic diversity, and higher frequency of adverbs but lower frequency of nouns, than the non-fiction curriculum reading corpus (CRC). Exploratory analysis of the most frequent parts of speech revealed that the IRC featured words about the concrete and the everyday, whereas the CRC had words that were more abstract and conceptual, suggesting that curriculum reading presents a different type of vocabulary challenge. The IRC was a better match to the exam texts than the CRC. As the English language GCSE exam is used as a measure of literacy competency for both future study and future employment, this suggests that the types of texts chosen for the exam are not a good match for this purpose. The choice of texts in assessments therefore needs careful consideration.

6.1 Introduction

In England, as part of a suite of General Certificate of Secondary Education (GCSE) exams, taken at the end of their full-time compulsory education, students (age 16) sit an English language GCSE that is taken to indicate their literacy competency and suitability for future study and employment. There is a separate English literature qualification to assess the critical analysis of literary texts. The English language GCSE exam was reformed by the Government in 2015, with the new exam introduced in 2017. This new specification changed

the form and age of the texts that have to be read in the exam and, instead of mostly modern and accessible texts in the old version (Isaacs, 2014), now texts are exclusively literature and literary non-fiction and have to be from all three of the 19th, 20th and 21st centuries. Students' ability to read and understand these types of texts is therefore an important area of research.

An important predictor of comprehension ability is reading experience (Acheson et al., 2008; Chateau & Jared, 2000; Davidse et al., 2011; Mol & Bus, 2011). This is explained by the lexical quality hypothesis (LQH) (Perfetti & Hart, 2002) as the gradual building of an increasingly secure and coherent, but also nuanced, understanding of words each time they are encountered. The lexical legacy hypothesis (LLH) (Nation, 2017) builds on this by specifying that encounters with words need to be in diverse contexts for greater quality to be built (Joseph & Nation, 2018; Pagán & Nation, 2019; Rosa et al., 2017, 2022). It is therefore important to examine students' actual reading experience, to understand how far it is providing exposures to words in order to build good vocabulary knowledge and comprehension ability, in preparation for the final exam.

6.2 Literature review

6.2.1 Reading Experience

The Simple View of Reading (SVR) (Gough & Tunmer, 1986; Hoover & Gough, 1990) describes reading as being the product of two parts, the ability to decode written words (either by sounding them out or by recognising them immediately) and the linguistic comprehension of the words. The importance of the two components of the SVR does not, however, remain consistent for readers across all ages. As students become more skilled, and their reading more proficient, the decoding element of the SVR (in which proficiency has been reached) declines in importance and the linguistic element becomes more important (Braze et al., 2007; Francis et al., 2005; Gough et al., 1996; Henderson et al., 2013; Nation & Snowling, 1998; Ouellette, 2006; Tilstra et al., 2009). This linguistic comprehension includes

vocabulary knowledge and, as the reading materials of secondary or high school education increase in difficultly, the language of the written text deviates more from spoken language (Braze et al., 2007; Cunningham, 2005; Landauer & Dumais, 1997; Tilstra et al., 2009) and includes more low frequency and exception words (that do not follow usual spelling rules) (Nation & Snowling, 1998). Perfetti and Hart's LQH (2002) defines the ability to read a word efficiently as when the reader is able to access high quality representations of three components of a word: its written form (orthography); its sound (phonology); and its meaning (semantic information). As specified by the LQH, building high-quality representations of words depends on experiences with them, each encounter enabling the components of high lexical quality to become more secure and coherent (Perfetti, 2007).

Reading experience is therefore an important predictor for reading ability (Acheson et al., 2008; Chateau & Jared, 2000; Davidse et al., 2011; Mol & Bus, 2011), and several studies show that it is fiction reading specifically that is a superior predictor of that ability (Mar & Rain, 2015; Martin-Chang et al., 2020; McGeown et al., 2015; Pfost et al., 2013). That reading experience predicts reading skill can be explained by the LLH (Nation, 2017) as it creates a bank of previous experiences with words that each reader has built up. If these experiences are diverse, then lexical quality is gradually increased through each new context or nuance of meaning encountered (Pagán & Nation, 2019; Rosa et al., 2017, 2022).

Whilst explicit teaching of vocabulary is, of course, an essential part of good classroom practice, using the theoretical background of the LQH and the LLH, it is clear that it is not enough to be taught words from lists. Instead, in order to build lexical quality (Perfetti & Hart, 2002) words must be experienced in diverse contexts (Nation, 2017). It has also been estimated that the number of words taught in classrooms each year is approximately 200-300 (Nagy & Herman, 1984), whereas the estimate of the number of words learnt by children each year is approximately 3000 (Nagy et al., 1987). The gap between words learnt

overall and words taught is filled, according to Nagy et al. (1987), by learning from context, that is through listening and reading. For older children most new words will be acquired through reading as they will have already encountered, by age 12, words that are found in spoken language (Landauer & Dumais, 1997). It is reading experience that is needed at this age, therefore, for a vocabulary growth to occur (Nagy et al., 1987).

Being able to read well and access the learning materials of the curriculum is crucial for students at the secondary levels of education (Shanahan & Shanahan, 2017). Analysis of the Programme for International Student Assessment research showed that 20% of 15-year-old students in England were below the reading level considered the minimum required to be able to participate in society (Ingram et al., 2023). The 2023 national Statutory Assessment Tests, taken in England by students at the end of their primary (elementary) education (age 11), showed even lower levels of proficiency, with only 73% of students meeting the expected standard in reading (Department for Education, 2023c). As adolescents progress through school, subjects are taught more discretely and reading materials become more complex, use increasingly specialised and more academic language (Schleppegrell, 2001, 2007). Many different types of words, for example technical or subject-specific vocabulary and also words that are used for cohesion like connectives, are found more frequently in written language than in spoken. It is reading experience therefore, that will provide encounters with this type of vocabulary (Tilstra et al., 2009).

6.2.2 Curriculum reading

Corpus studies of vocabulary in education have tended to focus more on higher education than on schools (Coxhead, 2000, 2011; Gardner & Davies, 2014). The main focus of these studies has been on creating lists of academic words (Coxhead, 2000; Gardner & Davies, 2014) and of disciplinary language (Hyland, 2008, 2017; Hyland & Tse, 2007). Some similar work has been done in secondary schools with the creation of lists of school

vocabulary and phrases (Green & Lambert, 2018, 2019). There have also been corpus studies of the language used in maths resources (Monaghan, 1999) and science textbooks (Coxhead et al., 2010; Deignan & Love, 2019). These studies and lists have provided teachers and students with valuable teaching and learning resources. However, as shown above (Nagy & Herman, 1984), being taught or learning words from lists in class is not sufficient for vocabulary to grow adequately. Diversity and meaningful context are lacking in lists of words or phrases, compared to the reading experience required to build lexical quality (Nation, 2017; Perfetti & Hart, 2002). Texts that students read, as part of their classes, form part of each students' bank of prior reading experience, or lexical legacy (Nation, 2017). Studying samples of class reading can therefore provide useful data about the vocabulary that students are (and are not) exposed to through the curriculum. The school studies outlined above have relied on collecting text from curriculum textbooks, to represent what is read in the classroom. However, the increasing use of technology, both from teachers' use of slides, worksheets and online quizzes in the classroom, and students' increasing use of their own devices and electronic resources, mean that textbooks can no longer be taken as a good example of the kind of reading that students are expected to do and are exposed to through the curriculum.

6.2.3 Independent Reading

Corpus studies of children's non-curriculum reading (reading for pleasure) have compared book language to spoken language and found that book language is more complex (Cameron-Faulkner & Noble, 2013; Dawson et al., 2021; Hsiao et al., 2022; Montag, 2019; Montag et al., 2015; Montag & MacDonald, 2015). These studies have generally used existing collections of texts written for children, like the Oxford Children's Corpus (Wild et al., 2013) or a children's reading subset of the Corpus of Contemporary American English (University of Arizona Libraries, 2021).

A report into the results of a large national (UK) survey by the National Literacy Trust (Clark et al., 2023) about children's reading practices shows that only 43.4% of children (aged 8-18) said they enjoyed reading, the lowest level recorded since the survey started in 2005. The number of young people enjoying reading drops as age increases. In the same survey only 28% of respondents said that they read daily in their free time, which followed the trend of gradually decreasing numbers since 2005. Fiction is still the most popular choice for free time reading (73.5%), but there was no further detail on the types of fiction that were being read. In an earlier report Clark and Rumbold (2006) showed that children's choices, when reading for pleasure, were diverse but that fiction dominated. A report based on data from the school reading programme, Accelerated Reader (Topping et al., 2023), for readers from the UK and the Republic of Ireland attending secondary school years 9-11 (age 13-16), showed that the most read titles were either fiction books that were likely to have been studied in class (e.g. *Of Mice and Men*), or titles by children's and YA authors (e.g. J.K. Rowling). It should be noted that this data will be affected by the books that are stocked by school libraries and the books that are listed on the Accelerated Reader platform itself.

6.2.4 The English Language GCSE

The new specification of the English language GCSE was first taken by students in England in 2017. A Grade 4 (equivalent to a C) in this qualification is needed by students to access most post-16 options, including further study, apprenticeships and employment. It is a government funding requirement for post-16 courses that any students who did not gain a Grade 4 or above, must continue to study English and ideally retake the qualification. GCSE results, with special focus on maths and English language, are published each year and are used as a measure by which to judge the quality of education being provided by each school. This means preparing for this exam is important for students, teachers, and schools. A corpus study of a selection of exams texts from the new English language GCSE identified 146

keywords, that appeared more frequently in the exam text corpus (ETC), created for the study, than a reference corpus and therefore were taken to typify the vocabulary in the exam texts (B. Jennings et al., 2024). These keywords were low in frequency in general language and were typically found in fictional texts, especially older classic fiction. The LQH and LLH show that if these words are to be understood, then students must have experienced them in their prior reading. Identifying the vocabulary content of students' reading experience therefore becomes key.

As the ability to proficiently comprehend the vocabulary in a text depends on previous reading experience having provided enough diverse exposure to that vocabulary (Acheson et al., 2008; Chateau & Jared, 2000; Davidse et al., 2011; Mol & Bus, 2011; Nation, 2017; Perfetti & Hart, 2002), adolescents preparing for their English language GCSE exam will be relying on their previous reading experience to enable their comprehension of the exam texts. That reading experience may have been gained inside and/or outside school. Previous corpus studies of academic language have focused on producing lists of vocabulary that are either common across disciplines (Coxhead, 2000; Gardner & Davies, 2014) or needed within disciplines (Green & Lambert, 2018, 2019; Hyland, 2008, 2017; Hyland & Tse, 2007). For independent reading or reading for pleasure outside school, there is evidence that fewer students, especially in this adolescent age group, choose to read in their free time (Clark et al., 2023). When children and young people do choose to read independently, fiction seems to remain the most popular choice. This is key for reading proficiency as previous research shows that fiction is a superior predictor of reading skill (Mar & Rain, 2015; Martin-Chang et al., 2020; McGeown et al., 2015; Pfost et al., 2013). It is also important to note that corpus studies of children's reading for pleasure usually depend on using collections of texts that are based on the target age for readers of the texts, rather than from any data about what children or adolescents are actually choosing to read. Although the data that we do have

would suggest that this is likely to be children's or YA fiction ((Topping et al., 2023), it is crucial to find out what young people are actually choosing to read so that we can have a more accurate picture of the vocabulary they are encountering, rather than just the vocabulary that they would encounter if they read the books and genres targeted at their age group.

6.2.5 This Study

The focus of this study is the vocabulary content of the reading experience of adolescents, both across the curriculum at school and in any independent reading. The intention was to collect a small but manageable number of texts and to carry out an exploratory analysis. One aim was to look at a sample of text drawn from a range of curriculum reading at school, rather than just have word lists or collections that only represent single or limited numbers of subjects, as previous corpus studies have done. A second aim was to add to the primary data on adolescent reading by creating a sample of students' actual independent reading, rather than looking at a collection that is defined by suggested age ranges or specific genres. It would then be possible to compare this new collection of adolescent reading materials to the corpus of exam texts created in a previous study (B. Jennings et al., 2024). This study therefore created two corpora of texts to explore students' actual reading materials: 1) from lesson materials, to explore curriculum reading in school; and 2) from students' independent reading, outside of school.

In order to explore the different reading experiences offered by the two different genres of reading (curriculum and independent), this study examined the linguistic content of the two new corpora created. The occurrences of different parts of speech were compared, as these can be an indication of linguistic register and could therefore suggest the types of registers present in each corpus (Biber et al., 1999). The lexical diversity of the two corpora were compared as a measure of linguistic richness, a high lexical diversity score indicates that there are more unique words in the text. This is important for reading experience as a

higher lexical diversity will provide more encounters with different words and therefore have the potential to build greater lexical quality (Nation, 2017; Perfetti & Hart, 2002) with a greater range of words. The most frequent words in the two corpora were then compared, these most frequent words lists were separated into the four main parts of speech to enable a close comparison. Again, this was a useful method to use to consider the reading experiences and potential vocabulary encounters offered by the two different types of texts.

The level of difficulty presented by the words on the most frequent words lists from the two corpora was analysed by comparing the average number of letters in the words. Longer words have been shown, by eye tracking studies to have longer reading times (e.g. Joseph et al., 2009), and this can impact comprehension due to the increase in processing time (Martin-Chang et al., 2020). The level of difficulty for nouns was measured using concreteness and imageability scores. Words with higher scores for these two measures are easier to comprehend as the reader can draw on perceptual memory (Brysbaert et al., 2014; Cortese & Fugett, 2004; Khanna & Cortese, 2021; Sadoski & Paivio, 2013). Using these three measures (word length, concreteness and imageability) allowed for a comparison of the difficulty of the words in two focus corpora, an important indication of the kind of reading experience being offered by them.

6.2.6 Research Questions:

1. What is the linguistic make-up of the corpora of students' independent and curriculum reading?
2. What types of words typify the student reading material that were collected?
3. How far do the student reading materials match the vocabulary in the English language GCSE exam?

6.3 Methodology

6.3.1 Ethical Approval

This study was granted ethical approval by the University of Reading's Institute of Education.

6.3.2 Independent Reading

Retrospective opt-out permission was used to access a list of reading materials submitted by students in a year 10 (age 14-15) mixed ability English class for a free choice reading homework task over a half-term holiday. Of the twenty-five students in the class: twenty-three submitted what they had read for homework (two students did not complete the original homework task); and twenty-one did not opt-out. One book was submitted twice (*One of Us is Lying* by Karen McManus), this left a list of twenty different source texts. These twenty texts consisted of: nine young adult (YA) fiction books; three newspaper articles, two autobiographies; two classic children's books; two modern literary fiction books; one crime/thriller fiction book; and one classic literary fiction book (see Table 14). One thousand words from the beginning of each text were collected, as Biber (1990) showed that 1,000-word sub samples from texts, when compared, had high level of linguistic stability.

Table 14

Texts used to create Independent Reading Corpus.

No.	Text	Genre
1	*Harry Potter and the Chamber of Secrets* by J. K. Rowling	Young Adult Fiction
2	*Checkmate* by Malorie Blackman	Young Adult Fiction
3	*Harry Potter and the Deathly Hallows* by J. K. Rowling	Young Adult Fiction
4	*One Of Us Is Lying* by Karen McManus	Young Adult Fiction
5	*Rule of Wolves* by Leigh Bardugo	Young Adult Fiction
6	*Divergent* by Veronica Roth	Young Adult Fiction
7	*Twilight* by Stephenie Meyer	Young Adult Fiction
8	*Harry Potter and the Philosopher's Stone* by J. K. Rowling	Young Adult Fiction
9	*The Maze Runner* by James Dashner	Young Adult Fiction
10	News article from online daily newspaper for young people	Non-fiction (news)
11	News article from online daily newspaper for young people	Non-fiction (news)
12	Sports article from an online newspaper	Non-fiction (news)
13	*The Storyteller: Tales of Life and Music* by Dave Grohl	Autobiography
14	*I am Malala* by Malala Yousafzai	Autobiography
15	*Biggles of the Camel Squadron* by W. E. Johns	Classic Children's Fiction
16	*The BFG* by Roald Dahl	Classic Children's Fiction
17	*Everything I Never Told You* by Celeste Ng	Modern Literary Fiction
18	*Woman in Black* by Susan Hill	Modern Literary Fiction
19	*Body Language* by A. K. Turner	Crime/thriller
20	*The Great Gatsby* by F. Scott Fitzgerald	Classic Fiction

6.3.3 Curriculum Reading

Curriculum materials were collected for year 10 classes (age 14-15) from an online platform used by teachers to share resources with their classes (Google Classroom). A week in June was chosen for expediency and resources were downloaded from each class. Resources were accessed from the following 16 subjects: Art, Computing, Media, Technology, English, Geography, History, Maths, Music, Physical Education, Religious Studies, Biology, Chemistry, Physics, Childcare, Graphics. Collecting data from a range of

subjects across the curriculum is important because students' experience with words and their lexical legacy (Nation, 2017), is formed by all their experiences with text, not just from subjects like English where reading is explicitly being taught. With recent reports suggesting that only 28% of children read every day in their free time (Clark et al., 2023), reading within the curriculum may represent the only reading that some children do, so the full range of subjects is essential to study. The types of resources downloaded included: worksheets; slides; pages from textbooks; quizzes; exam questions and answers; and coursework tasks. In subjects where there were more than 1000 words (11 subjects), the first 1000 words were taken as representative (Biber, 1990). Five subjects had less than 1000 words (see Table 15).

There were some challenges in converting the documents that were shared on the online platform into text files that were suitable for uploading to the corpus tool, *Sketch Engine* (Kilgarriff et al., 2014). Slides often used pictures and graphics with the text presented in separate boxes, so the process of extracting the text was difficult to automate. There were similar challenges with PDF files, pages from textbooks, exam papers, and worksheets; where the design and presentation of text meant that many manual adjustments were needed when converting the format. Considerable time was therefore needed to create a relatively small corpus.

Table 15

Curriculum Subjects in the Curriculum Reading Corpus with Word Counts

Subject	Word Count
Art	206
Computing	1000
Media	909
Technology	1000
English	1000
Geography	1000
History	1000
Maths	1000
Music	603
Physical Education	1000
Religious Studies	1000
Biology	437
Chemistry	1000
Physics	1000
Childcare	1000
Graphics	276

6.3.4 Creation of the Corpora

Two corpora were created from the texts collected: 1) the IRC, using the 20 independent reading documents collected from the homework task; 2) the CRC, using the 16 curriculum documents. Details of the two corpora are given in Table 16. Documents were uploaded to the corpus tool *Sketch Engine* (Kilgarriff et al., 2014).

Table 16
Corpora Contents

Corpus	Documents	Tokens	Words	Types (Unique Words)
Independent Reading Corpus	20	25,467	21,553	5,169
Curriculum Reading Corpus	16	15, 574	13,210	3,356

6.3.5 Frequencies of Parts of Speech and Lexical Diversity

To answer RQ1, what is the linguistic make up of students' curriculum and independent reading, total occurrences for nouns, verbs, adjectives, adverbs and other parts of speech were calculated for the IRC and CRC. A comparison of frequencies of parts of speech showed how the linguistic make up of these two corpora differed. In order to compare totals between corpora that are not the same size, frequencies need to be normalized. This was calculated by converting raw scores to frequency per million (fpm) (raw occurrences of part of speech/total words in corpus x 1,000,000). Chi squared tests of independence were used to compare whether differences between the frequencies of the parts of speech in each corpus were significant. Lexical diversity, which is measure of how many different (unique) words are used in a corpus was calculated using a type to token ratio (TTR) (Jarvis, 2013; Richards, 1987). This measure showed which of the two corpora contained the most unique words and therefore could potentially be a richer source of reading experience. A simple TTR can be calculated by dividing the number of types (unique words) by the number of tokens (total words) within a text or corpus, with higher scores representing higher diversity. However, this calculation does not account for the impact that the length of a text will have on this ratio (Covington & McFall, 2010; Kyle et al., 2021). To account for the sizes of the corpora, a moving-average type-token ratio (MATTR) (Covington & McFall, 2010) was calculated for

both corpora using the MATTR computer program (Covington & McFall, 2008) which averages the TTR for every rolling 500 words.

6.3.6 Most Frequent Words

Frequent words are important to study as these are the words that students are most likely to encounter in these different types of reading experiences and therefore gave an indication of how the vocabulary in the reading texts might be different. Word lists, which rank words by their frequency in the corpus, were produced from *Sketch Engine* for four parts of speech (nouns, verbs, adverbs and adjectives) from the IRC and CRC corpora. The parts of speech labels were allocated to the words in the corpora through the automatic tagger in *Sketch Engine*. The 100 most frequent words of each of the four parts of speech from the two corpora were compared using the MRC psycholinguistic database (Coltheart, 1981). The first measure used was word length to identify if there were significant differences in word lengths between the word lists from the two corpora. Word length is compared as longer words are an indication of greater difficulty (Carver, 1976). Longer words can also lead to longer processing times which can have a negative impact on comprehension (Martin-Chang et al., 2020). The two lists of the 100 most frequently occurring nouns were then compared to see if there were significant differences in concreteness and imageability. Concreteness is measure of the closeness of what the word denotes to a "perceptual entity" (Brysbaert et al., 2014, p. 904). A word that has a high concreteness score is understood to be easier to process because perceptual memory can be used, as compared to abstract words where it cannot (Brysbaert et al., 2014; Khanna & Cortese, 2021). Imageability scores give a measure of the extent to which the word is related to the senses and the formation of a mental image (Sadoski & Paivio, 2013). High scores for imageability indicate that the word is easier to process (Cortese & Fugett, 2004; Khanna & Cortese, 2021).

The word lists from the two corpora were then compared to identify which occurrences, in the 100 most frequent words in each of the four parts of speech, were common to both corpora and which occurrences were only in one of the corpora. Qualitative analysis was then conducted to further describe and compare the words on these eight lists.

6.3.7 Corpora Comparisons

In order to see how far students' prior reading, represented by the IRC and CRC, matched the texts that they would need to comprehend in their English language GCSE (RQ3), a comparison was run in *Sketch Engine* (Kilgarriff et al., 2014) between the two corpora created for this study, the ETC created for a previous study (B. Jennings et al., 2024) and a range of reference corpora. The *Sketch Engine* comparison tool compares the keyword scores (frequency per million in the focus corpus divided by the frequency per million in the reference corpus) of the 5000 most frequent words in each corpus and then creates an overall comparison score from the mean of the highest 500.

6.4 Findings

6.4.1 Parts of Speech in the Corpora

Raw numbers and fpm are reported for occurrences for each part of speech in both corpora created for this study and for the ETC created in a previous study (B. Jennings et al., 2024) (see Table 17).

Table 17

Raw Occurrences and Frequency per Million for Parts of Speech in the Corpora

Corpus	Nouns (fpm)	Verbs (fpm)	Adjectives (fpm)	Adverbs (fpm)	Other[a] (fpm)
Independent Reading Corpus	5,576 (258,711)	3,945 (183,037)	1,625 (75,396)	1,351 (62,683)	9,056 (420,174)
Curriculum Reading Corpus	4,467 (338,153)	2,424 (183,497)	1,025 (77,593)	406 (30,734)	4,888 (370,023)
Exam Text Corpus	8,191 (223,890)	7,085 (193,659)	2,535 (69,291)	2,396 (65,491)	16,378 (447,670)

Note. [a] other includes: conjunctions, prepositions, pronouns and numerals

A chi-square test of independence was performed to examine the relationship between the frequency of different parts of speech in the IRC and CRC. For adjectives, the difference was not significant, χ^2 (1, N = 34,763) = 0.56, p = 45. For verbs, the difference was also not significant, χ^2 (1, N = 34,763) = 0.01, p = 91. However, for adverbs, the difference was significant, χ^2 (1, N = 34,763) = 174.21, p = < 0.01: adverbs were significantly more frequent in the IRC compared to the CRC. For nouns, the difference was also significant, χ^2 (1, N = 34,763) = 251.60, p = < 0.01, with nouns significantly more frequent in the CRC compared to the IRC.

Reference data on the frequencies of parts of speech in different registers (Biber et al., 1999) identifies verbs and adverbs being most common in conversation and fiction, nouns as being most common in newspaper language and then academic prose, and adjectives being most common in academic prose and then newspaper language. Whilst none of the four registers used in Biber et al. (1999)(conversation, fiction, newspaper language and academic prose) are a complete match for the make-up of the IRC and CRC, the frequencies of the parts of speech in them generally follow the same pattern. The IRC, which contains mostly fiction, some narrative non-fiction and three newspaper articles, had a significantly higher

frequency of adverbs, which fits with adverbs being most common in fiction. In contrast, the CRC had significantly more nouns, as is found in newspaper language and academic prose. The frequencies were closer for verbs and adjectives, perhaps due to the mix of registers contained in the two corpora.

6.4.2 Lexical Diversity

Lexical diversity, measured by MATTR was slightly higher in the IRC (0.55) than the CRC (0.46), suggesting that the independent reading (mostly fiction) had a higher lexical diversity than the curriculum reading.

6.4.3 Comparing the Most Frequent Words in the Different Parts of Speech in the Two Corpora

Independent-samples t-tests were conducted to compare the number of letters, as an indication of difficulty, in the 100 most frequent words for each part of speech from the two corpora. Nouns in the IRC contained on average fewer letters than nouns in the CRC, and this was also the case for verbs and adjectives. For nouns there was a significant difference between the IRC (M = 4.87, SD = 1.35) and the CRC (M = 5.48, SD = 1.99), t (158) = 2.45, p = .02, two-sided. The effect size was small, with a Cohen's d of .36. Verbs had significantly fewer letters in the IRC (M = 4.37, SD = 1.20) than in the CRC (M = 5.44, SD = 1.84), t (171) = 4.87, p = < .001, two-sided. The effect size was medium, with a Cohen's d of .69. Adjectives also had significantly fewer letters in the IRC (M = 5.18, SD = 1.85) than in the CRC (M = 6.22, SD = 2.27), t (185) = 3.49, p = < .001, two-sided. The effect size was medium, with a Cohen's d of .50. For the length of adverbs there was no significant difference between the IRC (M = 5.63, SD = 2.00) and the CRC (M = 6.04, SD = 2.59), t (193) = 1.25, p = .211, two-sided. The effect size was small, with a Cohen's d of .18.

An independent samples t-test was also conducted to compare the concreteness and imageability scores of the 100 most frequent nouns in both corpora. Higher scores for both these attributes suggest lower difficulty. For concreteness the score was significantly higher for the IRC (M = 507.57, SD = 99.99) than the CRC (M = 450.54, SD = 100.18), t (131) = -3.28, p = .001, two-sided. The effect size was medium, with a Cohen's d of .57. For imageability the scores were also significantly higher for the IRC (M = 532.16, SD = 82.79) compared to the CRC (M = 469.48, SD = 91.15), t (132) = -4.17, p = < .001, two-sided. The effect size was medium, with a Cohen's d of .72. The lower scores in the CRC indicate that the words in the curriculum texts would be more difficult to comprehend.

Qualitative exploratory analysis, on the 100 most frequent words for each part of speech in the two corpora, was then conducted to identify any similarities and differences between them.

6.4.3.1 Nouns. Nouns are the most frequent word class (Biber et al., 1999) so it was perhaps to be expected that there was a high diversity of occurrences in the two noun frequency lists. Only 13 of the same nouns occurred in both corpora's top 100 for frequency. The nouns that appeared on both 100 most frequent word lists were all high frequency nouns and were concrete entities (e.g. *school*, *queen*) and qualities and states (e.g. *time*, *year*, *word*, *day*). In the IRC top 100 nouns by frequency (Appendix F) there were 21 proper nouns, 20 of which were for people (e.g. *Harry*, *Voldemort*, *Lydia*, *Drogba*) and one for a place (*Chelsea*). 18 common nouns were for domestic or everyday objects (e.g. *house*, *room*, *table*, *car*), eight common nouns were for the body or parts of it (e.g. *eye*, *hand*, *hair*, *head*), and 12 common nouns were for people (e.g. *man*, *mother*, *queen*, *brother*). Of the 22 nouns that denoted qualities or states, just over half related to time (e.g. *year*, *day*, *moment*, *night*). The CRC top 100 nouns by frequency (Appendix G) had far fewer proper nouns (5) (e.g. *London*, *Essex*, *Elizabeth*) than the IRC (21). Common nouns were very different to those in the IRC, with

very few about the domestic or everyday objects and more that were materials (e.g. *metal*, *copper*, *carbon*), were about space (e.g. *galaxy*, *earth*, *universe*, *sun*) or were to do with the classroom (e.g. *paper*, *line*, *mark*). There were more nouns that were about qualities or states, but time was not as dominant as in the IRC (only 2 - *day* and *year*), instead the nouns of quality or state were wide ranging in topic (e.g. *probability*, *aggression*, *personality*, *spectrum*).

6.4.3.2 Verbs. There were far more shared verbs in the top 100 frequency lists of the IRC (Appendix H) and the CRC (Appendix I) than there were for nouns, with 43 verbs appearing on both lists. These are mostly simple, high frequency actions and states verbs (e.g. *be*, *do*, *have*, *create*, *learn*). *Be* and *do* are also always likely to be very frequent due to their grammatical use in tense building. The verbs that only appear in the IRC top 100 are similar to the shared ones, mostly simple actions or states (e.g. *turn*, *feel*, *want*, *call*, *tell*). However, the verbs that appear in the CRC top 100 most frequent, that are not shared in the IRC top 100, are very different and include: verbs that are parts of instructions for class tasks (e.g. *explain*, *describe*, *write*, *extract*, *identify*); verbs that are part of a mark scheme or answer sheet (e.g. *accept*, *demonstrate*) and verbs that describe causation or relationships (e.g. *help*, *develop*, *involve*, *increase*, *produce*).

6.4.3.3 Adjectives. Just under half (42) of the adjectives were on both the IRC (Appendix J) and CRC (Appendix K) top 100 most frequent, and these were largely physical qualities (e.g. *red*, *long*, *small*, *green*, *big*) or simple qualitative attributes (e.g. *good*, *different*, *important*, *major*). The 58 adjectives in the IRC top 100 that did not appear in the CRC, were similarly about physical qualities (e.g. *little*, *tall*, *hard*, *black*, *pale*) and more complex qualitative attributes (e.g. *strange*, *magnificent*, *extraordinary*, *prominent*). The adjectives that only appeared in the CRC top 100 featured fewer that described physical attributes and the qualities described were more abstract (e.g. *relative*, *reactive*, *random*, *holistic*).

6.4.3.4 Adverbs. Adverbs were the part of speech that had the most crossover between the IRC (Appendix L) and CRC (Appendix M) top 100 frequency lists. This would be expected as it is the smallest word class (Biber et al., 1999). Those that appeared on both lists were simple adverbs, including of time and place (e.g. *now*, *then*, *back*, *down*), and of manner (e.g. *quickly*, *especially*, *directly*, *exactly*). Adverbs that only appeared in the IRC were also of time and place (e.g. *finally*, *soon*, *forever*, *upwards*, *behind*), but were mostly of manner (e.g. *obviously*, *purely*, *completely*, *barely*, *excitedly*). The adverbs that were only on the CRC list only contained more technical examples of manner (e.g. *randomly*, *artificially*, *extrinsically*, *functionally*, *aesthetically*).

6.4.4 Corpora Comparisons

The comparisons between corpora are presented in Table 18. The comparison score represents the mean of the highest 500 scoring keywords created by calculating the frequency per million in one corpus divided by the frequency per million in the other, the closer the score to 1 the more alike the corpora are.

Table 18

Comparison of Independent Reading Corpus and Curriculum Reading Corpus with Exam Text Corpus and Other Reference Corpora

	Independent Reading Corpus	Curriculum Reading Corpus
Independent Reading Corpus	~~1.0~~	5.56
Curriculum Reading Corpus	5.56	~~1.0~~
Exam Text Corpus	2.83	5.44
British National Corpus (spoken part)	4.66	7.34
British National Corpus	2.72	4.33
Brown Family	2.62	4.35
Project Gutenberg	2.73	5.02
English Web 2015	3.09	4.04
English Broadsheet Newspapers	2.88	4.51
Cambridge Academic English	3.89	4.30

Note. Comparison score is the mean of the highest 500 scoring keywords created by calculating the frequency per million in one corpus divided by the frequency per million in the other, the closer the score to 1 the more alike the corpora are.

The IRC is a closer match to all the other corpora than the CRC, suggesting that the CRC is very particular in its register. The IRC is a closer match to the exam text corpus than the CRC.

6.5 Discussion

Whilst the majority of the independent reading, chosen by the class of year 10 students, was, as expected, by children's and YA authors, there were exceptions with autobiographies, fiction written for adults, and newspaper articles included in the choices. From this small sample at least, this suggests that analysis of children's and adolescent's reading materials should not focus solely on texts that are targeted at their age group. This is especially important with the age group in this study, mid-adolescents, as they transition from reading books by children's and YA authors to more mainstream and general genres (e.g.

crime/thrillers) or to non-fiction genres (e.g. autobiography). Whilst acknowledging that the concept of genres is contested (Bawarshi & Reiff, 2010; Biber, 1990; Chandler, 1997; Sabao, 2014), and using genres to describe reading materials can only give an imperfect indication of the type of language that might be found within them, the different range of genres represented by the reading materials chosen does warrant attention.

The differences in the frequencies of nouns and adverbs in the two corpora, suggests that both types of reading, curriculum and independent, are important in a students' reading experience, as they contain different proportions of parts of speech. This suggests that the types of texts read will impact the number of encounters readers could have with different types of words. For example, if students only read curriculum texts, then they are less likely to have experienced a wide range of adverbs. The slightly higher lexical diversity of the IRC, as measured by the MATTR could also suggest that independent reading offers experience with a wider range of vocabulary than curriculum reading. As the numbers of students who read independently outside of school regularly is decreasing (Clark et al., 2023) this will mean that students who do not read fiction independently could potentially miss out on the most lexically diverse texts. It is important to note however that these results are from the comparison of two very small corpora and further research would be needed with larger collections of text to support these exploratory findings.

The significant differences between the number of letters in the 100 most frequent nouns, verbs and adjectives in the IRC and CRC, with the CRC nouns, verbs and adjectives having significantly more letters, suggested that, on the simple measure of word length, that the vocabulary challenge was higher in the CRC. Eye tracking studies have shown that, for adults and children, longer words have longer reading times (e.g. H. Joseph et al., 2009), and longer processing times for words can impact comprehension (Martin-Chang et al., 2020). The large overlap in the 100 most frequent adverbs in the two corpora probably accounts for

the lack of significant difference in the number of letters between the two corpora for this part of speech.

The significant differences between the concreteness and imageability scores for the 100 most frequent nouns in the two corpora, with the scores being higher in the IRC, again suggests that the challenge of the vocabulary is greater in the CRC. Concreteness and imageability can indicate the closeness of the meaning of a word to perceptual experience, the idea being that the closer the meaning of a word is to perceptual experience, the easier the word is to process (Brysbaert et al., 2014; Cortese & Fugett, 2004; Khanna & Cortese, 2021). Therefore, since concreteness and imageability scores were lower for the most frequent nouns in the CRC, it suggests that these are harder words to process as the words are further from perceptual experience.

The qualitative exploratory analysis, of the word lists of the 100 most frequent words in each part of speech (nouns, verbs, adjectives and adverbs), revealed interesting differences between the two corpora. In the CRC there was a particular vocabulary group that was specific to the classroom and learning tasks, both in nouns (e.g. *paper*, *line*, *mark*) and verbs (e.g. *explain*, *describe*, *write*, *extract*, *identify*, *accept*, *demonstrate)*. Not surprisingly, the CRC also had subject-specific tier three vocabulary (Beck et al., 2002) that was specific to topics being studied in the week the curriculum texts were collected (e.g. *metal*, *copper*, *carbon, galaxy*, *earth*, *universe*, *sun*). There was also more abstract vocabulary in the CRC, across nouns (e.g. *probability*, *aggression*, *personality*, *spectrum*), adjectives (e.g. *relative*, *reactive*, *random*, *holistic*) and adverbs (e.g. *randomly*, *artificially*, *extrinsically*, *functionally*, *aesthetically*), demonstrating the more theoretical content of the curriculum materials. Despite the IRC having higher lexical diversity, it could be argued that the challenge of the words in the top 100 most frequent word lists from the CRC was much higher. Not only was there a set of words that were specific to the classroom and learning tasks but also words that

require conceptual understanding, none of which were present in the top 100 frequency word lists of the IRC.

The whole corpora comparisons supported the findings, from comparing the parts of speech, that the two corpora were different linguistically. The CRC seemed particularly unlike any of the other corpora, even a corpus of academic English. This suggests that there might be a real particularity to curriculum resources in schools. The high scores, and therefore large difference, between both the IRC and CRC and the reference corpus of spoken language, supports the literature that reading is providing experience with different vocabulary to that which is experienced through listening (Braze et al., 2007; Cunningham, 2005; Landauer & Dumais, 1997; Tilstra et al., 2009). The IRC was a closer match to the ETC, that represents the vocabulary found in the English language GCSE, which is surprising as this qualification is meant to demonstrate proficiency for work and future study, rather than fiction reading ability, which is measured separately in the English literature GCSE. The closer relationship between independent (mostly fiction) reading and the corpus created from exam texts, suggests that it is the independent reading of fiction that is going to provide the best preparation for comprehending the reading texts in these high stakes exams. This is concerning as large numbers of students say that they do not read outside of school (Clark et al., 2023).

However, we also want to be careful not to create a deficit narrative with these findings. The word ‘gap’ has been an influential concept in education in England in recent years (e.g. Department for Education, 2023a; Ofsted, 2022; Quigley, 2018, 2020). This concept of groups of children or students having a deficit or ‘gap’, compared to other groups, dates back to B. Hart and Risley’s influential study (1995) in which they claimed that there was a thirty-million-word gap between the lowest socioeconomic group they studied (the ‘welfare’ group) and the highest (the ‘professional’ group). There has been further research on

this perceived ‘gap’ (e.g.Duff & Brydon, 2020; Fernald et al., 2013; Sullivan et al., 2021) and Cushing (2023) has shown that it has been a very influential concept in the English educational context from 2010 to the present day. However, there has also been extensive critique of the deficit narrative (Baugh, 2017; Cushing, 2023; García & Otheguy, 2017; E. J. Johnson, 2015), where the concept of a ‘gap’ is seen as positioning the linguistic practices of traditionally powerful and dominant groups above those used by more marginalised groups and defining the difference between their practices as the marginalised group’s deficit.

In order to avoid creating a simplistic deficit narrative in our findings, that of students’ lack of independent reading being judged as deficient, as far as preparation for the language in the exam is concerned, what should also be questioned or critiqued instead is the rationale behind the choices of what is included in the exam. As this qualification operates as a gatekeeper to future study, training and work opportunities, it is important to question any assumptions or value judgements about what have been deemed to be appropriate texts to include in the exam. Older, literature texts are now required, instead of the multi-modal and more deliberately accessible texts used in the past (Isaacs, 2014), non-fiction choices must be ‘extended literary’ and ‘transient’ (online) texts are specifically listed as not to be included (Department for Education, 2013a, p. 4). These choices reveal an inherent valuing of literary and traditional genre forms, to the exclusion of new and non-literary forms and genres, that are explicitly devalued. This narrow focus could be considered as much of a ‘deficit’ as any so called ‘gap’ in students’ reading. If experience of the vocabulary, found in the types of texts that have been specified for the exam, depends on the independent reading of fiction, then this could exclude students for all sorts of reasons. There can be financial, social or time barriers to adolescents accessing the kinds of reading materials that will most likely prepare them for the vocabulary in their exams and there are also huge swathes of alternative types of

texts and vocabulary that could be being read but are not currently being included in the texts in the exams.

What this exploratory analysis of the curriculum texts in the CRC has shown is that they contain a very particular set of vocabulary, that didn't match either the IRC, the ETC or any of the other reference corpora. This is obviously a very small corpus, that only represents one week of curriculum materials, so it's particular nature could be due to the high density of the specific topics covered in lessons that week. For example, in chemistry the topic was metals, in physics it was red shift and in history the 1601 rebellion by the Earl of Essex – all these topics featured in the top 100 most frequent word lists. However, there is no reason to think that these topics are not representative of the subjects from which they were taken. The policy ambition behind the construction of the new exams was 'to prepare young people better for the next steps in their education or employment' (Ofqual, 2013, p. 4). However, with the exam texts having so little in common with the curriculum materials collected for this study, it is hard to see how far the English language GCSE tests the comprehension abilities that will be needed for the curriculum materials in further or higher education – especially when it comes to the more abstract vocabulary found only on the CRC most frequent word lists. It is also hard to see how the close match to the vocabulary found in fiction links to the literacy needs of employers.

6.6 Limitations and Further Study

The limited scope of this study meant that the corpora created, and the findings generated, were only ever intended to be exploratory rather than representative. The very small sample of independent reading was collected from a just one class of students and a wider range of participants would be desirable in the future as different students may make very different reading choices. The curriculum materials accessed were also only from a very small number of lessons, that took place in just one week. A greater number of texts from a

greater number of lessons would create a larger corpus with which to test some of the initial findings from this paper. More sophisticated methods of extracting the text from highly designed formats like slides and PDFs could also help further study of classroom materials, as online resource formats continue to replace textbooks.

The difficulty with preparing the curriculum texts for uploading to the corpus tool highlights a research challenge now that classroom resources come in a wider variety of formats. With the growing use of slides and other formats that use sophisticated design features, collating and formatting classroom materials for corpus studies will be much more difficult than it was when there were standard textbooks that could be taken as a representation of what was being read in classrooms.

6.7 Conclusion

The English language GCSE is seen, in England, as an indication of a student's literacy ability and serves as a gatekeeping qualification for access to further and higher education and to employment and training. This study sought to create and explore two different types of reading that students are most likely to be exposed to: curriculum reading and independent reading. The curriculum reading was not as close a match for the vocabulary found in the exam texts as the independent reading. This suggests that unless students are reading independently outside of school, something that has been shown to be in decline, they will not have experience with, and therefore have had the chance to build sufficient knowledge of, the type of vocabulary that will be found in the exam.

However, instead of creating a simple deficit narrative, that some students are not reading enough independently or reading enough fiction, the choice of exam text should be critiqued too. The specification that the new exam should only have texts that are literature and literary non-fiction, prioritises and values one genre of reading over any others. Students' ability with a range of fictional texts is already assessed in the English literature GCSE,

instead of duplicating this valuing of fiction, maybe the English language GCSE should be filling in the 'gap' and including texts that are more like the curriculum texts that will be read in any future studies and also including texts that are common in the workplace and society. The exploratory analysis of the curriculum texts suggested that there may be a higher frequency of more abstract vocabulary, as well as a set of vocabulary that was exclusive to the classroom and learning activities. If the exam is used as an indication of having the reading skills needed for further study, then perhaps more vocabulary representative of curriculum materials should feature in the reading texts. There could also be an argument to consider other language practices, that will be useful in adult life, not just more formal and privileged language practices.

This exploratory study has shown that the collection and analysis of actual reading materials is possible, if challenging. Continued development in the methods and techniques of studying the content of reading experience, especially as it moves outside traditional formats, will help to improve our understanding of reading and reading content.

Chapter 7. Discussion

7.1 Contribution

This study has made contributions to the literature, to theory and to practice. There has been a contribution to primary data in the form of three new corpora and also new adolescent reading habit data collected from the survey. The contribution to theory in this study has been the use of methodologies, for example corpus linguistics and a genre specific ART, to explore reading experience. The contribution to practice has been the identification of the potential influence of assessment content on curriculum choices and on reading practices.

The primary data in the corpus of exam texts provided a new insight into the vocabulary challenge of the new English language GCSE exam texts. The two corpora of student reading materials were also new. The first student reading corpus, created from a sample of independently chosen student reading material, revealed the breadth of registers and genres that formed the actual reading of adolescents, as well as allowing for the analysis of the linguistic content of this reading. The second student reading corpus, a collection of reading materials from a wide range of curriculum subjects, contributed new data by representing cross-curricula reading rather than subject specific reading, which is more usual in the literature.

The survey of adolescent reading habits also makes several valuable contributions to the literature. Whilst there is already data about the independent reading habits of children and young people (e.g. Topping et al., 2023), this study also asked respondents to the survey about their genre reading within the curriculum, demonstrating that there is a key difference between these two areas. The two separate ART scores also contribute to the existing literature on the impact of different types of reading experience (Mar & Rain, 2015; McGeown et al., 2015; Pfost et al., 2013), demonstrating that there can be a difference

between the reading experience with different types of fiction read as well as between the reading of fiction and other genres.

The use of corpus linguistic methods to explore reading experience, using the LQH and the LLH as justification for its importance to building reading proficiency and skill, is a valuable theoretical contribution. The methods used in this study, including keywords, frequencies counts, corpora comparisons and qualitative analysis of vocabulary all generated new insights into both potential and actual adolescent reading experience. The positive relationship between the classic author ART and GCSE attainment also demonstrated a theoretical contribution to the ways in which this measure of reading experience can be used.

The final type of contribution that this study had made, that of contribution to practice, has been in the identification of the impact that the content of assessment material could have on curriculum content and the reading practices that are encouraged and endorsed. By demonstrating that some types of reading experience matched the exam texts better than other a potential for prioritising some kinds of reading over others was identified and its implications considered.

7.2 Overview of Results

The typical vocabulary in the English language GCSE exam texts was shown, in the first paper of this study (Chapter 4), to be low in frequency and to be most likely to be found in older, literary fiction. Since multiple, diverse encounters with a word are necessary to build the high-quality representations needed for good comprehension (Nation, 2017; Perfetti & Hart, 2002), this could mean that, in order to best prepare for the vocabulary demands of the exam, students might be encouraged or advised to read more older, literary texts. Focusing reading time on older, literary texts obviously has merits, for example for cultural capital and for knowledge and experience of historical contexts, historical vocabulary and historical syntactic structures, but these types of texts are already the focus of a second English exam,

English literature GCSE. A doubling up of the attention given to older literary texts means there will be less time for more modern texts and more modern forms and genres of texts, for example online information, potentially leaving students less than ready for the everyday literacy demands of adulthood. Another potential impact of the lack of time and attention to modern texts is that there are likely to be fewer opportunities for texts that have been written by or are about groups that have not been traditionally privileged or powerful.

The survey results, presented in the second paper (Chapter 5), supported the finding in the first paper - that there was a link between the vocabulary in the exam texts and older literary texts. Scores on an ART of classic authors were significantly higher for students who had attained the high grades (7-9) in their English language GCSE than for students who had attained mid-grades (4-6). However, the survey also revealed that respondents chose to read modern fiction when choosing what to read at home. A potential disparity was therefore revealed between the genre of reading that had a relationship with high attainment in the exam (older literary fiction – represented by the classic ART score) and the genre that students choose to read for their own independent reading (modern fiction). Whilst there is already evidence in the literature that independent reading for pleasure is beneficial to attainment (Clark & Rumbold, 2006; Kirsch et al., 2002; Spencer et al., 2017; Sullivan & Brown, 2015), examining the impact of different genres of reading is a novel contribution. This is a critical point because current recommendations around the promotion of independent reading focus on student choice and on diversity and representation (Casey, 2010; Gambrell et al., 1996; Moss & McDonald, 2004; Picton & Clark, 2022). However, attention to the detail of which genres of reading have the strongest relationship to attainment could compromise the work to diversify the content of what students read.

The third paper (Chapter 6) of this study focused more closely on the choices that students make in their own independent reading and the reading that they encounter across the

curriculum in school. The choices made in independent reading, from the small sample in this study, revealed the diversity in what is being chosen to be read. This suggests that, for this age group of mid-adolescents, reading experience is varied and disparate. The predominantly narrative texts in the IRC had frequent occurrences of vocabulary that focused on people, everyday objects and actions, and on the concrete. Exploratory analysis of the CRC showed that its most frequent vocabulary featured occurrences that were concerned with classroom and learning activities and also occurrences of more abstract vocabulary. This calls into question the decision to specify a focus on literary reading texts in a qualification that was designed to be, and is used as, an indication of a students' literacy competence for future study and employment. If the vocabulary of curriculum texts is substantially different to that of the literary reading texts, used to test comprehension in the exam, then its validity as a measure of preparedness for the demands of curriculum reading in post-16 study is questionable. Whilst the inclusion of older literary texts in the exam does increase the reading comprehension challenge, what these exploratory findings suggest is that it not the same challenge that is present in the curriculum materials. That students are able to meet the challenge of the vocabulary in the curriculum materials, should be important in a qualification that purports to measure reading ability for future study.

The political decision to specify that texts in the English language GCSE exam must be literary and from the 19th, 20th, and 21st centuries, not only risks prioritising older literary texts over more modern reading materials, but it also ignores the different linguistic content of academic and informational texts (Biber et al., 1999). The focus on literary texts means that the exam misses the mark in assessing the literacy preparedness of students for further study, work and productive citizenship.

7.3 Limitations

The limitations of this study are presented under two main headings: methodological and theoretical. The methodological limitations section covers: the size of the corpora created; the lack of representativeness of the respondents to the survey; the weaknesses in two of the measures in the survey (the vocabulary test and the ART); the impact of COVID; and the focus on one specific exam. The theoretical limitations section covers: the difficulties with collecting enough data to fully explore the two hypotheses used in this study (the LQH and the LLH); and the limitations of the focus on vocabulary.

7.3.1 Methodological Limitations

7.3.1.1 Corpora Size. The new corpora created for this study, in Paper 1 (Chapter 4) and Paper 3 (Chapter 6), were very specialised and small. At the time of its creation, the ETC used all available exam texts and so was representative as far as was possible, but it was still small (36,585 words) and represented only the first two years of content from a very new qualification. If the corpus was expanded, to include new exam texts as they become publicly available, then further analysis could prove informative. The two student reading corpora, the IRC and the CRC, were also very small (21,553 words and 13,210 words respectively) and could only represent small amounts of reading text. This was due to both the difficulty with collecting the data from participants, as it demanded high levels of cooperation, and also the difficulties with converting classroom resources (e.g. worksheets and slides) into text formats suitable to upload to a corpus tool. The findings from the linguistic analysis of the IRC and CRC therefore remain exploratory, rather than conclusive, but are useful in highlighting potential areas of interest and potential areas for future research.

7.3.1.2 Online Survey and Representative Participants. The use of an online survey for Paper 2 (Chapter 5) had the advantage of gaining relatively large numbers of responses in a timely and cost-efficient manner. However, this method of data collection

resulted in a sample that was unrepresentative: being disproportionately female, high attaining and from a high socioeconomic background. Finding alternative methods, that collect data from groups for whom reading may be less socially desirable or who do not have such high prior attainment in reading, would be advantageous in future research. This could be particularly challenging for research into adolescents as, due to the pressures of assessments, it is difficult to get agreement from educational settings to work with them. It is also difficult to get agreement from participants from this age group to agree to participate in research, as they often have many other pressures on their time, for example school assessments. This age group can also be generally less obliging and eager to please than younger participants and seemed to more readily take the legitimate option of not opting in.

It has been shown that there can be differences in outcome variables between groups of respondents ranked by whether they were easier or harder to reach, and therefore that there will be a difference in outcome variables for nonrespondents (Heffetz & Reeves, 2019). Traditionally there has been a bias towards participants from Western Industrial Educated Rich and Democratic (WEIRD) backgrounds (Henrich et al., 2010; Muthukrishna et al., 2020; Newson et al., 2021). As outlined in Chapter 3, the main components for successful response to an online survey were followed as far as was possible for the survey in this study. For example, keeping the survey short, ensuring the software chosen was compatible with completion on a mobile phone, communicating the purpose clearly, and providing information about protection of data (Evans & Mathur, 2018; Kılınç & Fırat, 2017). However, these strategies did not attract participants outside of the privileged group that did respond. It is unlikely, therefore, that simply circulating the survey to a larger number of students would have altered the nonresponse bias for any of the characteristics that were not representative of the population (e.g. gender, socioeconomic background, attainment). Future research in this area might have to use different data collection methods, for example matched samples

(Bethlehem, 2016) or Worldwide In-site Local and Diverse (WILD) strategies (Newson et al., 2021) to counter both low response rates in general and nonresponse bias in particular.

7.3.1.3 Vocabulary Test. The vocabulary test created for the survey in Paper 2 (Chapter 5) did not have high internal validity and ceiling effects were observed in this high-performing sample. The consequential weakness of this measure could account for the lack of a significant relationship between attainment and vocabulary scores. The vocabulary score was expected to be a mediator between reading experience and attainment in the English language GCSE, but the weakness of this measure meant it was difficult to reach a conclusion on this aspect of the study. A standardised vocabulary measure might have collected more reliable data, but it would not have aligned as closely to the vocabulary in the exam texts as the one uniquely designed for this study did.

The use of a standardised vocabulary test could be useful in future research to further explore the relationship between reading experience and attainment in the English language GCSE. Spencer et al. (2017) did find a relationship between vocabulary, comprehension measures and attainment in the previous version of the English language GCSE. Testing these relationships on attainment in the new specification of the exam would be a valuable contribution to the literature. It should be acknowledged though that no single vocabulary test will capture the complexity of the lexical knowledge that is needed for reading comprehension and while the test used in Paper 2 was imperfect, it is not necessarily the case that a single standardised receptive or expressive vocabulary test would have been superior as an index of the kind of vocabulary knowledge needed to comprehend the unseen texts found in this specific examination.

7.3.1.4 ARTs. The creation of two new and distinct ARTs for the survey in Paper 2 (Chapter 5), meant that there was no prior data or prior uses to support their validity. However, it would not have been possible to use previously created ARTs that tested

independent reading (reading for pleasure) as they are time and population-sensitive and become out of date quickly (Acheson et al., 2008). No previous ARTs were found that had targeted this age group specifically, especially in the UK, which was highly relevant given the focus of the study. Curriculum or classic authors have not been the focus of ARTs in the literature, which have generally been designed to measure reading experience with contemporary popular authors. The ARTs used in the current research did have high Cronbach alpha coefficients for the correct answers of the whole ART, the YA ART and the classic authors ART. There was also a strong positive correlation between scores on the YA ART and scores on the classic authors ART. Both ART scores also had significant relationships with the self-reported reading measures and the vocabulary test, which indicates that the two ARTs, created for this project, were capturing reading experience. Whilst these all indicate the potential strength of the ARTs created for this study and whilst new tests were essential for the research questions and participants, it is still important to acknowledge the limitation inherent in this designing of new measures.

7.3.1.5 COVID-19. Compared to other studies conducted at the same time, it was very fortunate that the data collection and analyses in this thesis were not too negatively impacted by COVID-19 and school closures. The one area where the data was affected was the English language GCSE attainment of the survey respondents who were part of the 2020 GCSE cohort. In 2020 the exams themselves were cancelled, due to COVID-19, and students received a grade based on a submission by their schools (Department for Education, 2020). There was not a uniform procedure for schools to follow in awarding this grade, so it was impossible to know how much was based on reading comprehension skill. There was also widespread grade inflation due to the removal of any norm referencing from the results (Ofqual, 2020). The statistical analysis for the survey results was therefore affected because it

had to account for cohort year as well as grade group. The grade inflation in the 2020 results could also explain the lack of responses from students with lower grades.

7.3.1.6 Focus on the English Language GCSE. Whilst the focus on a national exam has meant that the relevance of this study to practice is strong, it has also been a limitation because it has not considered all aspects of the English curriculum. The English literature GCSE content and curriculum has a substantial influence on the reading experience of adolescents at this stage of their education. It would have been interesting to have had time to consider the choices of texts for the literature GCSE alongside the unseen texts in the English Language exam. There are a range of literature texts from which schools can choose and the impact of these choices, on the vocabulary that students are exposed to through their reading experience, could be a very useful way of exploring the relationship between these two qualifications. The recent paper by Korochkina and Rastle (2024) has begun to look at this area of assessment by comparing a corpus created from the literature texts, from the specifications of two exam boards, with a corpus of popular books aimed at 13-16-year-old readers. Their analysis found that the GCSE literature books had a higher density of different words than the popular books and also contained a high numbers of words that did not feature in either the popular books or in spoken language. This suggested that the literature texts could potentially be very good sources for encounters with new vocabulary. However, further analysis of the differences between the literature texts showed that they were very unlike each other, that is they did not have a high overlap of vocabulary, meaning that reading one of the texts may not necessarily be good preparation, in terms of vocabulary knowledge, for reading another literature text.

Considering the results of the study by Korochkina and Rastle (2024) and any other future studies in this area, alongside the finding of this thesis, could bring greater understanding of both the qualifications, curriculum and adolescent reading experience.

7.3.2 Theoretical Limitations

This study used two main hypotheses about vocabulary and reading: the LQH (Perfetti & Hart, 2002) and the LLH (Nation, 2017). These hypotheses supported this study's focus on vocabulary knowledge and on reading experience. However, the data collected could not directly test these hypotheses but instead used them as a rationale. The decision to focus on vocabulary, as a component of reading skill, also had limitations. These theoretical limitations are discussed further in three sections below, that cover independent reading, curriculum reading, and vocabulary.

7.3.2.1 Lexical Legacies – Independent Reading. The significant relationship between the ART scores, the vocabulary scores and the self-reported reading measures, in Paper 2 (Chapter 5), did support the existing literature on the relationship between reading experience and lexical quality (O'Connor et al., 2019; Perfetti, 2007; Reichle & Perfetti, 2003). However, since each individual reader's lexical legacy is different, the data collected from the survey could only attempt to represent possible legacies or reading experiences rather than actual ones. The data collected for Paper 3 (Chapter 6), does begin, in an exploratory way, to represent actual lexical legacies by creating a collection of texts chosen for independent reading by a class of students. Whilst this was not representative of the detail of what one student might have read, it instead represented a range of different interests and choices across a group of students at one particular point in time. A larger project could use a similar methodology but with a greater range of students or across a greater time span in order to create more detail in descriptions of individual trajectories of reading. Future studies could also further explore the links between students' reading experience and their attainment in GCSE exams or their scores on other reading measures.

7.3.2.2 Lexical Legacies – Curriculum Reading. Similarly for the curriculum texts, not all students study all subjects, but the content for just one student's choices would have

skewed the content of the data in the curriculum corpus to their limited number of option subject choices. However, covering a large number of subjects, as was done in Paper 3 (Chapter 6), spreads the content beyond what any one student would experience. The data studied is therefore both wider (because it covers more subjects) and narrower (because it just covered one week) than any one reader's experience of curriculum reading. Using the LQH and the LLH and their complex and nuanced explanations of the building of vocabulary knowledge, makes straightforward, definitive and simple findings and conclusions difficult. However, the data collected for this study, especially for Paper 3 (Chapter 6) did provide interesting and useful exploratory findings.

7.3.2.3 Vocabulary. The focus on vocabulary was also a limitation. Although vocabulary is an important component in reading comprehension, it is not the only aspect that warrants study. This study's focus on vocabulary was necessary for practical reasons, but it is important to note that aspects such as collocations, syntactic and morphological structures, cohesive devices and anaphora all contribute to successful reading comprehension as well.

7.4 Implications

The implications of this study are presented under three main headings: implications for future research; implications for theory; and implications for practice. The implications for future research section considers how the findings of this study might inform future research in the area of adolescent reading comprehension. It covers: the choices regarding which types of reading texts are used in assessments; the impact these assessment texts can have on curriculum content; the importance of measuring reading experience in nuanced ways; the challenges for future reading research with adolescents; older literary texts as a genre of reading experience; the use of corpus linguistics techniques; and teachers' understanding of this area. The second implication section, implications for theory, examines the implications for the LLH and for research into adolescent reading more generally. The

final implications section, implications for practice, considers curriculum reading and independent reading separately and then vocabulary more generally.

7.4.1 Implications for Future Research

7.4.1.1 Analysis of Reading Texts Used in Assessments. This study has shown that the choices made about the kinds of the texts that are chosen for and used in assessments can have serious implications. In England, the aims, content and assessment objectives of exams are specified and regulated by the Government. The Secretary of State for Education, at the time of the introduction of the new GCSE specification that has been the focus of this study, was conservative Member of Parliament, Michael Gove. In Gove's policy steer letter to Ofqual (Gove, 2013) he outlined the reasons behind his desire to reform the GCSE qualifications. These were to: improve the reputation of the qualifications; align with other jurisdictions; increase stretch, challenge and ambition for pupils; and hold schools accountable for the performance of their pupils. The inclusion of literary texts and the use of older texts was intended to ensure that the reading texts in the exam were "high-quality" and "challenging" (Department for Education, 2013a, p. 4).

This study has suggested that the decision about the types of texts included in the exam could have an impact on the reading habits, curriculum content and the suitability of the qualification for what comes next in students' lives. Future research into assessment, as well as research into reading, that use passages of text as part of a comprehension measure, should likewise pay close attention to the age and genre of the text used, as both of these will impact the vocabulary and therefore the functional readability of the text. The selection of suitable texts for assessments involves making assumptions about particular genres of past reading experience, or even of life experience (as shown by the *hostel* example in the prologue to this thesis), that may or may not match that of the candidates taking the test. These aspects (assessment text age, assessment text genre, and implicit assumptions about candidates'

previous reading and life experiences) are important considerations in research into reading skill, assessment content and curricula. The implication for future research in this area is that there should be consideration of the reciprocal impact that these aspects have and, therefore, the impact of changing one (e.g. the age and genres of texts chosen for an exam) should also consider implications for the others (e.g. candidates previous reading experiences).

7.4.1.2 Impact of Assessments on Curriculum Choices. One of the key findings of this research study, presented in Paper 1 (Chapter 4), was that older literary texts provided the best match to the key vocabulary found in the English language GCSE. As this qualification has now had more time, since this study began, to become integrated into schools' schemes of work and curricula, it would be interesting to see research into the extent to which the new exam texts have changed reading in English classrooms, if at all. As this study has shown that it is possible for choices about assessment content to possibly influence choices about what is included in the curriculum, this has implications for future research into curricula with this age group.

7.4.1.3 Nuanced Measures of Reading Experience. This study has shown, through measuring different genres of reading experience separately, that it was reading experience with the genre of older literary fiction that had a closer relationship to high attainment in the English language GCSE. Future research into reading should continue to measure reading experience in a nuanced way that enables differentiation between types of reading experience. Possibilities for doing this include designing ARTs that discriminate between different genres of reading, in the same way as the one designed for Paper 2 (Chapter 5) in this study did, or designing measures that can account for other differences in reading experience.

Not all types of reading lend themselves to measurement using an ART. It would be difficult to build an ART of non-fiction authors as many popular non-fiction authors are famous or well known outside of their role as the author of their book or books. For example,

Dave Grohl, who was the author of one of the books chosen for independent reading in Paper 3 (Chapter 6), would probably be recognised first as a musician rather than as a writer of an autobiography. ARTs work as measures of reading experience on the premise that recognition of an author's name indicates reading experience with that author (Cunningham & Stanovich, 1997), recognising names for other reasons, as a famous musician for example, would therefore undermine the way in which the measure works. Future research should, therefore, look for ways in which this type of reading could be objectively measured.

Objective measures of other types of reading experience would also be very valuable, for example measuring reading experience online, especially as some digital reading practices (such as the reading of social media posts) have been shown to have a negative or weaker relationship to reading skill than fiction (Mar & Rain, 2015; McGeown et al., 2015; Pfost et al., 2013). Future research and the future design of reading measures should look for accurate, innovative and objective ways of capturing these types of reading experience in order to explore their relationship to reading skill. There could possibly be technological solutions to developing these measures, especially as these types of texts are read on a digital device. Vuorinen et al. (2023) used a specially developed e-reading web application to monitor reading behaviours, such as the time spent with the e-reader open, the navigation of the pages (linear/nonlinear), and the number of times that another browser was opened. This study was used to compare e-reading behaviours for the reading of just one story, with reader characteristics, but similar technology could be used to track a greater range of reading over a longer period of time. Freeman and Saunders (2015) also used e-reader data to track reading and compare reader behaviour via a university library e-book system. They were able to access data, for a whole year, on the e-books that had been borrowed, the time spent reading per session and the sequential order in which pages were viewed. The analysis focused on time spent reading, with some comparison between books from different subjects, but there

could be potential, with this type of e-reader data, to consider the vocabulary content of reading too. Whatever the solutions, reading experience online is now so ubiquitous that it should form part of future reading experience research.

7.4.1.5 Further Research into Adolescent Reading. There is far more reading research with younger readers and adult readers than there is with adolescents. This study has found that this was an age group that was difficult to reach and research. The demands of school, in the lead up to important exams, and the high levels of cooperation needed for data collection, meant that getting consent to participate, from both schools and individuals, was challenging. As this study has shown, however, research into the challenges of reading, specifically for this adolescent age group, is timely and interesting. Research into linguistic comprehension and vocabulary are important for this age group as, once decoding proficiency has been reached, they are more likely to be the cause of any reading problems that children experience when progressing through the later stages of primary and into their secondary education. Research that focuses on linguistic comprehension, vocabulary acquisition and reading experience, along with other components important to reading comprehension, particularly with adolescents, will therefore be very valuable in future studies.

7.4.1.6 Older Literary Texts. Following the identification of older literary texts as a good source of vocabulary for the exam texts in Paper 1 (Chapter 4) and the strong relationship between scores on the classic author ART and high attainment in the English language GCSE exam found in Paper 2 (Chapter 5), future research into reading experience with older literary texts is warranted. There could now be closer research into what it is about reading experience with older literary texts that is driving this relationship with higher attainment. It could be hypothesised that there is content in classic texts, maybe the level of vocabulary, use of figurative language, complexity of syntactic structures, that means that

reading experience with them is inherently superior for the acquisition of higher levels of reading skill. Or it could be hypothesised that reading experience with modern literary texts (with the possible exception of YA fiction) offers the same level of benefit. It should be noted that, as found by the recent study into GCSE literature texts (Korochkina & Rastle, 2024), there is great diversity within genres as well as between them. As outlined in the section above, nuanced measures of reading experience will allow for a differentiation both between different genres and within genres of reading experience. The use of standardised measures of reading skill could explore the relationship to reading experience with older literary texts further.

7.4.1.7 Use of Corpus Linguistics Techniques. This study has used corpus linguistics methods to create and study unique corpora and has also used existing reference corpora for comparisons and searches. Since the publication of Paper 1 (Chapter 4), in the first part of this study, at least one other study has also used corpus methods to research the language of texts in a GCSE exam (Korochkina & Rastle, 2024). Research that continues to use corpus linguistics to study and analyse the language of assessments and the content of adolescents' actual reading materials, whether independently chosen or read as part of their curriculum of study, could continue to critically evaluate and inform practice in this area. Building more corpora that reflect the diversity of texts that students in this age groups are actually reading, as was done for Paper 3 (Chapter 6) in this study, rather than relying on using texts that are aimed at this age group, would also provide useful and informative data about the reading experience of adolescents.

The other implication for future corpus linguistics research in this area is in the collection of texts that represented curriculum reading. Whilst a few decades ago, textbooks would have been the dominant reading material students experienced in classrooms, now reading materials are very different. School spending on books has been shown to be

decreasing in the 21st century while school spending on technology has increased (Coxhead et al., 2010). The current study found, in the data collected for the curriculum materials corpus in Paper 3 (Chapter 6), a range of online formats being used across subjects. Whilst only one week of curriculum materials was collected, the data from this week gives an indication of the kinds of reading formats used in classes: slides, worksheets, online or scanned pdfs, online quizzes and word documents, all featured. This has implications for future research in this area because the materials that students read in class, as part of the curriculum, could be becoming less standardised. There are also multiple platforms and online resources that students can access for studying outside of school, adding to the variety and diversity in curriculum reading experiences. Future research into the nature of these online resources could explore how many are commercially produced and how many are produced by individual teachers. It would also be of interest to explore the quality of the e-resources used, for example using measures such as vocabulary difficulty and syntactic complexity.

It was also noted, when uploading the data, that the formats of the curriculum materials meant that there was very little continuous prose content, something that warrants future research as this could have implications for students' experiences with words in context. This could be particularly relevant for the learning of vocabulary from meaningful contexts (Nation, 2017), as continuous prose could provide more contextual information than, for example, bullet points on a slide. There are also practical and technical implications for the collection, formatting and uploading of the materials into a corpus. The variety of formats and the structure of the materials, that included images and separated text boxes, made converting the text into a suitable format to upload more challenging, something that will need to be considered when planning future research projects of this nature.

7.4.1.8 Teachers' Understanding of Reading Comprehension and Assessment.

This study was motivated, from the outset, by the needs of practitioners and students. Further

studies that focus on real assessments, actual student reading experiences and potential application to real classrooms would be valuable. There is very little research that focuses specifically on the language content in and the importance of reading skills for GCSE exams. Spencer et al. (2017), Paper 1 (Chapter 4) of this study, and Korochkina and Rastle (2024) are recent exceptions. Research that is grounded in real classroom concerns and on real assessments can be immediately relevant to teachers and to classroom practice. Recommendations for changes or improvements to learning activities or learning behaviours, that come from research grounded in the context of real classrooms, could potentially be much more accessible to practitioners and could therefore have a more immediate impact on students and their learning and progress. More research in this vein, that is directly concerned with actual assessments, classroom practices or learning activities and behaviours, could therefore be beneficial.

7.4.2 Implications for Theory

This section, on implications for theory, will first examine some of the results from this study about different genres of reading experience and the implications for the LLH. Then the implications for theory of the new primary data collected from adolescents' actual reading materials will be covered.

7.4.2.1 The LLH. The first implication for theory is related to the LLH (Nation, 2017), which alongside the LQH (Perfetti & Hart, 2002), has provided the theoretical underpinning to the entirety of this study. This study was not a test of these two hypotheses, but the data speaks to them and is consistent with them both. In Paper 1 (Chapter 4) the LLH informed the decision to identify typical vocabulary from the exam texts and then search the reference corpus to identify the genres and registers in which this vocabulary was most likely to be found. Identifying the types of texts in which the vocabulary was most likely to be found could then suggest what prior reading experience was most advantageous, in terms of

providing encounters with this exam vocabulary. Having prior reading experience with the types of words that would be found in the exam would be essential, according to the LQH and the LLH, because efficient processing of words depends on readers having high quality representations of them and representations acquire high quality through repeated diverse reading experiences with those words.

Paper 1 (Chapter 4) demonstrated that not all reading experiences are as likely to provide repeated encounters with certain types of words. The vocabulary that was typical of the exam texts, the keywords, were found to be low in frequency, sometimes archaic, and were concerned with people, social gatherings and dramatic events. Low frequency words can be particularly challenging to comprehend, because it is harder to encounter them often enough to build high quality lexical representations of them (Perfetti & Hart, 2002). The best place to experience these words, to build a lexical legacy with them (Nation, 2017), was identified to be the imaginative texts in the BNC. Further searches in more specific genres identified prose fiction, poetry and drama as having the highest density of the keywords. This was in line with the literature that the reading of fiction has the closest relationship to reading skills (McGeown et al., 2015; Pfost et al., 2013; Torppa et al., 2020). Comparisons with other reference corpora identified the closest match, to the ETC, to be a corpus of older literary fiction (Project Gutenberg). Paper 1 (Chapter 4) was therefore able to demonstrate that the reading experience that was most likely to have prepared students for the comprehension demands of the vocabulary in the English language GCSE exam, included imaginative texts and specifically older literary fiction texts.

Paper 2 (Chapter 5) used an ART as a measure of reading experience. Following the methodology in previous research, that had measured different genres of reading experience (Mar & Rain, 2015; Martin-Chang et al., 2020; McGeown et al., 2015; Pfost et al., 2013), the ART was split into two distinct lists of authors and gave two scores: one for reading

experience with classic authors; and one for reading experience with YA authors. The significant correlation between scores on the classic ART, which represented reading experience with older literary fiction, and being in the high-grade group for English language GCSE attainment, can be understood, through the LLH, as showing that a specific type of prior reading experience, or a specific lexical legacy, is linked with high attainment in the assessment.

Paper 3 (Chapter 6) collected actual reading materials from adolescents' independent reading, from a homework task, and from adolescents' curriculum reading, from an online platform used by the participant school to share lesson materials. The linguistic analysis of these texts was informed by the LLH, as this would position any reading as providing a lexical legacy that would build reading expertise through increases in lexical quality (Nation, 2017). Using the LLH in this way, as the theoretical basis for the creation of two unique corpora and as the theoretical basis for studying the linguistic content of these corpora, develops and expands the uses of this hypothesis. Paper 3 (Chapter 6) did find, as would be expected by the literature (Biber et al., 1999), that there was a difference in the linguistic content between the two corpora created for the study. Students' independent reading was mainly fiction and therefore contained a significantly higher proportion of adverbs than the curriculum materials. The curriculum reading materials matched the linguistic content of the two non-fiction registers analysed by Biber et al. (1999), newspaper language and academic prose, by having significantly higher occurrences of nouns than the students' independent reading texts. This analysis demonstrates that lexical legacies are going to be different depending on the types of genres that form reading experiences. For example, as well as the different frequencies of adverbs and nouns in the different genres, exploratory analysis of the most frequent parts of speech in the two corpora identified more concrete vocabulary in the independent reading texts and more abstract vocabulary in the curriculum texts,

demonstrating another difference in the linguistic contents of the reading experiences provided by different genres.

The theoretical implication, for future research with the LLH, is therefore that the comprehension of different genres of reading will require different lexical legacies. Reading skill attainment could be impacted by the different genres of reading that might form different readers' lexical legacies. Gaining repeated experiences of words in diverse contexts, needed for reading skill proficiency, could therefore depend on the types of genres of texts that are read and the types of words that they contain (e.g. adverbs or nouns or concrete or abstract words), as well as other factors not explored in this thesis.

The LLH, in its clear theoretical description of the relationship between reading experience, vocabulary knowledge and reading skill, has given the analysis in this study its theoretical structure. Through the LLH, the vocabulary challenge of the main subject of this study, the English language GCSE, was linked to reading experience and then reading experience, in turn, linked to attainment. This successful application of theory to a real educational context could provide a useful model for future research.

7.4.2.2 Adolescent Reading. This study has collected new primary data on adolescent reading. These were: 1) the creation of a corpus from the English language GCSE exam texts, analysed in Paper 1 (Chapter 4); 2) adolescent reading habits data collected in Paper 2 (Chapter 5); 3) reading experience data, as measured by the ART, and its relationship to attainment in the English language GCSE as analysed in Paper 2 (Chapter 5); 4) actual adolescent independent reading choices collected in Paper 3 (Chapter 6); 5) the creation of a corpus of actual adolescent reading in Paper 3 (Chapter 6); and 6) the creation of a corpus of curriculum materials collected for Paper 3 (Chapter 6). In addition to the analysis of this data within this study, this new primary data, which will be made available after the completion of

this PhD study, could also potentially contribute to further research and theory in adolescent reading.

The findings in this study will also potentially have implications for theories that apply to adolescent reading. Paper 1 (Chapter 4) found that the specified texts for the new exam appeared to have impacted the type of vocabulary and therefore the type of reading experience that will prepare students for it. This suggests that it would be useful if the content of assessments is the focus of future research because it can potentially impact curriculum choices. One of the concerns raised by the first paper in this study was that the match to older literary fiction could mean that the traditional literary canon is promoted in reading choices. This is a concern because due to the sociocultural and ideological nature of its construction (Ervin, 2022; Löffler, 2017), that prioritised traditionally privileged groups, this could have a negative impact on the diversity and representative nature of curriculum materials and the independent reading choices of adolescents. This is an important implication for theory because it explicitly sites reading and reading texts in a social, political and ideological context. It reminds us of the real-world implications of assessments, curricula, and reading and also the real-world implication of research in these areas.

7.4.3 Implication for Practice

This study came from, and was motivated by, a problem observed from practice. The new English GCSE examination, with a reading component that is now worth 50% of the mark, and the change to the content of the exam texts, so that they now have to include literature or literary non-fiction and be from all three of the 19th, 20th and 21st centuries, changed the demands on students' reading skills. The impetus behind this project was, therefore, to find out more about this new reading challenge and how students might be better prepared for it. The implications for practice therefore focus on reading within the

curriculum, independent reading, and vocabulary, as they apply to students, teachers and school systems.

7.4.3.1 Curriculum Reading. The finding that the keywords from the new exam texts were high challenge, being typically low in frequency and predominantly found in older literary texts, has implications for students' preparation for the assessment. If classroom learning activities, planned to prepare students for the reading challenges of the exam, focus more on the older literary texts that have been shown to be a good match for the exam texts, this could have an impact on diversity and representation in the curriculum. The specification by Ofqual (2013) that students should read "literature and other high-quality writing" (Ofqual, 2013, p. 4) side-steps the sociocultural and ideological nature of decisions about what is considered to be literature and what types of writing are deemed high-quality (Ervin, 2022; Löffler, 2017). Traditionally privileged and powerful demographic groups are more likely to have been published and are therefore more likely to have a stronger reputation and have their texts disseminated in curriculum materials. At a time when diversity and representation is recognised as key for reasons of social justice and equality (Centre for Literacy in Primary Education, 2023; Dyches, 2018; Ervin, 2022; L. Johnson, 2018; Picton & Clark, 2022), the choices made about the texts that will feature in an exam, that is taken by the vast majority of 16-year-olds in England, seem retrograde in this respect.

The prioritising of literature and extended literary non-fiction (Ofqual, 2015) narrows the variety of texts which students will experience in their preparations for this qualification. Different registers of texts have very different grammatical and linguistic content (Biber et al., 1999). For example, as was shown in the difference between the IRC and CRC in Paper 3 (Chapter 6), the corpus that represented students' independent reading had a significantly higher frequency of adverbs and a significantly lower frequency of nouns, compared to the curriculum materials corpus. One of the stated aims of the new English language GCSE

qualification is that students should be able to read a wide range of texts (Ofqual, 2015), but the specification that exam texts be literature and literary non-fiction, effectively excludes texts that are informational or academic (i.e. about the curriculum), and more modern modes like online newsfeeds that are considered by the specification documentation, in a pejorative way, to be "transient" (Ofqual, 2015, p. 4). This qualification is used, by employers and by post-16 education providers, as an indication of the literacy capabilities of students. However, the narrowed range of texts means that, if lessons are purposefully targeted at providing experience with the types of texts and vocabulary that will feature in the exams, students are not gaining as much experience with or being tested on texts that have the different linguistic features, such as more abstract nouns.

Informational, academic, instructional, commercial and online texts are all surely valid texts to read and are going to feature in life post-16. In curriculum texts, represented by the CRC corpus in Paper 3 (Chapter 6), there were, in the most frequent occurrences of words, more abstract and technical words, when compared to the IRC which had parts of speech that focused more on the concrete and every day. It seems likely that the curriculum texts that students will encounter in their post-16 studies will have similar abstract and technical vocabulary, but currently the exam that is taken to indicate literacy ability and suitability for further academic study is not likely to include testing of proficiency with this type of vocabulary. A greater range of registers of texts in the examination would promote the greater use of a wide range of texts within the English curriculum. This wider range would make it much more likely that students will experience and be tested on a wider range of linguistic features and therefore would, arguably, be better prepared for the reading demands of their lives post-16.

It is understandable perhaps, for teachers and schools to yield to pressure to deliver a curriculum that they think will best prepare their students for high stakes end tests. In

England, as in many other geographical contexts, exam scores are used to judge teachers and schools, as well as the students themselves. However, cramming or excessive exam practice is not necessarily in the best interests of the students or will even provide the best learning activities (Sacks, 2000; Volante, 2004). For example, teachers might feel that their students need to improve their knowledge and understanding of archaic words. However, focusing lesson time on rote learning the meanings of archaic words is likely to be inefficient, as Nagy (1987) showed that it is learning words from context, particularly reading, that allows for sufficient vocabulary growth.

The aims behind the reform of the GCSEs, started a little over a decade ago in 2013, were to: 1) have qualifications that matched other high performing countries (although these countries were not specified); 2) improve respect for the qualifications; 3) evidence pupils' achievements; 4) provide a foundation for further and higher study; 5) hold schools accountable; 6) have more challenging content; and 7) more rigorous assessment structures (even if these jeopardised reliability) (Gove, 2013). The drive for what was seen to be more rigorous assessment, meant that there were more and longer exams at the end of the courses and that these exams had fewer short questions and answers and more focus on extended written answers (Gove, 2013). What did not seem to have been covered in the scope of the policy steer, by the Government at the time, was the impact of the new style of qualifications on the curriculum, apart from to be generally more "stretching" (Gove, 2013, p. 2). Future reviews of the qualification and curriculum landscape should, therefore, carefully consider the impact that test content and style will have on the curriculum.

Curriculum time is limited and precious. The design of any high stakes test should pay careful attention to the influence that it could bring to bear on curriculum content and activities. Any potential narrowing or unthinking prioritisation of sociocultural or ideological content, that has been traditionally privileged, needs to be carefully thought through and

assessed. Teachers and students should not be left trying to square a circle – where the content that would help attainment in an exam is perhaps contrary both to the stated aim of the curriculum and also to what might best prepare students for their lives after the qualification.

7.4.3.2 Independent Reading. There are of course many benefits to spending time independently reading (reading for pleasure). During primary education the focus on the promotion of reading is often on creating a love of reading and therefore promoting autonomy, choice and enjoyment (Casey, 2010; Department for Education, 2023a; Gambrell et al., 1996; Moss & McDonald, 2004) and the National Literacy Trust survey has found that children, including adolescents, are motivated to read by curiosity, for mindful reasons (e.g. to feel relaxed and happy), and also for social reasons (e.g. to feel connected to the world and to others). Whilst acknowledging all these wider benefits, the focus here is on the implications for preparation and attainment in a key exam. The finding in Paper 2 (Chapter 5) of this study, that there was a significant relationship between the classic author ART score and attaining a high grade in the English language GSCE exam, demonstrates that what is chosen to be read, the content and genre of reading experience, might matter. Reading for pleasure, already a somewhat problematic concept (Cremin et al., 2022), is not going to be enough if what a reader finds pleasurable to read is not providing enough encounters with words that will enable them to acquire adequate levels of reading skill. If independent choice, guided by what readers find pleasurable to read, cannot necessarily be relied upon to create an adequate lexical legacy for crucial exams then it might be that students' independent reading would benefit from being guided and specified to ensure that it does include reading experience adequate to create the needed legacy. As seen above, this is a change from the message currently promoted in schools.

There is also the issue that not all students are reading independently. The homework reading task, that was used to collect data on students' choices for independent reading for Paper 3 (Chapter 6), was not completed by two students from the class and it is also worth noting that the monitoring of the completion of the homework task depended entirely on the accurate self-report of the students. This supports the National Literacy Trust finding that the numbers of respondents to their survey saying that they enjoy reading and regularly read from choice is falling (Clark et al., 2023). Some of the students, for the independent reading task, said they had chosen to read a newspaper article, which while it met the requirements of the task, when viewed in the light of the findings from Paper 1 (Chapter 4), becomes a cause for concern, in so far as preparation for the English language GCSE exam is concerned. This is because newspaper writing is not a good match for the vocabulary in the exam texts. The tension, outlined above, between promoting choice and giving students agency in their reading and trying to match the reading experience that is the best preparation for the high stakes exam that they will face, is evident in this data.

Whilst fiction continues to dominate as the most popular choice for independent reading, both in the data from the survey in Paper 2 (Chapter 5) and the data from the homework task in Paper 3 (Chapter 6) in this study, as well as in the literature, non-fiction is often chosen, read and enjoyed too (Clark et al., 2023). The question then becomes whether it is better to encourage a student to continue non-fiction reading that they are enjoying, or to focus on promoting fiction reading because it is has been found to be more beneficial to measures of reading skill (McGeown et al., 2015; Torppa et al., 2020). With children's and adolescents' low levels of participation in reading being described as a crisis (Clark et al., 2023), is non-fiction reading better than none?

Questions of diversity and representation also feature in the literature about children's reading choices. Respondents to the National Literacy Trust survey said that they felt that

reading about people who were different to them was important, as well as saying they had increased confidence if they read about characters who were like them (Picton & Clark, 2022). Due to historic biases, older literary texts, that are the best match for the vocabulary in the English language GSCE, are not likely to be diverse or representative. Even today, according to the Centre for Literacy in Primary Education, more needs to be done to ensure that children's books are representative (Centre for Literacy in Primary Education, 2023). The range of books chosen by the small sample in Paper 3 (Chapter 6), suggests that mid-adolescents (ages 14-15), whilst still choosing what might be considered typical YA reading (e.g. Harry Potter books and *Twilight*), were also making choices that were outside typical teenage reading (e.g. Grohl's autobiography *The Storyteller* and a crime novel *Body Language).* This variety of choices is not always captured by the literature, where data might be collected from reading schemes that can only include a finite number of books and so inevitably focus on the most popular genre, YA fiction (Topping et al., 2023). Attention to these questions of diversity and representation, as well as attention to the actual choices being made by adolescents in what they are reading, should continue.

7.4.3.3 Vocabulary. The CRC in Paper 3 (Chapter 6) revealed the challenging nature of the vocabulary in the curriculum reading texts, which students in this age group are faced with in their lessons every day in school. Whilst the exploratory analysis showed that the IRC contained vocabulary about people, places and everyday occurrences, much like the ETC in Paper 1 (Chapter 4) contained characters and social events, the CRC instead had vocabulary that was more conceptual and abstract. If older literary fiction becomes the reading experience that is prioritised by schools or teachers who are predominantly concerned about attainment in the English language GCSE, then students might be missing out on building reading experience with the kinds of words that feature in most curriculum (academic) texts. It seems likely that this different set of vocabulary will also appear in post-16 curriculums

and, as indicated by the description of the linguistic content of academic texts (Biber et al., 1999), in higher education too. The focus in the exam on literature and literary non-fiction could potentially leave reading experience with these types of texts and vocabulary lacking.

7.5 Conclusion

When a 16-year-old student in England, at the end of their compulsory schooling, sits their English language GCSE exam, they have to comprehend three previously unseen texts. These are required to include literature and literary non-fiction and to be drawn from across all three of the 19th, 20th and 21st centuries. In order to have the required reading skill to comprehend these texts in their exams, students will be drawing on their prior reading experience. They will have, hopefully, mastered the decoding element of reading very early on in their reading journey, when at primary school. Building the linguistic comprehension of the words before them, will have required them to have experienced these words multiple times in different contexts (Nation, 2017; Perfetti & Hart, 2002). This will have required reading experience, not just experience of spoken language, as written texts are needed for many low frequency words that are not often used in our speech (Braze et al., 2007; Landauer & Dumais, 1997). If a student's linguistic and reading experience has been poor then they will struggle to access, quickly and fluently, the meaning of the vocabulary in the exam texts.

This study has shown the current English language GCSE exam texts present a high vocabulary challenge. The words that typify the exam texts are low in frequency and are most likely to have been experienced before in older fiction texts. Not all students are keen readers, with many saying that they do not read independently or for pleasure (Clark et al., 2023). Those students who do choose to read independently tend to choose modern fiction and YA authors, but these may not provide the same experience with words that the older literary texts, that were identified as the best match to the exam texts, do. Reading in lessons, or curriculum reading, for students who do not read outside of school independently, could then

become for them an important source of experience with words. Study of the linguistic content of curriculum documents, however, showed that they were not a good match for the English language GCSE exam texts.

The results of this study might suggest that the remedy, for students struggling with their reading comprehension skill in this important exam, should be reading more older, literary fiction. This recommendation, however, may fall into the trap of describing a deficit narrative, that criticises students lack of the 'right' kind of reading experience, without considering the power structures at play behind what is considered 'right'. A focus on older literary texts, a genre that is already covered by a separate English literature GCSE, risks perpetuating the privilege of groups that have traditionally benefited from being socially and ideologically powerful, and whose writings have become to be considered the literary canon. Instead, we should be challenging the deficit in the exam texts. They are currently drawn from a narrow pool that does nothing to improve or challenge the historical lack of representation and diversity in the curriculum (Centre for Literacy in Primary Education, 2023; Dyches, 2018; V. Elliott et al., 2021; Ervin, 2022; L. Johnson, 2018; Picton & Clark, 2022). The focus on literature and the literary also excludes the linguistic content of academic and non-fiction texts (Biber et al., 1999), surely important for students' future study, work and civic engagement.

This study was born out of a desire to help struggling students improve their reading skill in this gate-keeping qualification. However, the results here have found that it is the exam itself that is struggling. It's struggling, due to its narrow and dated content, to adequately measure what it purports to, the literacy skills needed by students post-16. It is also struggling to foster, in fact it seems to actively work against, the diverse and representative curriculum that all students deserve. This thesis will hopefully draw attention

to this issue and could bring about some change, especially with the recent decision by the new Government to review the curriculum and assessments.

Word count: 72,852

Reference List

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*. https://doi.org/10.3758/BRM.40.1.278

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science*. https://doi.org/10.1111/j.1467-9280.2006.01787.x

Alfassi, M. (1998). Reading for meaning: The efficacy of reciprocal teaching in fostering reading comprehension in high school students in remedial reading classes. *American Educational Research Journal*. https://doi.org/10.3102/00028312035002309

Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple Indicators of Children's Reading Habits and Attitudes: Construct Validity and Cognitive Correlates. In *Journal of Educational Psychology* (Vol. 84, Issue 4, pp. 489–503). https://doi.org/10.1037/0022-0663.84.4.489

Anderson, R. C., & Freebody, P. (1981). *Comprehension and Teaching Research Reviews* (J. T. Guthrie, Ed.). International Reading Assn.

Anderson, R. C., & Freebody, P. (1982). *Reading Comprehension and the Assessment and Acquisition of Word Knowledge. Technical Report No. 249.*

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in Reading and How Children Spend Their Time Outside of School. *Reading Research Quarterly*, *23*(3), 285–303. https://doi.org/10.1598/rrq.23.3.2

Anthony, L. (2013). A critical look at software tools in corpus linguistics. *Linguistic Research*, *30*(2), 141–161.

Apthorp, H., Randel, B., Cherasaro, T., Clark, T., McKeown, M., & Beck, I. (2012). Effects of a Supplemental Vocabulary Program on Word Knowledge and Passage

Comprehension. *Journal of Research on Educational Effectiveness*. https://doi.org/10.1080/19345747.2012.660240

Aston, G. (2001). Text categories and corpus users: A response to David Lee. *Language Learning and Technology*, *5*(3), 73–76.

Baird, J.-A., Caro, D., Elliott, V., El Masri, Y., Ingram, J., Issacs, T., Pinot de Moira, A., Randhawa, A., Stobart, G., Meadows, M., Morin, C., & Taylor, R. (2019). *Examination Reform: Impact of Linear and Modular Examinations at GCSE* (Issue April).

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Balota, D. A., & Spieler, D. H. (1998). The Utility of Item-Level Analyses in Model Evaluation: A Reply to Seidenberg and Plaut. *Psychological Science*, *9*(3). https://doi.org/10.1111/1467-9280.00047

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General*, *128*(1). https://doi.org/10.1037/0096-3445.128.1.32

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The english lexicon project. In *Behavior Research Methods* (Vol. 39, Issue 3). https://doi.org/10.3758/BF03193014

Batterink, L., & Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuroscience*, *23*(11), 3181–3196. https://doi.org/10.1162/jocn_a_00013

Baugh, J. (2017). Meaning-Less Differences: Exposing Fallacies and Flaws in “The Word Gap” Hypothesis That Conceal a Dangerous “Language Trap” for Low-Income

American Families and Their Children. *International Multilingual Research Journal*, *11*(1), 39–51. https://doi.org/10.1080/19313152.2016.1258189

Baumann, J. F., Edwards, E. C., Font, G., Tereshinski, C. A., Kame'enui, E. J., & Olejnik, S. (2002). Teaching morphemic and contextual analysis to fifth-grade students. *Reading Research Quarterly*. https://doi.org/10.1598/rrq.37.2.3

Bawarshi, A. S., & Reiff, M. J. (2010). *Genre: An Introduction to History, Theory, Research, and Pedagogy*. Parlor Press.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing Words to Life*. The Guildford Press.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*. https://doi.org/10.1037/0022-0663.74.4.506

Beech, J. R. (2002). Individual Differences in Mature Readers in Reading, Spelling, and Grapheme-phoneme Conversion. *Current Psychology*, *21*(2), 121–132.

Berkeley, S., Mastropieri, M. A., & Scruggs, T. E. (2011). Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities*, *44*(1), 18–32. https://doi.org/10.1177/0022219410371677

Bethlehem, J. (2016). Solving the Nonresponse Problem With Sample Matching? *Social Science Computer Review*, *34*(1), 59–77. https://doi.org/10.1177/0894439315573926

Biber, D. (1990). Methodological Issues Regarding Corpus-based Analyses of Linguistic Variation. *Literary and Linguistic Computing*, *5*(4), 257–269. https://doi.org/10.1093/llc/5.4.257

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Pearson Education.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*. https://doi.org/10.1037/0022-0663.93.3.498

Bisson, N., Tobin, S., & Grondin, S. (2012). Prospective and retrospective time estimates of children: A comparison based on ecological tasks. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0033049

Blair, E., & Burton, S. (1987). Cognitive Processes Used by Survey Respondents to Answer Behavioral Frequency Questions. *Journal of Consumer Research*, *14*(2), 280. https://doi.org/10.1086/209112

Brabham, E. G., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology*. https://doi.org/10.1037//0022-0663.94.3.465

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities*, *40*(3), 226–243. https://doi.org/10.1177/00222194070400030401

Brown, C. P. (2015). Taking and teaching the test are not the same: A case study of first-year teachers experiences in high-stakes contexts. *Teachers and Teaching: Theory and Practice*, *21*(8), 1026–1044. https://doi.org/10.1080/13540602.2015.1005870

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Current Directions in Psychological Science*, *27*(1), 45–50. https://doi.org/10.1177/0963721417727521

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Cameron-Faulkner, T., & Noble, C. (2013). A comparison of book text and Child Directed Speech. *First Language*, *33*(3), 268–279. https://doi.org/10.1177/0142723713487613

Carnine, D., Kameenui, E. J., & Coyle, G. (1984). Utilization of Contextual Information in Determining the Meaning of Unfamiliar Words. *Reading Research Quarterly*. https://doi.org/10.2307/747362

Carver, R. P. (1976). Word length, prose difficulty, and reading rate. *Journal of Literacy Research*, *8*(2). https://doi.org/10.1080/10862967609547176

Casey, R. (2010). *A Phenomenological study of high school students' perceptions of English*.

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

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language Deficits in Poor Comprehenders: A Case for the Simple View of Reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278–293. https://doi.org/10.1044/1092-4388(2006/023)

Centre for Literacy in Primary Education. (2023). *REFLECTING REALITIES Survey of Ethnic Representation within UK Children's Literature 2022*. www.clpe.org.uk

Chandler, D. (1997). *An Introduction to Genre Theory*. http://visual-memory.co.uk/daniel//Documents/intgenre/chandler_genre_theory.pdf

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory and Cognition*. https://doi.org/10.3758/BF03211582

Chesney, T., & Penny, K. (2013). The impact of repeated lying on survey results. *SAGE Open*. https://doi.org/10.1177/2158244012472345

Choi, W., Lowder, M. W., Ferreira, F., & Henderson, J. M. (2015). Individual differences in the perceptual span during reading: Evidence from the moving window technique. *Attention, Perception, and Psychophysics*, *77*(7), 2463–2475. https://doi.org/10.3758/s13414-015-0942-1

Cipielewski, J., & Stanovich, K. E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology*, *54*(1), 74–89. https://doi.org/10.1016/0022-0965(92)90018-2

Clark, C., Picton, I., & Galway, M. (2023). *Children and young people's reading in 2023*.

Clark, C., & Rumbold, K. (2006). *Reading for Pleasure: A research overview*.

Clarke, P. J., Paul, S. A. S., Smith, G., Snowling, M. J., & Hulme, C. (2017). Reading Intervention for Poor Readers at the Transition to Secondary School. *Scientific Studies of Reading*, *21*(5), 408–427. https://doi.org/10.1080/10888438.2017.1318393

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating Children's Reading-Comprehension Difficulties: A Randomized Controlled Trial. *Psychological Science*, *21*(8), 1106–1116. https://doi.org/10.1177/0956797610375449

Coltheart, M. (1981). The mrc psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4). https://doi.org/10.1080/14640748108400805

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. In *Behavior Research Methods, Instruments, and Computers* (Vol. 36, Issue 3). https://doi.org/10.3758/BF03195585

Covington, M. A., & McFall, J. (2008). *MATTR 2.0.3018.28419 2008/04/07* (2.0). CASPR Project.

Covington, M. A., & McFall, J. D. (2010). Cutting the gordian knot: The moving-average type-token ratio (MATTR). In *Journal of Quantitative Linguistics* (Vol. 17, Issue 2, pp. 94–100). https://doi.org/10.1080/09296171003643098

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, *34*(2), 213. https://doi.org/10.2307/3587951

Coxhead, A. (2011). The academic word list 10 years on: Research and teaching implications. *TESOL Quarterly*, *45*(2), 355–362. https://doi.org/10.5054/tq.2011.254528

Coxhead, A., & Boutorwick, T. J. (2018). Longitudinal Vocabulary Development in an EMI International School Context: Learners and Texts in EAL, Maths, and Science. *TESOL Quarterly*, *52*(3), 588–610. https://doi.org/10.1002/tesq.450

Coxhead, A., Stevens, L., & Tinkle, J. (2010). Why might secondary science textbooks be difficult to read. *New Zealand Studies in Applied Linguistics*, *16*(2), 1.

Cremin, T., Hendry, H., Chamberlain, L., & Twiner, A. (2022). Reading for pleasure: Exploring the concept. In *the Mercers' Company*.

Cremin, T., & Scholes, L. (2024). Reading for pleasure: scrutinising the evidence base–benefits, tensions and recommendations. *Language and Education*. https://doi.org/10.1080/09500782.2024.2324948

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing Print Exposure and Orthographic Processing Skill in Children: A Quick Measure of Reading Experience. *Journal of Educational Psychology*, *82*(4), 733–740. https://doi.org/10.1037/0022-0663.82.4.733

Cunningham, Anne. E. (2005). Vocabulary growth through independent reading and reading aloud to children. In Kamil M. & Hiebert E (Eds.), *Teaching and learning new vocabulary: Bringing research to practice* (pp. 45–68). Erlbaum.

Cunningham, Anne. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945. https://doi.org/10.1037/0012-1649.33.6.934

Cushing, I. (2023). Word rich or word poor? Deficit discourses, raciolinguistic ideologies and the resurgence of the 'word gap' in England's education policy. *Critical Inquiry in Language Studies*, *20*(4), 305–331. https://doi.org/10.1080/15427587.2022.2102014

Davidse, N. J., de Jong, M. T., Bus, A. G., Huijbregts, S. C. J., & Swaab, H. (2011). Cognitive and environmental predictors of early literacy skills. *Reading and Writing*, *24*(4). https://doi.org/10.1007/s11145-010-9233-3

Dawson, N., Hsiao, Y., Wei, A., Tan, M., & Banerji, N. (2021). Features of lexical richness in children's books: Comparisons with child-directed speech. *Language Development Research*. https://doi.org/10.34842/5we1-yk94

De Naeghel, J., Van Keer, H., Vansteenkiste, M., & Rosseel, Y. (2012). The relation between elementary students' recreational and academic reading motivation, reading frequency, engagement, and comprehension: A self-determination theory perspective. *Journal of Educational Psychology*, *104*(4), 1006–1021. https://doi.org/10.1037/a0027800

Deignan, A., & Love, R. (2019). Using corpus methods to identify subject specific uses of polysemous words in English secondary school science materials. *Corpora*. https://eprints.whiterose.ac.uk/154115/

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. In *Educational Research Review* (Vol. 25, pp. 23–38). Elsevier Ltd. https://doi.org/10.1016/j.edurev.2018.09.003

Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., Fletcher, J. M., & Carter, S. (2015). Text-Processing Differences in Adolescent Adequate and Poor Comprehenders Reading Accessible and Challenging Narrative and Informational Text. *Reading Research Quarterly*, *50*(4), 393–416. https://doi.org/10.1002/rrq.105

Department for Education. (2012). *On Reading for Pleasure* (Issue May).

Department for Education. (2013a). English language: GCSE subject content and assessment objectives. In *English language GCSE subject content and assessment objectives*. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/254497/GCSE_English_language.pdf

Department for Education. (2013b). National curriculum in England key stages 1 to 4. In *11 September* (Issue September). https://www.gov.uk/government/publications/national-curriculum-in-england-framework-for-key-stages-1-to-4/the-national-curriculum-in-england-framework-for-key-stages-1-to-4

Department for Education. (2020, August 27). *Taking exams during the coronavirus (COVID-19) outbreak.* https://www.gov.uk/government/publications/coronavirus-covid-19-cancellation-of-gcses-as-and-a-levels-in-2020/coronavirus-covid-19-cancellation-of-gcses-as-and-a-levels-in-2020#awarding-grades-in-2020

Department for Education. (2023a). *The reading framework*.

Department for Education. (2023b). *The reading framework*.

Department for Education. (2023c, December 14). *Key stage 2 attainment*. National Statistics. https://explore-education-statistics.service.gov.uk/find-statistics/key-stage-2-attainment/2022-23

Duff, D., & Brydon, M. (2020). Estimates of individual differences in vocabulary size in English: how many words are needed to 'close the vocabulary gap'? *Journal of Research in Reading*, *43*(4), 454–481. https://doi.org/10.1111/1467-9817.12322

Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). *British Picture Vocabulary Scale* (Third Edit). GL Assessment Ltd.

Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test. In *American Guidance Service* (3rd ed.).

Dyches, J. (2018). Critical Canon Pedagogy: Applying Disciplinary Inquiry to Cultivate Canonical Critical Consciousness. *Harvard Educational Review*, *88*(4), 538–564.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*(1), 262–300. https://doi.org/10.3102/0034654308325998

Elliott, G., Joseph, H., & Powell, D. (2024). *Ameliorating the reading comprehension difficulties of adolescent poor comprehenders: A randomised control trial [Manuscript in Preparation]*.

Elliott, V., Nelson-Addy, L., Chantiluke, R., & Courtney, M. (2021). *Diversity in Literature in English Schools*.

Ervin, J. (2022). Critically Reading the Canon: Culturally Sustaining Approaches to a Prescribed Literature Curriculum. *Journal of Adolescent and Adult Literacy*, *65*(4), 321–329. https://doi.org/10.1002/jaal.1208

Evans, J. R., & Mathur, A. (2018). The value of online surveys: a look back and a look ahead. *Internet Research*, *28*(4), 854–887. https://doi.org/10.1108/IntR-03-2018-0089

Evison, J. (2015). What are the basics of analysing a corpus? In *The Routledge Handbook of Corpus Linguistics*. Routledge. https://doi.org/10.4324/9780203856949.ch10

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*(2), 234–248. https://doi.org/10.1111/desc.12019

Foster, C. (2024). Methodological pragmatism in educational research: from qualitative-quantitative to exploratory-confirmatory distinctions. *International Journal of Research and Method in Education*, *47*(1), 4–19. https://doi.org/10.1080/1743727X.2023.2210063

Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin, J. B. (2005). Dimensions affecting the assessment of reading comprehension. In *Children's Reading Comprehension and Assessment*. https://doi.org/10.4324/9781410612762

Freeman, R. S., & Saunders, E. S. (2015). E-book reading practices in different subject areas: An exploratory log analysis. In *Academic E-Books: Publishers, Librarians, and Users*.

Frick, A., Bächtiger, M. T., & Reips, U. D. (2001). Financial incentives, personal information and dropout in online studies. In *Dimensions of Internet Science* (pp. 209–219). https://www.researchgate.net/profile/Andrea_Frick/publication/290392761_Financial_incentives_personal_information_and_drop-out_in_online_studies/links/569e352d08ae950bd7a9479a.pdf%0Ahttp://deusto.academia.edu/UlfDietrichReips/Papers/106583/Financial_incen

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., & Allen, N. E. (2017). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *American Journal of Epidemiology*, *186*(9), 1026–1034. https://doi.org/10.1093/aje/kwx246

Gambrell, L. B., Palmer, B. M., Codling, R. M., & Mazzoni, S. A. (1996). Assessing Motivation to Read. *The Reading Teacher*, *49*(7), 518–533. https://doi.org/10.4324/9781315752815-3

García, O., & Otheguy, R. (2017). Interrogating the Language Gap of Young Bilingual and Bidialectal Students. *International Multilingual Research Journal*, *11*(1), 52–65. https://doi.org/10.1080/19313152.2016.1258190

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, *35*(3), 305–327. https://doi.org/10.1093/applin/amt015

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., & Yoon, H. J. (2018). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism*, *21*(3), 563–584. https://doi.org/10.1017/S1366728917000219

Gorelova, Y. N., Zalyaeva, E. O., & Sungatullina, D. D. (2015). Examination text structure as an indicator of examination task type and complexity (On the basis of GCSE English language unit 1 section A reading). *Mediterranean Journal of Social Sciences*, *6*(3), 414–420. https://doi.org/10.5901/mjss.2015.v6n3s2p414

Gough, P. B., Hoover, W., & Peterson, C. (1996). Some observations on a simple view of reading. In *Reading comprehension difficulties: Processes and Intervention*.

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

Gove, M. (2013). *Letter-from-michael-gove-reform-of-ks4-qualifications*. Department for Education.

Gray, S. H., Ehri, L. C., & Locke, J. L. (2018). Morpho-phonemic analysis boosts word reading for adult struggling readers. *Reading and Writing*, *31*(1), 75–98. https://doi.org/10.1007/s11145-017-9774-9

Green, C., & Lambert, J. (2018). Advancing disciplinary literacy through English for academic purposes: Discipline-specific wordlists, collocations and word families for eight secondary subjects. *Journal of English for Academic Purposes*, *35*, 105–115. https://doi.org/10.1016/j.jeap.2018.07.004

Green, C., & Lambert, J. (2019). Position vectors, homologous chromosomes and gamma rays: Promoting disciplinary literacy through Secondary Phrase Lists. *English for Specific Purposes*, *53*, 1–12. https://doi.org/10.1016/j.esp.2018.08.004

Harlaar, N., Deater-Deckard, K., Thompson, L. A., Dethorne, L. S., & Petrill, S. A. (2011). Associations Between Reading Achievement and Independent Reading in Early Elementary School: A Genetically Informative Cross-Lagged Study. *Child Development*, *82*(6), 2123–2137. https://doi.org/10.1111/j.1467-8624.2011.01658.x

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Hart, P. D., & Teeter, R. M. (2001). A measured response: Americans speak on education reform. *Education Testing Service*.

Hawkins, R. O., Musti-Rao, S., Hale, A. D., McGuire, S., & Hailley, J. (2010). Examining listening previewing as a classwide strategy to promote reading comprehension and vocabulary. *Psychology in the Schools*. https://doi.org/10.1002/pits.20513

Heffetz, O., & Reeves, D. B. (2019). Difficulty of Reaching Respondents and Nonresponse Bias: Evidence from Large Government Surveys. *Review of Economics and Statistics*, *101*(1), 176–191. https://doi.org/10.1162/rest_a_00748

Heiervang, E., & Goodman, R. (2011). Advantages and limitations of web-based surveys: Evidence from a child mental health survey. *Social Psychiatry and Psychiatric Epidemiology*. https://doi.org/10.1007/s00127-009-0171-9

Henderson, L., Snowling, M., & Clarke, P. (2013). Accessing, Integrating, and Inhibiting Word Meaning in Poor Comprehenders. *Scientific Studies of Reading*, *17*(3), 177–198. https://doi.org/10.1080/10888438.2011.652721

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? In *Behavioral and Brain Sciences* (Vol. 33, Issues 2–3, pp. 61–83). Cambridge University Press. https://doi.org/10.1017/S0140525X0999152X

Hoffman, P., Lambon Ralph, M. A., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, *45*(3), 718–730. https://doi.org/10.3758/s13428-012-0278-x

Hoffman, P., & Woollams, A. M. (2015). Opposing effects of semantic diversity in lexical and semantic relatedness decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 385–402. https://doi.org/10.1037/a0038995

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

Hsiao, Y., Dawson, N. J., Banerji, N., & Nation, K. (2022). The nature and frequency of relative clauses in the language children hear and the language children read: A developmental cross-corpus analysis of English complex grammar. *Journal of Child Language*. https://doi.org/10.1017/S0305000921000957

Hulme, R. C., Barsky, D., & Rodd, J. M. (2019). Incidental Learning and Long-Term Retention of New Word Meanings From Stories: The Effect of Number of Exposures. *Language Learning*, *69*(1), 18–43. https://doi.org/10.1111/lang.12313

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Hyland, K. (2017). English in the disciplines: Arguments for specificity. *ESP Today*, *5*(1), 5–23. https://doi.org/10.18485/esptoday.2017.5.1.1

Hyland, K., & Tse, P. (2007). Is There an "Academic Vocabulary"? In *Quarterly* (Vol. 41, Issue 2).

Iarossi, G. (2006). *Power of survey design : A user's guide for managing surveys, interpreting results, and influencing respondents*. World Bank Publications. https://doi.org/10.1596/978-0-8213-6392-8

Ingram, J., Stiff, J., Cadwallader, S., Lee, G., & Kayton, H. (2023). *PISA 2022: National Report for England December 2023*.

Isaacs, T. (2014). Curriculum and assessment reform gone wrong: The perfect storm of GCSE English. *Curriculum Journal*, *25*(1), 130–147. https://doi.org/10.1080/09585176.2013.876366

James, E., Currie, N. K., Tong, S. X., & Cain, K. (2020). The relations between morphological awareness and reading comprehension in beginner readers to young adolescents. *Journal of Research in Reading*, 1–21. https://doi.org/10.1111/1467-9817.12316

Jarvis, S. (2013). Capturing the Diversity in Lexical Diversity. *Language Learning*, *63*(SUPPL. 1), 87–106. https://doi.org/10.1111/j.1467-9922.2012.00739.x

Jennings, B., Powell, D., Jaworska, S., & Joseph, H. (2024). A Corpus Study of English Language Exam Texts: Vocabulary Difficulty and the Impact on Students' Wider Reading (or Should Students be Reading More Texts by Dead White Men?). *Journal of Adolescent and Adult Literacy*, *67*, 303–316. https://doi.org/10.1002/jaal.1331

Jennings, J. L., & Bearak, J. M. (2014). "Teaching to the Test" in the NCLB Era: How Test Predictability Affects Our Understanding of Student Performance. *Educational Researcher*, *43*(8), 381–389. https://doi.org/10.3102/0013189X14554449

Jitendra, A. K., Kay Hoppes, M., & Yan Ping Xin. (2000). Enhancing Main Idea Comprehension for Students with Learning Problems. *The Journal of Special Education*, *34*(3), 127–139. https://doi.org/10.1177/002246690003400302

Johns, B. T., Dye, M., & Jones, M. N. (2016). The influence of contextual diversity on word learning. *Psychonomic Bulletin and Review*, *23*(4), 1214–1220. https://doi.org/10.3758/s13423-015-0980-7

Johnson, E. J. (2015). Debunking the “language gap.” *Journal for Multicultural Education*, *9*(1), 42–50. https://doi.org/10.1108/JME-12-2014-0044

Johnson, L. L. (2018). *Where Do We Go from Here? Toward a Critical Race English Education*. *53*(2), 102–124. https://doi.org/10.2307/26802728

Joseph, H., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, *49*(16), 2078–2086. https://doi.org/10.1016/j.visres.2009.05.015

Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology*, *166*, 190–211. https://doi.org/10.1016/j.jecp.2017.08.010

Kameenui, E. J., Carnine, D. W., & Freschi, R. (1982). Effects of Text Construction and Instructional Procedures for Teaching Word Meanings on Comprehension and Recall. *Reading Research Quarterly*. https://doi.org/10.2307/747525

Kaminski, R. A. (2002). Dynamic Indicators of Basic Early Literacy Skills TM 6 th Edition. *Measurement*.

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

Khanna, M. M., & Cortese, M. J. (2021). How well imageability, concreteness, perceptual strength, and action strength predict recognition memory, lexical decision, and reading aloud performance. *Memory*, *29*(5), 622–636. https://doi.org/10.1080/09658211.2021.1924789

Kilgarriff, A. (2001). Comparing Corpora. *International Journal of Corpus Linguistics*, *6*(1), 97–133. https://doi.org/10.1075/ijcl.6.1.05kil

Kilgarriff, A. (2005). Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory*, *1*(2), 263–276. https://doi.org/10.1515/cllt.2005.1.2.263

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*. https://doi.org/10.1007/s40607-014-0009-9

Kintsch, W., Rawson, K. A., Nation, K., Perfetti, C. A., Landi, N., & Oakhill, J. (2005). *The Science of Reading* (M. J. Snowling & C. Hulme, Eds.). Blackwell Publishing Ltd.

Kirsch, I., Jong, J. De, Lafontaine, D., McQueen, J., Mendelovits, J., & Monseur, C. (2002). Reading for Change Performance and Engagement Across Countries Results From Pisa 2000 Organisation for Economic Co-Operation and Development Organisation for Economic Co-Operation and Development. *Oecd*, 262. https://doi.org/10.1787/9789264099289-en

Kılınç, H., & Fırat, M. (2017). Opinions of expert academicians on online data collection and voluntary participation in social sciences research. *Kuram ve Uygulamada Egitim Bilimleri*, *17*(5), 1461–1486. https://doi.org/10.12738/estp.2017.5.0261

Koester, A. (2015). Building small specialised corpora. In *The Routledge Handbook of Corpus Linguistics*. Routledge. https://doi.org/10.4324/9780203856949.ch6

Kong, Y., Seo, Y. S., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers and Education*, *123*, 138–149. https://doi.org/10.1016/j.compedu.2018.05.005

Korochkina, M., Marelli, M., Brysbaert, M., & Rastle, K. (2024). The Children and Young People's Books Lexicon (CYP-LEX): A large-scale lexical database of books read by children and young people in the United Kingdom [E-PUB AHEAD OF PRINT]. *Q J Exp Psychol (Hove)*. https://doi.org/https://doi.org/10.31234/osf.io/nha8t

Korochkina, M., & Rastle, K. (2024). *The vocabulary barrier in the General Certificate of Secondary Education (GCSE) in English Literature [Manuscript in Preparation]*. https://osf.io/hqvd3/

Kuhn, M. R., & Stahl, S. A. (1998). Teaching children to learn word meanings from context: A synthesis and some questions. *Journal of Literacy Research*, *30*(1), 119–138. https://doi.org/10.1080/10862969809547983

Kyle, K., Crossley, S. A., & Jarvis, S. (2021). Assessing the Validity of Lexical Diversity Indices Using Direct Judgements. *Language Assessment Quarterly*, *18*(2), 154–170. https://doi.org/10.1080/15434303.2020.1844205

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240. https://doi.org/10.1037/0033-295X.104.2.211

Lee, D. Y. W. (2001). Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning and Technology*, *5*(3), 37–72. https://doi.org/10.1163/9789004334236_021

Lee, D. Y. W. (2015). What corpora are available? In *The Routledge Handbook of Corpus Linguistics*. https://doi.org/10.4324/9780203856949.ch9

Lenhard, W., & Schneider, W. (2006). *ELFE 1–6: Ein Leseverständnistest für Erst- bis Sechstklässer [ELFE 1–6: A reading comprehension test for Grades 1 to 6]*. Hogrefe.

Leppänen, U., Aunola, K., & Nurmi, J. E. (2005). Beginning readers' reading performance and reading habits. In *Journal of Research in Reading* (Vol. 28, Issue 4, pp. 383–399). https://doi.org/10.1111/j.1467-9817.2005.00281.x

Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the Developmental Relationship Between Oral Language Skills and Reading Comprehension: It's Simple, But Complex. *Child Development*, *89*(5). https://doi.org/10.1111/cdev.12861

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of Academic Vocabulary Instruction for Linguistically Diverse Adolescents: Evidence From a Randomized Field Trial. *American Educational Research Journal*. https://doi.org/10.3102/0002831214532165

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. In *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X99001776

Löffler, P. (2017). *Reading the Canon: Literary History in the 21st Century*. Universitatsverlag Winter.

Lubliner, S., & Smetana, L. (2005). The effects of comprehensive vocabulary instruction on students' metacognitive word-learning skills and reading comprehension. *Journal of Literacy Research*. https://doi.org/10.1207/s15548430jlr3702_3

MacGinitie, W. H., & MacGinitie, R. K. (1992). *Gates-MacGinitie reading tests* (2nd Canadi). Nelson.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2000). *Gates–MacGinitie Reading Tests* (4th ed.). Riverside.

Mano, Q. R., & Guerin, J. M. (2018). Direct and indirect effects of print exposure on silent reading fluency. *Reading and Writing*, *31*(2), 483–502. https://doi.org/10.1007/s11145-017-9794-5

Mar, R. A., & Rain, M. (2015). Narrative Fiction and Expository Nonfiction Differentially Predict Verbal Ability. *Scientific Studies of Reading*, *19*(6), 419–433. https://doi.org/10.1080/10888438.2015.1069296

Markwardt, F. C. Jr. (1998). *Peabody Individual Achievement Test-Revised*. American Guidance Service.

Martin-Chang, S., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading*, *31*(3), 273–284. https://doi.org/10.1111/j.1467-9817.2008.00371.x

Martin-Chang, S., Kozak, S., & Rossi, M. (2020). Time to read Young Adult fiction: print exposure and linguistic correlates in adolescents. *Reading and Writing*, *33*(3), 741–760. https://doi.org/10.1007/s11145-019-09987-y

McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press.

McGeown, S. P., Duncan, L. G., Griffiths, Y. M., & Stothard, S. E. (2015). Exploring the relationship between adolescent's reading skills, reading motivation and reading habits. *Reading and Writing*, *28*(4), 545–569. https://doi.org/10.1007/s11145-014-9537-9

McGeown, S. P., Osborne, C., Warhurst, A., Norgate, R., & Duncan, L. G. (2016). Understanding children's reading activities: Reading motivation, skill and child characteristics as predictors. *Journal of Research in Reading*, *39*(1), 109–125. https://doi.org/10.1111/1467-9817.12060

McIntyre, D., & Walker, B. (2015). How can corpora be used to explore the language of poetry and drama? In *The Routledge Handbook of Corpus Linguistics*. Routledge. https://doi.org/10.4324/9780203856949.ch37

Mitchell, K. J. (1997). What happens when school reform and accountability testing meet? *Theory into Practice*, *36*(4), 262–265. https://doi.org/10.1080/00405849709543777

Miyamoto, A., Pfost, M., & Artelt, C. (2018). Reciprocal relations between intrinsic reading motivation and reading competence: A comparison between native and immigrant students in Germany. *Journal of Research in Reading*, *41*(1), 176–196. https://doi.org/10.1111/1467-9817.12113

Mol, S. E., & Bus, A. G. (2011). To Read or Not to Read: A Meta-Analysis of Print Exposure From Infancy to Early Adulthood. *Psychological Bulletin*. https://doi.org/10.1037/a0021890

Monaghan, F. (1999). Judging a word by the company it keeps: The use of concordancing software to explore aspects of the mathematics register. *Language and Education*, *13*(1), 59. https://doi.org/10.1080/09500789908666759

Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First Language*, *39*(5), 527–546. https://doi.org/10.1177/0142723719849996

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science*, *26*(9), 1489–1496. https://doi.org/10.1177/0956797615594361

Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8-and 12-year-old children and adults. *Journal of Experimental Psychology: General*, *144*(2), 447–468. https://doi.org/10.1037/xge0000054

Moore, M., & Gordon, P. C. (2014). Reading ability and print exposure: item response theory analysis of the author recognition test. *Behavior Research Methods*, *47*(4), 1095–1109. https://doi.org/10.3758/s13428-014-0534-3

Morrison, C. M., & Ellis, A. W. (1995). Roles of Word Frequency and Age of Acquisition in Word Naming and Lexical Decision. *JournalofExperimentalPsychology*, *21*(1), 116–133.

Moss, G., & McDonald, J. W. (2004). The borrowers: Library records as unobtrusive measures of children's reading preferences. *Journal of Research in Reading*, *27*(4), 401–412. https://doi.org/10.1111/j.1467-9817.2004.00242.x

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*.

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) Psychology: Measuring and Mapping Scales of Cultural and Psychological Distance. *Psychological Science*, *31*(6), 678–701. https://doi.org/10.1177/0956797620916782

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning Word Meanings From Context During Normal Reading. *American Educational Research Journal*, *27*(2), 237–270. https://doi.org/10.3102/00028312024002237

Nagy, W. E., & Herman, P. A. (1984). Limitations of Vocabulary Instruction. In *Centre for the Study of Reading*.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning Words from Context. *Reading Research Quarterly*, *20*(2), 233. https://doi.org/10.2307/747758

Nash, H., & Snowling, M. (2006). Teaching new words to children with poor existing vocabulary knowledge: A controlled evaluation of the definition and context methods. *International Journal of Language and Communication Disorders*. https://doi.org/10.1080/13682820600602295

Nation, K. (2017). Nurturing a lexical legacy: reading experience is critical for the development of word reading skill. *Npj Science of Learning*, *2*(1), 3. https://doi.org/10.1038/s41539-017-0004-7

Nation, K., & Snowling, M. J. (1998). Semantic Processing and the Development of Word-Recognition Skills: Evidence from Children with Reading Comprehension Difficulties. *Journal of Memory and Language*, *39*(1), 85–101. https://doi.org/10.1006/jmla.1998.2564

Neale, M. D. (1997). *Analysis of Reading Ability: Second Revised British Edition.* NFER-Nelson.

Nelson, M. (2015). Building a written corpus. In *The Routledge Handbook of Corpus Linguistics*. Routledge. https://doi.org/10.4324/9780203856949.ch5

Nelson, M. J., & Denny, E. E. (1973). *The Nelson-Denny Reading Test*. Houghton Miffin Company.

Newson, M., Buhrmester, M., Xygalatas, D., & Whitehouse, H. (2021). Go WILD, Not WEIRD. *Journal for the Cognitive Science of Religion*, *6*(1–2). https://doi.org/10.1558/jcsr.38413

Norman, R., Hulme, R. C., Sarantopoulos, C., Chandran, V., Shen, H., Rodd, J. M., Joseph, H., & Taylor, J. S. H. (2023). Contextual diversity during word learning through reading benefits generalisation of learned meanings to new contexts. *Quarterly Journal of Experimental Psychology*, *76*(7), 1658–1671. https://doi.org/10.1177/17470218221126976

O'Connor, M., Geva, E., & Koh, P. W. (2019). Examining Reading Comprehension Profiles of Grade 5 Monolinguals and English Language Learners Through the Lexical Quality Hypothesis Lens. *Journal of Learning Disabilities*, *52*(3), 232–246. https://doi.org/10.1177/0022219418815646

OECD. (2021). *21st-Century Readers: Developing Literacy Skills in a Digital World* (PISA). OECD. https://doi.org/10.1787/a83d84cb-en

Ofqual. (2011). *Functional Skills Criteria for English*.

Ofqual. (2013). *Reforms to GCSEs in England from 2015* (Issue November 2013).

Ofqual. (2015). *GCSE Subject Level Guidance for English Language July 2015* (Issue July).

Ofqual. (2020). *Summer 2020 results analysis – GCSE , AS and A level*.

Ofsted. (2022). *Research Review Series: English*. https://www.gov.uk/government/publications/curriculum-research-review-series-english/curriculum-research-review-series-english

Ouellette, 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*. https://doi.org/10.1037/0022-0663.98.3.554

Pagán, A., Bird, M., Hsiao, Y., & Nation, K. (2019). Both Semantic Diversity and Frequency Influence Children's Sentence Reading. *Scientific Studies of Reading*, 1–9. https://doi.org/10.1080/10888438.2019.1670664

Pagán, A., & Nation, K. (2019). Learning Words Via Reading: Contextual Diversity, Spacing, and Retrieval Effects in Adults. *Cognitive Science*, *43*(1). https://doi.org/10.1111/cogs.12705

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2002). *The lexical quality hypothesis*. 189–213. https://doi.org/10.1075/swll.11.14per

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading* (pp. 227–247). Blackwell Publishing Ltd.

Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, *26*, 89–102. https://doi.org/10.1016/j.lindif.2013.04.008

Picton, I., & Clark, C. (2022). *Seeing yourself in what you read: diversity and children and young people's reading in 2022*. https://explore-education-statistics.service.gov.uk/find-

Porter, S. R., & Whitcomb, M. E. . (2003). The Impact of Lottery Incentives on Student Survey Response Rates. *Research in Higher Education*, *44*(4), 389–407. https://www.jstor.org/stable/40197312 ON STUDENT SURVEY RESPONSE RATES

Psychological Corporation. (1999). *Wechsler abbreviated scale of intelligence*.

Quigley, A. (2018). *Closing the [Vocabulary] Gap*. Routledge.

Quigley, A. (2020). *Closing the Reading Gap*. Routledge.

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in Word Identification: A Word-Experience Model That Accounts for Morpheme Frequency Effects. *Scientific Studies of Reading*, *7*(3), 219–237. https://doi.org/10.1207/s1532799xssr0703_2

Reppen, R. (2015). Building a corpus. In *The Routledge Handbook of Corpus Linguistics*. Routledge. https://doi.org/10.4324/9780203856949.ch3

Richards, B. (1987). Type/token ratios: What do they really tell us? *Journal of Child Language*, *14*(2), 201–209. https://doi.org/10.1017/S0305000900012885

Richter, T., Isberner, M. B., Naumann, J., & Neeb, Y. (2013). Lexical Quality and Reading Comprehension in Primary School Children. *Scientific Studies of Reading*, *17*(6), 415–434. https://doi.org/10.1080/10888438.2013.764879

Rosa, E., Salom, R., & Perea, M. (2022). Contextual diversity favors the learning of new words in children regardless of their comprehension skills. *Journal of Experimental Child Psychology*, *214*. https://doi.org/10.1016/j.jecp.2021.105312

Rosa, E., Tapia, J. L., & Perea, M. (2017). Contextual diversity facilitates learning new words in the classroom. *PLoS ONE*, *12*(6). https://doi.org/10.1371/journal.pone.0179004

Rose, J. (2006). Independent Review of the Teaching of Early Reading Final Report. In *Independent review* (Issue March). www.standards.dcsf.gov.uk/phonics/report.pdf

Rosner, J., & Simon, D. (1971). The Auditory Analysis Test: An Initial Report. *Journal of Learning Disabilities*. https://doi.org/10.1177/002221947100400706

Sabao, C. (2014). Towards a theory of genre ? Reflections on the problems and debates on theorising ‘ genre .’ *The Dyke*, *8*(2).

Sacks, P. (2000). Predictable Losers in Testing Schemes. *School Adminstrator*, *57*(11), 6–9.

Sadoski, M., & Paivio, A. (2013). Imagery and text: A dual coding theory of reading and writing: Second edition. In *Imagery and Text: A Dual Coding Theory of Reading and Writing: Second Edition*. https://doi.org/10.4324/9780203801932

Salmeron, L., Altamura, L., Delgado, P., Karagiorgi, A., & Vargas, C. (2024). Reading Comprehension on Handheld Devices Versus on Paper: A Narrative Review and Meta-Analysis of the Medium Effect and Its Moderators. *Journal of Educational Psychology*, *116*(2), 153`72. https://doi.org/10.1037/edu0000830.supp

Sampson, M. R., Valmont, W. J., & Allen, R. Van. (1982). The Effects of Instructional Cloze on the Comprehension, Vocabulary, and Divergent Production of Third-Grade Students. *Reading Research Quarterly*. https://doi.org/10.2307/747526

Saris, W. E., & Galhofer, I. N. (2014). *Design, Evaluation, and Analysis of Questionnaires for Survey Research : Design, Evaluation, and Analysis of Questionnaires for Survey Research* (2nd ed.). John Wiley & Sons, Incorporated. https://ebookcentral.proquest.com

Sax, L., Gilmartin, S., & Bryant A. N. (2003). *Assessing Response Rates and Nonresponse Bias in Web and Paper Surveys*. *44*(4), 409–432.

Schaffner, E., Philipp, M., & Schiefele, U. (2016). Reciprocal effects between intrinsic reading motivation and reading competence? A cross-lagged panel model for academic track and nonacademic track students. *Journal of Research in Reading*, *39*(1), 19–36. https://doi.org/10.1111/1467-9817.12027

Schleppegrell, M. J. (2001). *Linguistic Features of the Language of Schooling*.

Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: A research review. In *Reading and Writing Quarterly* (Vol. 23, Issue 2, pp. 139–159). https://doi.org/10.1080/10573560601158461

Schwartz, N., Hippler, H., Deutsch, B., & Strack, F. (1985). Response Scales : Effects of Category Range on Reported Behavior and Comparative Judgments. *Public Opinion Quarterly*, *49*(3), 388–395.

Schwartz, R. M., & Raphael, T. E. (1985). Concept of definition: A key to improving students' vocabulary. *The Reading Teacher*.

Scott, M. (2015). What can corpus software do? In *The Routledge Handbook of Corpus Linguistics*. https://doi.org/10.4324/9780203856949.ch11

Semel, E., Wigg, E. H., & Secord, W. (1995). *The clinical evaluation of language fundamentals* (3rd ed.). Psychological Corporation.

Sénéchal, M., LeFevre, J. A., Hudson, E., & Lawson, E. P. (1996). Knowledge of Storybooks as a Predictor of Young Children's Vocabulary. *Journal of Educational Psychology*, *88*(3), 520–536. https://doi.org/10.1037/0022-0663.88.3.520

Shanahan, T., & Shanahan, C. (2017). Disciplinary Literacy: Just the FAQs. *Educational Leadership*, *74*, 18–22.

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

Snowling, M. J., Stothard, S. E., Clarke, P. J., Bowyer-Crane, C., Harrington, A., Truelove, E., & Hulme, C. (2009). *York Assessment of Reading for Comprehension: GL Assessment*.

Spear-Swerling, L., Brucker, P. O., & Alfano, M. P. (2010). Relationships between sixth-graders' reading comprehension and two different measures of print exposure. *Reading and Writing*, *23*(1), 73–96. https://doi.org/10.1007/s11145-008-9152-8

Spencer, S., Clegg, J., Stackhouse, J., & Rush, R. (2017). Contribution of spoken language and socio-economic background to adolescents' educational achievement at age 16 years. *International Journal of Language and Communication Disorders*. https://doi.org/10.1111/1460-6984.12264

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science*, *8*(6). https://doi.org/10.1111/j.1467-9280.1997.tb00453.x

Stainthorp, R. (1997). A Children's Author Recognition Test: A Useful Tool in Reading Research. *Journal of Research in Reading*, *20*(2), 148–158. https://doi.org/10.1111/1467-9817.00027

Stanovich, K. E. (1986). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/rrq.21.4.1

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition*, *20*(1), 51–68. https://doi.org/10.3758/BF03208254

Stanovich, K. E., & West, R. F. (1989). Exposure to Print and Orthographic Processing. *Source: Reading Research Quarterly*, *24*(4), 402–433.

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: some current controversies. *TrendsinCognitiveSciences*, *5*(4), 156–162. http://tics.trends.com

Sternberg, R. J. (1987). Most vocabulary is learned from context. *The Nature of Vocabulary Acquisition*.

Stothard, S. E., Hulme, C., Clarke, P., Barmby, P., & Snowling, M. J. (2010). *YARC York Assessment of Reading for Comprehension (Secondary): GL Assessment*.

Stutz, F., Schaffner, E., & Schiefele, U. (2016). Relations among reading motivation, reading amount, and reading comprehension in the early elementary grades. *Learning and Individual Differences*, *45*, 101–113. https://doi.org/10.1016/j.lindif.2015.11.022

Sullivan, A., & Brown, M. (2015). Reading for pleasure and progress in vocabulary and mathematics. *British Educational Research Journal*, *41*(6), 971–991. https://doi.org/10.1002/berj.3180

Sullivan, A., Moulton, V., & Fitzsimons, E. (2021). The intergenerational transmission of language skill. *British Journal of Sociology*, *72*(2), 207–232. https://doi.org/10.1111/1468-4446.12780

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: components of the simple view of reading across grade levels. *Journal of Research in Reading*, *32*(4), 383–401. https://doi.org/10.1111/j.1467-9817.2009.01401.x

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of Specific Language Impairment in Kindergarten Children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245–1260. https://doi.org/10.1044/jslhr.4006.1245

Topping, K., Clark, C., Picton, I., & Cole, A. (2023). *What and How Kids are Reading: The Book-Reading Behaviours of Pupils*. https://renaissance.widen.net/view/pdf/priipdurrj/UK-What-Kids-Are-Reading-report-2023.pdf?t.download=true&u=zceria

Torgesen, J. K. , Wagner, R. R., & Rashotte, C. (1999). *Test of Word Reading Ability (TOWRE)*. The Psychological Corp.

Torppa, M., Niemi, P., Vasalampi, K., Lerkkanen, M., Tolvanen, A., & Poikkeus, A. (2020). Leisure Reading (But Not Any Kind) and Reading Comprehension Support Each

Other—A Longitudinal Study Across Grades 1 and 9. *Child Development*, *91*(3), 876–900. https://doi.org/10.1111/cdev.13241

Troyer, M., Kim, J. S., Hale, E., Wantchekon, K. A., & Armstrong, C. (2019). Relations among intrinsic and extrinsic reading motivation, reading amount, and comprehension: a conceptual replication. *Reading and Writing*, *32*(5), 1197–1218. https://doi.org/10.1007/s11145-018-9907-9

University of Arizona Libraries. (2021). *Corpus of Contemporary American English (COCA) 1990 to 2012 Dataset Version 2).* . University of Arizona Research Data Repository.

van Bergen, E., Snowling, M. J., Zeeuw, E. L., Beijsterveldt, C. E. M., Dolan, C. V., & Boomsma, D. I. (2018). Why do children read more? The influence of reading ability on voluntary reading practices. *Journal of Child Psychology and Psychiatry*, *59*(11), 1205–1214. https://doi.org/10.1111/jcpp.12910

van Bergen, E., Vasalampi, K., & Torppa, M. (2021). How Are Practice and Performance Related? Development of Reading From Age 5 to 15. *Reading Research Quarterly*, *56*(3), 415–434. https://doi.org/10.1002/rrq.309

Verhoeven, B. (2022). The politics of GCSE English Language. *English Today*, *38*(4), 244–253. https://doi.org/10.1017/S0266078421000110

Volante, L. (2004). *Teaching To the Test: What Every Educator and Policy-maker Should Know*. http://www.umanitoba.ca/publications/cjeap/articles/volante.html

Vuorinen, P. T. E., Tatler, B. W., & Keller, F. (2023). Tracking e-reading behavior: uncovering the effects of task context, electronic experience, and motivation. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1302701

Walberg, H. J., & Tsai, S. ling. (1984). Reading achievement and diminishing returns to time. *Journal of Educational Psychology*, *76*(3), 442–451. https://doi.org/10.1037/0022-0663.76.3.442

Wechsler, D. (2005). *Wechsler Individual Achievement Test 2nd Edition (WIAT II)*. London Psychological Corp.

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the Real World and Its Correlates. *Reading Research Quarterly*, *28*(1), 34–50.

Westbrook, J., Sutherland, J., Oakhill, J., & Sullivan, S. (2019). 'Just reading': the impact of a faster pace of reading narratives on the comprehension of poorer adolescent readers in English classrooms. *Literacy*, *53*(2), 60–68. https://doi.org/10.1111/lit.12141

Wiederholt, J. L., & Bryant, B. R. (2001). *Gray Oral Reading Test* (4th ed.). PRO-ED.

Wiig, E. H., & Secord, W. (1992). *Test of Word Knowledge*. The Psychological Corporation.

Wild, K., Kilgarriff, A., & Tugwell, D. (2013). The Oxford Children's Corpus: Using a Children's Corpus in Lexicography. *International Journal of Lexicography*, *26*(2), 190–218. https://doi.org/10.1093/ijl/ecs017

Wilder, A. A., & Williams, J. P. (2001). Students with severe learning disabilities can learn higher order comprehension skills. In *Journal of Educational Psychology* (Vol. 93, Issue 2, pp. 268–278). https://doi.org/10.1037/0022-0663.93.2.268

Wilkinson, G. (1993). Wide Range Achievement Test WRAT. *Wide Range Inc*.

Williams, J. P., Brown, L. G., Silverstein, A. K., & de Cani, J. S. (1994). An instructional program in comprehension of narrative themes for adolescents with learning disabilities. *Learning Disability Quarterly*, *17*(3), 205–221. https://doi.org/10.2307/1511074

Wolke, D., Waylen, A., Samara, M., Steer, C., Goodman, R., Ford, T., & Lamberts, K. (2009). Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *British Journal of Psychiatry*, *195*(3), 249–256. https://doi.org/10.1192/bjp.bp.108.053751

Wright, T. S., & Cervetti, G. N. (2017). A Systematic Review of the Research on Vocabulary Instruction That Impacts Text Comprehension. *Reading Research Quarterly*, *52*(2), 203–226. https://doi.org/10.1002/rrq.163

Zipke, M., Ehri, L. C., & Cairns, H. S. (2009). Using Semantic Ambiguity Instruction to Improve Third Graders' Metalinguistic Awareness and Reading Comprehension: An Experimental Study. *Reading Research Quarterly*. https://doi.org/10.1598/rrq.44.3.4

Appendices

Appendix A
Exam Test Corpus Keyword List

No.	Keyword score	Keyword	Part of speech	Fpm in BNC
1	74.49	breaker	noun	1.96
2	69.44	nasally	adverb	0.04
3	42.59	thrill	noun	4.9
4	36.3	boulder	noun	4.69
5	35.23	uncontrollably	adverb	1.04
6	32.93	clang	verb	0.59
7	32.63	napkin	noun	2.66
8	32.41	majestically	adverb	0.49
9	31.47	cravat	noun	0.53
10	30.56	balloon	noun	8.75
11	29.58	slosh	noun	0.11
12	27.34	dispirit	verb	0.77
13	26.41	incessantly	adverb	0.83
14	25.84	swimmer	noun	2.81
15	25.37	Lucy	noun	22.36
16	25.14	handshake	noun	1.89
17	24.72	giddy	adjective	0.95
18	23.74	molten	adjective	2.03
19	23.03	solitary	adjective	0.47
20	22.93	rut	noun	1.4
21	21.93	crockery	noun	1.2
22	21.75	rekindle	verb	1.22
23	21.16	divest	verb	1.28
24	20.99	idleness	noun	1.34
25	20.91	spoonful	noun	1.38
26	20.75	plank	noun	4.77
27	20.67	nightfall	noun	1.35
28	20.39	homework	noun	7.24
29	20.28	endurance	noun	2.64
30	19.91	gobble	verb	1.36
31	19.68	shriek	verb	3.29
32	19.64	ox	noun	2.9
33	19.21	speck	noun	1.58
34	19.21	yank	verb	1.51
35	19.14	spiky	adjective	1.52
36	19.14	savoury	adjective	1.22
37	18.74	smelt	verb	1.34
38	18.69	Iceland	noun	4.14
39	18.06	wade	verb	2.98
40	17.82	stocky	adjective	1.71
41	17.74	trifle	noun	2.79
42	17.7	wistfully	adverb	1.73

No	Keyword score	Keyword	Part of speech	Fpm in BNC
43	17.29	burnt	adjective	3.16
44	17.25	nightdress	noun	1.81
45	17.18	motionless	adjective	3.19
46	16.72	crumpled	adjective	1.89
47	16.67	sickening	adjective	1.9
48	16.57	horrid	adjective	1.9
49	15.8	gust	noun	2.07
50	15.77	quicken	verb	3.56
51	15.75	hoarse	adjective	2.07
52	15.69	bedside	noun	5.12
53	15.66	Pat	noun	17.21
54	15.22	hearty	adjective	2.01
55	14.72	agony	noun	8.94
56	14.49	drip	verb	6.3
57	14.37	fragrant	adjective	2.36
58	14.19	dangle	verb	3.85
59	14	neglected	adjective	2.45
60	13.86	rosy	adjective	2.48
61	13.86	amiable	adjective	2.48
62	13.75	upside	adverb	3.74
63	13.72	mantelpiece	noun	2.52
64	13.72	housekeeping	noun	3.12
65	13.5	ledge	noun	4.44
66	13.45	sane	adjective	2.59
67	13.38	hue	noun	2.99
68	13.12	thrilling	adjective	2.68
69	13.09	tea	noun	74.39
70	13.02	weary	adjective	5.66
71	13.02	throb	verb	3.57
72	12.99	miraculous	adjective	2.71
73	12.64	online	adjective	4.69
74	12.59	spoon	noun	8.28
75	12.58	humiliate	verb	6.6
76	12.49	hasty	adjective	2.87
77	12.46	horribly	adverb	2.88
78	12.43	envelop	verb	2.88
79	11.96	scorch	verb	2.78
80	11.89	wardrobe	noun	9.54
81	11.73	ooze	verb	2.44
82	11.68	tweed	noun	6.97
83	11.63	cork	noun	8.46
84	11.59	crest	noun	6.03
85	11.49	fiercely	adverb	7.32
86	11.42	prisoner	noun	39.57
87	11.36	wail	verb	4.1
88	11.27	soup	noun	12.48

No	Keyword score	Keyword	Part of speech	Fpm in BNC
89	11.27	float	verb	17.03
90	11.17	arrogant	adjective	5.44
91	11.00	landing	noun	20.96
92	10.9	fro	adverb	3.42
93	10.72	sock	noun	9.84
94	10.72	meaningless	adjective	5.71
95	10.69	ghastly	adjective	3.18
96	10.67	kitten	noun	3.59
97	10.61	frantically	adverb	3.55
98	10.61	siren	noun	3.63
99	10.42	jersey	noun	10.32
100	10.34	hopelessly	adverb	3.67
101	10.3	shabby	adjective	3.69
102	10.21	instant	noun	7.66
103	10.21	muffle	verb	3.68
104	10.17	hideous	adjective	3.75
105	10.02	radiate	verb	3.82
106	9.99	mighty	adjective	7.74
107	9.98	seep	verb	3.4
108	9.87	fury	noun	9.53
109	9.87	hillside	noun	6.47
110	9.84	feeble	adjective	3.9
111	9.82	ache	verb	7.19
112	9.77	vessel	noun	25.89
113	9.74	shore	noun	18.9
114	9.73	boat	noun	63.79
115	9.63	tow	noun	3.19
116	9.61	consonant	noun	4.03
117	9.61	wrestle	verb	3.85
118	9.59	rotten	adjective	6.5
119	9.59	amusing	adjective	6.5
120	9.58	Nelson	noun	9.05
121	9.5	ocean	noun	21.88
122	9.43	prison	noun	62.42
123	9.34	eyelid	noun	4.2
124	9.29	sofa	noun	9.35
125	9.28	pizza	noun	5.43
126	9.28	rejoin	verb	4.2
127	9.25	moonlight	noun	4.2
128	9.25	dwindle	verb	4.22
129	9.2	packed	adjective	4.25
130	9.11	blanket	noun	14.21
131	9.11	tub	noun	4.28
132	9.09	dreadful	adjective	12.1
133	9.06	woollen	adjective	3.88
134	9.05	stocking	noun	7.19

No	Keyword score	Keyword	Part of speech	Fpm in BNC
135	9.02	dart	verb	4.53
136	8.84	utmost	adjective	3.12
137	8.8	nostril	noun	4.49
138	8.8	expedition	noun	12.78
139	8.8	defiance	noun	4.58
140	8.77	midday	noun	4.56
141	8.66	ice	noun	35.43
142	8.64	deck	noun	14.34
143	8.59	creak	verb	3.58
144	8.57	foam	verb	1.41
145	8.5	vigour	noun	4.73
146	8.48	distressed	adjective	4.69

Appendix B

Relative Frequencies of Keywords in the David Lee Categories

David Lee Category	**Average relative frequency of keywords**
S_brdcast_discussn	46.78
S_brdcast_documentary	46.70
S_brdcast_news	36.51
S_classroom	61.94
S_consult	19.93
S_conv	52.61
S_courtroom	6.28
S_demonstratn	76.91
S_interview	22.95
S_interview_oral_history	43.29
S_lect_commerce	6.11
S_lect_humanities_arts	58.15
S_lect_nat_science	36.01
S_lect_polit_law_edu	21.83
S_lect_soc_science	25.52
S_meeting	21.87
S_parliament	15.22
S_pub_debate	15.22
S_sermon	37.43
S_speech_scripted	45.43
S_speech_unscripted	39.51
S_sportslive	35.48
S_tutorial	20.01
S_unclassified	32.17
W_ac_humanities_arts	51.33
W_ac_medicine	20.26
W_ac_nat_science	34.93
W_ac_polit_law_edu	17.65
W_ac_soc_science	46.94
W_ac_tech_engin	15.02
W_admin	6.55
W_advert	69.08
W_biography	134.22
W_commerce	30.84
W_email	49.03
W_essay_school	108.73
W_essay_univ	36.30
W_fict_drama	113.11
W_fict_poetry	371.66

David Lee Category	Average relative frequency of keywords
W_fict_prose	227.17
W_hansard	23.67
W_institut_doc	25.79
W_instructional	104.48
W_letters_personal	110.54
W_letters_prof	23.74
W_misc	92.96
W_news_script	62.25
W_newsp_brdsht_nat_arts	145.21
W_newsp_brdsht_nat_commerce	42.83
W_newsp_brdsht_nat_editorial	65.10
W_newsp_brdsht_nat_misc	113.61
W_newsp_brdsht_nat_report	55.52
W_newsp_brdsht_nat_science	66.51
W_newsp_brdsht_nat_social	79.17
W_newsp_brdsht_nat_sports	82.39
W_newsp_other_arts	92.48
W_newsp_other_commerce	40.82
W_newsp_other_report	68.36
W_newsp_other_reportage	56.76
W_newsp_other_science	53.74
W_newsp_other_social	92.87
W_newsp_other_sports	73.83
W_newsp_tabloid	125.12
W_non_ac_humantities_arts	90.01
W_non_ac_medicine	76.70
W_non_ac_nat_science	74.06
W_non_ac_polit_law_edu	40.99
W_non_ac_soc_science	51.20
W_non_ac_teach_engin	29.42
W_pop_lore	110.12
W_religion	80.83

Appendix C

Student Reading Survey

Introduction

The aim of this survey is to find out how much students read and what they read. First you will be asked some general questions about yourself. You will then be asked about how often you read and what types of things you read. You will also be asked about different authors that you may have heard of. Finally, there will be some vocabulary questions.
The survey should take no more than 12 minutes to complete.
Thank you very much for your help.

Ethics and Data Protection

This survey is completely anonymous, nobody will see your answers apart from the research team. The anonymous results may be published, presented at conferences and shared with other researchers. Your details will not be shared with any third parties.

You do not have to take part in this survey and can leave the survey at any time if you change your mind.

The research has been granted ethical approval by the University of Reading. If you have any concerns or would like any further information please contact the researcher Beverley Jennings (email: b.j.jennings@pgr.reading.ac.uk) or her supervisor Dr Holly Joseph (h.joseph@reading.ac.uk). If you would like to receive an electronic copy of the research project, then you can ask for a copy to be emailed to you.

Data Protection

The organisation responsible for protection of your personal information is the University of Reading (the Data Controller). Queries regarding data protection and your rights should be directed to the University Data Protection Officer at imps@reading.ac.uk, or in writing to: Information Management & Policy Services, University of Reading, Whiteknights, P O Box 217, Reading, RG6 6AH.
The University of Reading collects, analyses, uses, shares and retains personal data for the purposes of research in the public interest. Under data protection law we are required to inform you that this use of the personal data we may hold about you is on the lawful basis of being a public task in the public interest and where it is necessary for scientific or historical research purposes. If you withdraw from a research study, which processes your personal data, dependant on the stage of withdrawal, we may still rely on this lawful basis to continue using your data if your withdrawal would be of significant detriment to the research study aims. We will always have in place appropriate safeguards to protect your personal data.
You have certain rights under data protection law which are:

- Withdraw your consent, for example if you opted in to be added to a participant register
- Access your personal data or ask for a copy
- Rectify inaccuracies in personal data that we hold about you
- Be forgotten, that is your details to be removed from systems that we use to process your personal data

- Restrict uses of your data
- Object to uses of your data, for example retention after you have withdrawn from a study

Some restrictions apply to the above rights where data is collected and used for research purposes.
You can find out more about your rights on the website of the Information Commissioners Office (ICO) at https://ico.org.uk
You also have a right to complain the ICO if you are unhappy with how your data has been handled. Please contact the University Data Protection Officer in the first instance.

Section 1 - Consent to take part in the survey

I have read the information provided and understand what I am being asked to do
Yes [ ]
No [ ]

I agree to take part in the survey
Yes [ ]
No [ ]

Section 2 – Some questions about you

What is your gender?
Female [ ]
Male [ ]
Other [ ]
I would prefer not to say [ ]
[Other] You can enter your gender here _____ (gender)

Are you (or have you ever been) entitled to free school meals and/or pupil premium funding?
Yes [ ]
No [ ]
Don't know [ ]
I would prefer not to say [ ]

Do you speak any languages other than English?
Yes [ ]
No [ ]
I would prefer not to say [ ]

What is your postcode? This information will be used to see what areas survey responses are coming from. (leave box blank if you would prefer not to say)
[_____]

What is your ethnic group?
White [ ]
Links to following subcategories:

English/Welsh/Scottish/Northern Irish/British [ ]
Irish [ ]
Gypsy or Irish Traveller [ ]
Any other White background, please describe _____
Mixed/Multiple ethnic groups [ ]
Links to following subcategories:
White and Black Caribbean [ ]
White and Black African [ ]
White and Asian [ ]
Any other Mixed/Multiple ethnic background, please describe _____
Asian/Asian British [ ]
Links to following subcategories:
Indian [ ]
Pakistani [ ]
Bangladeshi [ ]
Chinese [ ]
Any other Asian background, please describe _____
Black/ African/Caribbean/Black British [ ]
Links to following subcategories:
African [ ]
Caribbean [ ]
Any other Black/African/Caribbean background, please describe _____
Other ethnic group [ ]
Links to following subcategories:
Arab [ ]
Any other ethnic group, please describe _____

Prefer not to say [ ]

Section 3 – some questions about English language GCSE

Have you taken your English language GCSE?

Yes [ ]
No [ ]
Prefer not to say [ ]

If yes

In what year did you take English language GCSE?

2018 [ ]
2019 [ ]
2020 [ ]
2021 [ ]
Other [ ]
Prefer not to say [ ]
[Other} _____ (year of GCSE exam)

What grade did you achieve in your English language GCSE? (If you have taken the qualification more than once please enter your highest grade achieved.)

9	[ ]
8	[ ]
7	[ ]
6	[ ]
5	[ ]
4	[ ]
3	[ ]
2	[ ]
1	[ ]
U	[ ]
Prefer not to say	[ ]

If no

Have you been given a target or predicted grade for English language GCSE?

Yes	[ ]
No	[ ]
Don't know	[ ]
Prefer not to say	[ ]

If yes – What is the target or predicted grade that you have been given by your school/college for GCSE English language?

9	[ ]
8	[ ]
7	[ ]
6	[ ]
5	[ ]
4	[ ]
3	[ ]
2	[ ]
1	[ ]
U	[ ]
Prefer not to say	[ ]

Section 4 - Some questions about how much time you spend reading

Think about all the different times you read something yesterday (at home, and at school/college).

Roughly how many minutes in total do you think you spent reading?

By reading I mean reading a reading book, an e-book, a newspaper or online news site, school textbooks or resources used in lesson or for homework, poems, song lyrics or any other large piece of writing.

Minutes spent reading (at home and at school/college) yesterday: e.g.

1 hour = 60 minutes
2 hours = 120 minutes
3 hours = 180 minutes
4 hours = 240 minutes
[_____] (minutes spent reading yesterday)

Is [minutes spent reading yesterday] more, less or the same as the amount of time you spend reading on a typical day?
More than usual [ ]
The same as usual [ ]
Less than usual [ ]

On average, how much time do you think you usually spend reading each day?
This should include both reading at school/college and at home.
[_____] (average minutes reading each day)

Section 5 – Two questions about how your reading compares to other students your age
How does the time you spent reading each day compare to other students your age?
Do you read:
A lot more than others [ ]
A little more than others [ ]
About the same as others [ ]
A little less than others [ ]
A lot less than others [ ]
Not sure/prefer not to say [ ]

How does your enjoyment of reading compare to other students your age?
Do you enjoy reading:
A lot more than others [ ]
A little more than others [ ]
About the same as others [ ]
A little less than others [ ]
A lot less than others [ ]
Not sure/prefer not to say [ ]

Section 6 – Some questions about the type of reading you do
What kinds of different things do you usually read?
Please tick all that apply.
Modern fiction (stories) [ ]
Older/Classic fiction (before 1900) [ ]
Poetry [ ]
Drama/play scripts [ ]
Biographies and autobiographies [ ]
Non-fiction (information) books [ ]

Newspaper articles [ ]
Online articles/blogs [ ]
Magazines [ ]
Other [ ] Specify _____

What type of reading of reading do you do most often?
Please just choose one option that best represents the type of reading you do most often.
Modern fiction (stories) [ ]
Older/Classic fiction (before 1900) [ ]
Poetry [ ]
Drama/play scripts [ ]
Biographies and autobiographies [ ]
Non-fiction (information) books [ ]
Newspaper articles [ ]
Online articles/blogs [ ]
Magazines [ ]
Other [ ] Specify _____

What type of reading of reading do you enjoy the most?
Please just choose one option that best represents the type of reading you enjoy the most.
Modern fiction (stories) [ ]
Older/Classic fiction (before 1900) [ ]
Poetry [ ]
Drama/play scripts [ ]
Biographies and autobiographies [ ]
Non-fiction (information) books [ ]
Newspaper articles [ ]
Online articles/blogs [ ]
Magazines [ ]
Other [ ] Specify _____

What type of reading of reading do you do most for school or college?
Please just choose one option that best represents the type of reading you do for school or college.
Modern fiction (stories) [ ]
Older/Classic fiction (before 1900) [ ]
Poetry [ ]
Drama/play scripts [ ]
Biographies and autobiographies [ ]
Non-fiction (information) books [ ]
Newspaper articles [ ]
Online articles/blogs [ ]
Magazines [ ]

Other [ ] Specify _____

What type of reading of reading do you do most of for pleasure (not for school or college)? Please just choose one option that best represents the type of reading you do for pleasure (not for school or college).

Modern fiction (stories) [ ]
Older/Classic fiction (before 1900) [ ]
Poetry [ ]
Drama/play scripts [ ]
Biographies and autobiographies [ ]
Non-fiction (information) books [ ]
Newspaper articles [ ]
Online articles/blogs [ ]
Magazines [ ]
Other [ ] Specify _____

Section 7 – some names that might be authors

Which of these do you know are real authors?

Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Albany Gardner [ ]
Alfred Thomson [ ]
Annabella Lyons [ ]
Anthony Horowitz [ ]
Anthony McGowan [ ]
Sir Arthur Conan Doyle [ ]
Barbara Darlington [ ]
Benjamin Zephaniah [ ]
Bess Wright [ ]
Caitlin Jones Brewer [ ]

Which of these do you know are real authors?

Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Celestine Hobbs [ ]
Chadwik Humphreys [ ]
Charles Dickens [ ]
Charles Myatt [ ]
Charles Stainton [ ]
Charlotte Bronte [ ]
Charlotte Garcia [ ]
Chris Bradford [ ]
Chris Riddell [ ]
Clara Armstrong [ ]

Which of these do you know are real authors?
Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Corrinne Tucker	[ ]
Dai Glover	[ ]
Daphne Du Maurier	[ ]
Darren Shan	[ ]
Dennis Kelly	[ ]
Derek Landy	[ ]
E G Andrews	[ ]
Edgar Allan Poe	[ ]
Elizabeth Acevedo	[ ]
Elliott Ward	[ ]

Which of these do you know are real authors?
Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Emily Walton	[ ]
Dame Emma Daniel	[ ]
Frank Rees	[ ]
George Eliot	[ ]
George Orwell	[ ]
Geraldine McCaughrean	[ ]
Grover Hayes	[ ]
H G Wells	[ ]
Harper Lee	[ ]
Henry Bowman	[ ]

Which of these do you know are real authors?
Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Hughie Butler	[ ]
J A Reeves	[ ]
J B Priestly	[ ]
J K Rowling	[ ]
Jacqueline Wilson	[ ]
Jane Austen	[ ]
Jason Reynolds	[ ]
Jayden Martinez	[ ]
Jocelyn Benton	[ ]
John Bottles	[ ]

Which of these do you know are real authors?

Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

John Green [ ]
John Keats [ ]
Kamakshi Thomson [ ]
Kazuo Ishiguro [ ]
Kennice Smith Peters [ ]
Kiera Cass [ ]
Kristen Banks [ ]
Leonora Alexander [ ]
Linda Pongham [ ]
Linsay Fletcher [ ]

Which of these do you know are real authors?
Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Lisbet Abraham [ ]
Lois Lowry [ ]
Loriel Johnson [ ]
Louis Sacher [ ]
Louisa May Alcott [ ]
Lulu Brown [ ]
Malachi Frost [ ]
Malorie Blackman [ ]
Mary Hutchins [ ]
Mary Shelley [ ]

Which of these do you know are real authors?
Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Meghan Horton [ ]
Michael Morpurgo [ ]
Morris Gleitzman [ ]
Nicholas Winkler [ ]
Norman Morton [ ]
Oscar Hall [ ]
Patricia Cobham [ ]
Patrick Ness [ ]
Philip Pullman [ ]
Philip Reeve [ ]

Which of these do you know are real authors?
Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Prabhakar Hale [ ]
Rick Riordan [ ]
Robert Louis Stevenson [ ]
Stephenie Meyer [ ]
Sterling Connolly [ ]
Susan Hill [ ]
Suzanne Collins [ ]
Tara Khan [ ]
Terrell Mills [ ]
Sir Thomas Hammond [ ]

Which of these do you know are real authors?
Select the names that you know are real authors. Do not guess, but only select those that you know are authors. Remember not all of the names are real authors.

Thomas Hardy [ ]
Tracy Dowle [ ]
Victor Kaur [ ]
Wilfred Owen [ ]
Wilkie Collins [ ]
William Adkins [ ]
William Blake [ ]
William Golding [ ]
William Shakespeare [ ]
Willy Russell [ ]

Section 8 (last one) – Some vocabulary questions
Which answer is nearest in meaning to the work in italics?

1. Having once got hold they never let go but struggled and wrestled and rolled *incessantly*.
    a. continually
    b. suddenly
    c. freely
    d. separately

Which answer is nearest in meaning to the work in italics?

2. She is weeping *uncontrollably* and making horrible noises
    a. quietly
    b. privately
    c. every now and then
    d. strongly

Which answer is nearest in meaning to the work in italics?

3. The rode in *majestically*.
    a. clumsily

b. impressively
c. carefully
d. slowly

Which answer is nearest in meaning to the work in italics?

4. She had large, grave, *wistfully* attentive eyes
    a. mistily
    b. annoyingly
    c. thoughtfully
    d. beautifully

Which answer is nearest in meaning to the work in italics?

5. The *dispirited* horses plodded sullenly to his command
    a. depressed
    b. cheerful
    c. vicious
    d. ghostly

Which answer is nearest in meaning to the work in italics?

6. It *rekindled* the few expiring embers of freedom
    a. killed
    b. woke-up
    c. treated kindly
    d. enjoyed

Which answer is nearest in meaning to the work in italics?

7. They swam out to the first line of *breakers*, and then diving down were seen no more.
    a. boats
    b. rocks
    c. enemies
    d. waves

Which answer is nearest in meaning to the work in italics?

8. The *boulder* then crushes my right hand and ensnares my right arm at the wrist.
    a. predator
    b. enemy
    c. rock
    d. trap

Which answer is nearest in meaning to the work in italics?

9. She hid her face in her *napkin*, and left the table.
    a. hands
    b. book

c. shame
d. serviette

Which answer is nearest in meaning to the work in italics?

10. He has thick black whiskers, and wears a frock coat, billowy shirt, and *cravat*.
    a. tie
    b. belt
    c. weapon
    d. jewellery

Which answer is nearest in meaning to the work in italics?

11. I was so sick and *giddy* that I was always on the verge of falling.
    a. steady
    b. dizzy
    c. tired
    d. injured

Which answer is nearest in meaning to the work in italics?

12. We can see the river of *molten* lava which snakes away for several miles.
    a. liquid
    b. furry
    c. solid
    d. poisonous

Which answer is nearest in meaning to the work in italics?

13. There stood the figure of a *solitary* woman.
    a. busy
    b. pretty
    c. bored
    d. lone (alone)

Which answer is nearest in meaning to the work in italics?

14. I was going to make a little *savoury* mince meat for supper or breakfast.
    a. salty or spicy
    b. sweet
    c. home-made
    d. vegetarian

Which answer is nearest in meaning to the work in italics?

15. There were quiet lumps of sheep licking the *spiky* grass.
    a. soft
    b. damp
    c. pointed
    d. tasty

Appendix D

Individual Author Scores on the ART

F = foil

YA = Young adult author

C = Classic/curriculum author

Order in ART	**Author**	**Subset**	**Score**
44	J K Rowling	YA	319
99	William Shakespeare	C	317
13	Charles Dickens	C	304
45	Jacqueline Wilson	YA	275
46	Jane Austen	C	274
35	George Orwell	C	234
72	Michael Morpurgo	YA	228
28	Edgar Allan Poe	C	225
70	Mary Shelley	C	214
4	Anthony Horowitz	YA	212
94	Wilfred Owen	C	208
6	Sir Arthur Conan Doyle	C	189
38	H G Wells	C	188
51	John Green	YA	183
43	J B Priestly	C	181
16	Charlotte Bronte	C	176
87	Suzanne Collins	YA	174
97	William Blake	C	169
68	Malorie Blackman	YA	168
79	Philip Pullman	YA	135
83	Robert Louis Stevenson	C	131
39	Harper Lee	YA	128
91	Thomas Hardy	C	123
84	Stephenie Meyer	YA	121
52	John Keats	C	120
82	Rick Riordan	YA	106
78	Patrick Ness	YA	100
98	William Golding	C	97
34	George Eliot	C	72
54	Kazuo Ishiguro	C	71
65	Louisa May Alcott	C	67
23	Daphne Du Maurier	C	54
80	Philip Reeve	YA	50
8	Benjamin Zephaniah	C	46
86	Susan Hill	C	45

Order in ART	Author	Subset	Score
19	Chris Riddell	YA	33
100	Willy Russell	C	31
30	Elliott Ward	F	28
64	Louis Sacher	YA	28
18	Chris Bradford	YA	27
5	Anthony McGowan	YA	23
56	Kiera Cass	YA	21
26	Derek Landy	YA	20
24	Darren Shan	YA	19
25	Dennis Kelly	C	19
73	Morris Gleitzman	YA	17
62	Lois Lowry	YA	15
2	Alfred Thomson	F	15
42	J A Reeves	F	15
31	Emily Watson	F	14
27	E G Andrews	F	14
95	Wilkie Collins	C	12
40	Henry Bowman	F	11
3	Annabella Lyons	F	10
20	Clara Armstrong	F	10
57	Kirsten Banks	F	10
60	Linsay Fletcher	F	8
47	Jason Reynolds	YA	8
1	Albany Gardener	F	6
36	Geraldine McCaughrean	YA	5
10	Caitlin Jones Brewer	F	5
29	Elizabeth Acevedo	YA	5
90	Sir Thomas Hammond	F	5
9	Bess Wright	F	5
21	Corrinne Tucker	F	5
67	Malachi Frost	F	5
17	Charlotte Garcia	F	4
69	Mary Hutchins	F	4
92	Tracy Dowle	F	4
76	Oscar Hall	F	4
33	Frank Rees	F	3
53	Kamakshi Thomson	F	3
11	Celestine Hobbs	F	3
22	Dai Glover	F	3
71	Meghan Horton	F	3

Order in ART	Author	Subset	Score
88	Tara Khan	F	3
37	Grover Hayes	F	2
7	Barbara Darlington	F	2
32	Dame Emma Daniel	F	2
41	Hughie Butler	F	2
58	Leonora Alexander	F	2
77	Patricia Cobham	F	2
85	Sterling Connolly	F	2
12	Chadwick Humphries	F	1
14	Charles Myatt	F	1
48	Jayden Martinez	F	1
15	Charles Stainton	F	1
49	Jocelyn Benton	F	1
50	John Bottles	F	1
55	Kennice Smith Peters	F	1
61	Lisbet Abraham	F	1
66	Lulu Brown	F	1
75	Norman Morton	F	1
93	Victor Kaur	F	1
96	Williams Adkins	F	1
74	Nicholas Winkler	F	0
59	Linda Pongham	F	0
63	Loriel Johnson	F	0
81	Prabhakar Hale	F	0
89	Terrell Mills	F	0

Appendix E
Keyword Vocabulary Test Results

Question	Vocabulary tested	Frequency per million (fpm)	Correct answers (n= 330)	Percentage correct (%)	External Answer Validation result (%)
1	incessantly	0.83	290	88	100
2	uncontrollably	1.04	311	94	98
3	majestically	0.49	299	91	99
4	wistfully	1.73	229	69	82
5	dispirited	0.77	188	57	80
6	rekindled	1.22	303	92	98
7	breakers	1.96	202	61	80
8	boulder	4.69	319	97	99
9	napkin	2.66	204	62	86
10	cravat	0.53	207	63	96
11	giddy	0.95	283	86	94
12	molten	2.03	232	70	83
13	solitary	0.47	204	62	88
14	savoury	1.22	312	95	95
15	spiky	1.52	321	97	100

Appendix F

Independent Reading Corpus

Top 100 nouns by frequency

No.	Noun	Frequency
1	time*	50
2	dursley	40
3	year*	38
4	eye	38
5	people	35
6	harry	34
7	house	31
8	film*	30
9	room	30
10	man	29
11	word*	28
12	mother	27
13	lydia	27
14	day*	26
15	world	26
16	hand	26
17	thing	25
18	way*	25
19	mr	24
20	cheetah	23
21	something	23
22	face	22
23	mrs	21
24	general	20
25	table	19
26	car	19
27	life*	19
28	body	18
29	foot	18
30	sophie	18
31	window	18
32	hour	18
33	place*	18
34	school*	17
35	mind	17
36	hair	17
37	club	17
38	snape	17
39	chelsea	17
40	floor	16
41	father	16
42	air	16
43	today	16
44	cassie	16
45	moment	15
46	wall	15
47	yaxley	14
48	step	14
49	line*	14
50	art	14
51	nothing	14
52	door	14
53	voldemort	14
54	cat	14
55	boy	14
56	street	14
57	lukaku	13
58	[number]	13
59	uncle	13
60	head	13
61	summer	13
62	makhi	13
63	month	12
64	animal	12
65	human	12
66	child*	12
67	week	12
68	light*	12
69	queen*	12
70	cup	12
71	simon	12
72	dudley	12
73	charlie	12
74	front	12

No.	Noun	Frequency
75	course	12
76	road	12
77	thomas	12
78	drogba	11
79	voice	11
80	vernon	11
81	name	11
82	brother	11
83	country	11
84	lot	11
85	age	11
86	nought	10
87	minister	10
88	home	10
89	town	10
90	marilyn	10
91	night	10
92	jazz	10
93	city	10
94	person	10
95	minute	10
96	cinema	10
97	end	10
98	phone	10
99	corner	10
100	right	10

* also in CRC top 100 nouns by frequency

Appendix G

Curriculum Reading Corpus

Top 100 nouns by frequency

No.	Noun	Frequency
1	galaxy	47
2	probability	47
3	metal	46
4	film*	44
5	development	42
6	network	40
7	year*	38
8	copper	36
9	london	35
10	essex	34
11	child*	33
12	example	33
13	carbon	30
14	chord	30
15	question	29
16	type	28
17	skill	28
18	aggression	26
19	datum	26
20	policy	26
21	idea	25
22	earth	24
23	note	23
24	line*	23
25	product	22
26	creation	21
27	universe	20
28	firewall	20
29	people	20
30	information	20
31	god	20
32	area	20
33	design	19
34	paper	19
35	oxide	19
36	answer	19
37	point	19
38	ore	19
39	sport	18
40	number	18
41	use	18
42	life*	18
43	p	17
44	plant	17
45	way*	16
46	time*	16
47	elizabeth	16
48	source	16
49	work	16
50	school*	16
51	day*	16
52	personality	15
53	space	15
54	activity	15
55	counter	15
56	mark	15
57	list	15
58	shift	14
59	size	14
60	queen*	14
61	task	14
62	reason	13
63	c	13
64	performance	13
65	light*	13
66	image	13
67	fruit	13
68	big	12
69	world	12
70	2	12
71	b	12
72	part	12
73	bang	12
74	3	12

No.	Noun	Frequency
75	student	12
76	scale	11
77	1	11
78	place*	11
79	rule	11
80	figure	11
81	theory	11
82	[number]	11
83	group	11
84	level	10
85	word*	10
86	host	10
87	impact	10
88	sun	10
89	story	10
90	melody	10
91	land	9
92	ii	9
93	spectrum	9
94	earl	9
95	book	9
96	gibberellin	9
97	coin	9
98	site	9
99	theme	9
100	rhythm	9

* also in IRC top 100 nouns by frequency

Appendix H

Independent Reading Corpus

Top 100 verbs by frequency

No.	Verb	Frequency
1	be*	819
2	have*	304
3	do*	125
4	say*	104
5	look*	56
6	know*	56
7	make*	55
8	see*	51
9	go*	47
10	take*	46
11	come*	44
12	think*	39
13	get*	37
14	turn	27
15	want	26
16	feel	26
17	find*	26
18	give*	25
19	call	22
20	tell	21
21	stop	20
22	watch	20
23	sit	19
24	play*	19
25	stand	16
26	try*	16
27	like	15
28	happen	15
29	ask*	15
30	buy	14
31	move*	14
32	use*	14
33	put	14
34	seem	14
35	reach	14
36	start*	14
37	become*	13
38	win*	13
39	let	13
40	run	13
41	hear	12
42	leave	12
43	stare	12
44	pull	12
45	hope	12
46	speak	12
47	wear	11
48	walk	11
49	decide*	11
50	believe*	11
51	keep	11
52	hold	11
53	open	11
54	learn*	11
55	grow*	11
56	wait	11
57	miss	11
58	remember	11
59	work*	10
60	enjoy	10
61	fall*	10
62	pass	10
63	follow*	10
64	drive	10
65	love	9
66	show*	9
67	spend	9
68	begin	8
69	mean*	8
70	allow*	8
71	understand*	8
72	bring	8
73	expect*	8
74	cross	8

No.	Verb	Frequency
75	live*	7
76	catch	7
77	return	7
78	talk	7
79	draw*	7
80	hang	7
81	die	7
82	read	7
83	lead*	7
84	lie	6
85	lose*	6
86	forget	6
87	scream	6
88	slide	6
89	smile	6
90	cry	6
91	shoot	6
92	need*	6
93	meet	6
94	place	6
95	listen	6
96	change*	6
97	rise	6
98	create*	6
99	choose*	6
100	bear	6

* also in CRC top 100 verbs by frequency

Appendix I

Curriculum Reading Corpus

Top 100 verbs by frequency

No.	Verb	Frequency
1	be*	544
2	have*	93
3	do*	61
4	use*	53
5	make*	47
6	create*	38
7	give*	32
8	explain	30
9	move*	28
10	need*	28
11	help	27
12	describe	26
13	extract	22
14	show*	21
15	write	20
16	know*	19
17	find*	19
18	choose*	18
19	develop	17
20	see*	16
21	get*	15
22	identify	15
23	volunteer	15
24	mean*	15
25	understand*	15
26	follow*	15
27	believe*	15
28	look*	14
29	try*	14
30	complete	13
31	take*	12
32	involve	12
33	expand	12
34	win*	12
35	start*	11
36	change*	11
37	include	11
38	think*	11
39	produce	10
40	allow*	10
41	accept	10
42	play*	10
43	increase	10
44	work*	10
45	gather	9
46	build	9
47	go*	9
48	demonstrate	9
49	become*	9
50	select	9
51	break	9
52	provide	8
53	contain	8
54	live*	8
55	support	8
56	bear	8
57	draw*	7
58	base	7
59	fill	7
60	answer	7
61	come*	7
62	grow*	7
63	record	7
64	promote	7
65	lose*	7
66	land	7
67	associate	7
68	appear	6
69	cause	6
70	observe	6
71	improve	6
72	number	6
73	edit	6
74	receive	6

No.	Verb	Frequency
75	ask*	6
76	compare	6
77	learn*	6
78	collect	6
79	vary	6
80	add	6
81	measure	5
82	relate	5
83	analyse	5
84	expect*	5
85	fall*	5
86	refer	5
87	affect	5
88	represent	5
89	require	5
90	act	5
91	prevent	5
92	ripen	5
93	gain	5
94	suggest	5
95	say*	5
96	decide*	5
97	design	5
98	lead*	5
99	determine	5
100	marry	5

* also in IRC top 100 verbs by frequency

Appendix J

Independent Reading Corpus

Top 100 adjectives by frequency

No.	Adjective	Frequency
1	good*	33
2	other*	31
3	first*	27
4	last	24
5	old*	23
6	new*	23
7	more*	23
8	long*	23
9	high*	21
10	few*	20
11	young	19
12	dark*	19
13	same*	19
14	great*	17
15	small*	17
16	little	13
17	next*	13
18	most*	12
19	large*	12
20	tall	11
21	big*	11
22	own*	10
23	late*	10
24	whole*	9
25	different*	9
26	hard	9
27	least*	9
28	many*	9
29	only	8
30	important*	8
31	white*	8
32	green*	8
33	thin	7
34	black	7
35	fast	7
36	fine	7
37	popular	7

No.	Adjective	Frequency
38	bad	7
39	tiny	6
40	pale	6
41	strange	6
42	short*	6
43	deep	6
44	single*	6
45	past	6
46	such*	6
47	open	6
48	magnificent	6
49	certain	6
50	possible*	6
51	dead	5
52	normal	5
53	front	5
54	major*	5
55	outside	5
56	upstairs	5
57	key*	5
58	final*	5
59	difficult	5
60	simple	5
61	warm	5
62	extraordinary	5
63	wrong	5
64	familiar	4
65	able*	4
66	bright	4
67	successful	4
68	clean	4
69	nice	4
70	early*	4
71	blue*	4
72	third	4
73	sure*	4
74	clear*	4

No.	Adjective	Frequency
75	silent	4
76	stupid	4
77	red*	4
78	royal	4
79	several	4
80	low	4
81	wide*	4
82	enormous	4
83	prominent	4
84	full*	4
85	usual	4
86	wild	4
87	close	4
88	cool	4
89	second	4
90	human*	4
91	right	4
92	sharp	4
93	straight	3
94	pink	3
95	easy	3
96	enough	3
97	all-male	3
98	special	3
99	unable	3
100	rare	3

* also in CRC top 100 adjectives by frequency

Appendix K

Curriculum Reading Corpus

Top 100 adjectives by frequency

No.	Adjective	Frequency
1	red*	21
2	other*	21
3	different*	17
4	many*	17
5	more*	16
6	new*	15
7	such*	14
8	important*	13
9	physical	13
10	good*	13
11	green*	12
12	yellow	11
13	primary	10
14	able*	10
15	human*	10
16	most*	9
17	same*	9
18	own*	9
19	specific	9
20	great*	8
21	relative	8
22	possible*	8
23	indirect	8
24	next*	8
25	key*	8
26	white*	7
27	blue*	7
28	likely	7
29	big*	7
30	short*	7
31	reactive	7
32	light	7
33	random	7
34	small*	7
35	secondary	6
36	single*	6
37	scientific	6
38	less	6
39	personal	6
40	long*	6
41	holistic	6
42	social	6
43	relevant	5
44	first*	5
45	appropriate	5
46	early*	5
47	fair	5
48	distant	5
49	positive	5
50	dark*	5
51	high*	5
52	direct	5
53	[number]	5
54	individual	5
55	cheap	5
56	whole*	5
57	large*	5
58	powerful	4
59	rich	4
60	late*	4
61	biological	4
62	major*	4
63	future	4
64	basic	4
65	complex	4
66	final*	4
67	total	4
68	pedal	4
69	least*	4
70	emotional	4
71	sure*	4
72	interesting	4
73	old*	4
74	wide*	4

No.	Adjective	Frequency
75	industrial	3
76	active	3
77	graphic	3
78	ready	3
79	reserved	3
80	furious	3
81	further	3
82	talkative	3
83	environmental	3
84	equal	3
85	useful	3
86	clear*	3
87	national	3
88	cognitive	3
89	visible	3
90	sociable	3
91	few*	3
92	flowering	3
93	available	3
94	practical	3
95	former	3
96	full*	3
97	backup	3
98	religious	3
99	responsible	3
100	written	3

* also in IRC top 100 adjectives by

frequency

Appendix L

Independent Reading Corpus

Top 100 adverbs by frequency

No.	Adverb	Frequency
1	not*	207
2	so*	57
3	now*	45
4	back*	42
5	then*	36
6	as*	34
7	still*	31
8	only*	31
9	just*	27
10	very*	24
11	even*	24
12	never*	23
13	too*	21
14	always*	20
15	well*	18
16	down*	16
17	again*	16
18	here*	15
19	all*	15
20	more*	15
21	away*	15
22	ever*	14
23	about*	14
24	much*	13
25	there*	13
26	once*	12
27	right	12
28	most*	11
29	also*	11
30	already*	11
31	up*	10
32	sometimes*	10
33	quite	9
34	no*	9
35	often*	9
36	ago	8
37	instead*	8

No.	Adverb	Frequency
38	maybe	7
39	else*	7
40	finally	7
41	quickly*	7
42	out*	7
43	perhaps*	7
44	rather*	6
45	especially*	6
46	somehow	6
47	really	6
48	almost*	6
49	little	6
50	soon	6
51	alone*	6
52	enough	5
53	anywhere	5
54	actually	5
55	yet*	5
56	though	5
57	together*	5
58	upwards	5
59	nearly*	4
60	obviously	4
61	exactly*	4
62	anyway	4
63	probably	4
64	purely	4
65	completely	4
66	hardly	4
67	straight	4
68	directly*	4
69	later*	4
70	forward	4
71	long	3
72	either	3
73	before*	3
74	behind	3

No.	Adverb	Frequency
75	below*	3
76	ahead*	3
77	far*	3
78	perfectly	3
79	first*	3
80	forever	3
81	slowly	3
82	apparently	3
83	hard*	3
84	strangely	3
85	suddenly	3
86	sure	3
87	however*	3
88	twice	3
89	usually*	3
90	lately	3
91	barely	2
92	longer*	2
93	easily	2
94	merely	2
95	nothing	2
96	excitedly	2
97	outward	2
98	anymore	2
99	alike	2
100	particularly*	2

* also in CRC top 100 adverbs by frequency

Appendix M

Curriculum Reading Corpus

Top 100 adverbs by frequency

No.	Adverb	Frequency
1	not*	51
2	away*	18
3	more*	15
4	then*	13
5	so*	12
6	also*	12
7	as*	11
8	only*	11
9	usually*	11
10	here*	9
11	well*	9
12	there*	8
13	back*	7
14	now*	7
15	very*	7
16	even*	6
17	about*	6
18	less	6
19	all*	5
20	often*	5
21	therefore	5
22	too*	5
23	just*	5
24	rather*	4
25	most*	4
26	naturally	4
27	further	4
28	always*	4
29	quickly*	4
30	later*	4
31	however*	4
32	longer*	3
33	equally	3
34	never*	3
35	before*	3
36	once*	3
37	exactly*	3
38	out*	3
39	far*	3
40	faster	3
41	randomly	3
42	directly*	3
43	down*	3
44	much*	2
45	else*	2
46	newly	2
47	especially*	2
48	almost*	2
49	again*	2
50	already*	2
51	below*	2
52	particularly*	2
53	perhaps*	2
54	clearly	2
55	still*	2
56	ahead*	2
57	strongly	2
58	instead*	2
59	together*	2
60	up*	2
61	first*	2
62	artificially	2
63	yet*	2
64	efficiently	1
65	neither	1
66	nearly*	1
67	nearby	1
68	effectively	1
69	normally	1
70	no*	1
71	everywhere	1
72	ever*	1
73	alone*	1
74	outdoors	1
75	originally	1
76	overall	1

No.	Adverb	Frequency
77	overleaf	1
78	over	1
79	extrinsically	1
80	quick	1
81	quicker	1
82	freely	1
83	rapidly	1
84	functionally	1
85	fully	1
86	but	1
87	boldly	1
88	aesthetically	1
89	rudely	1
90	secretly	1
91	significantly	1
92	sometimes*	1
93	furthest	1
94	stepwise	1
95	hard*	1
96	heavily	1
97	closer	1
98	closely	1
99	through	1
100	uncontrollably	1

* also in IRC top 100 adverbs by frequency

Appendix N

Ethical Approval Documentation

University of Reading
Institute of Education
Ethical Approval Form A (version May 2019)

Tick one:
Staff project: _____ PhD _✓_ EdD ____

Name of applicant (s): **Beverley Jennings**

Title of project: **Adolescent reading practices: a corpus linguistics approach to defining success**

Name of supervisor (for student projects): **Dr Holly Joseph & Dr Daisy Powell**

Please complete the form below including relevant sections overleaf.

	YES	NO
Have you prepared an Information Sheet for participants and/or their parents/carers that:		
a) explains the purpose(s) of the project	✓	
b) explains how they have been selected as potential participants	✓	
c) gives a full, fair and clear account of what will be asked of them and how the information that they provide will be used	✓	
d) makes clear that participation in the project is voluntary	✓	
e) explains the arrangements to allow participants to withdraw at any stage if they wish	✓	
f) explains the arrangements to ensure the confidentiality of any material collected during the project, including secure arrangements for its storage, retention and disposal	✓	
g) explains the arrangements for publishing the research results and, if confidentiality might be affected, for obtaining written consent for this	✓	
h) explains the arrangements for providing participants with the research results if they wish to have them	✓	
i) gives the name and designation of the member of staff with responsibility for the project together with contact details, including email . If any of the project investigators are students at the IoE, then this information must be included and their name provided	✓	
k) explains, where applicable, the arrangements for expenses and other payments to be made to the participants	n/a	
j) includes a standard statement indicating the process of ethical review at the University undergone by the project, as follows: 'This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct'.	✓	
k) includes a standard statement regarding insurance: "The University has the appropriate insurances in place. Full details are available on request".	✓	
Please answer the following questions		
1) Will you provide participants involved in your research with all the information necessary to ensure that they are fully informed and not in any way deceived or misled as to the purpose(s) and nature of the research? (Please use the subheadings used in the example information sheets on blackboard to ensure this).	✓	
2) Will you seek written or other formal consent from all participants, if they are able to provide it, in addition to (1)?	✓	
3) Is there any risk that participants may experience physical or psychological distress in taking part in your research?		✓
4) Staff Only - have you taken the online training modules in data protection and information security (which can be found here: http://www.reading.ac.uk/internal/humanresources/PeopleDevelopment/newstaff/humres-MandatoryOnlineCourses.aspx Please note: students complete a Data Protection Declaration form and submit it with this application to the ethics committee.		
5) Have you read the Health and Safety booklet (available on Blackboard) and completed a Risk Assessment Form to be included with this ethics application?	✓	

6) Does your research comply with the University's Code of Good Practice in Research?	✓		
	YES	NO	N.A.
7) If your research is taking place in a school, have you prepared an information sheet and consent form to gain the permission in writing of the head teacher or other relevant supervisory professional?	✓		
8) Has the data collector obtained satisfactory DBS clearance?	✓		
9) If your research involves working with children under the age of 16 (or those whose special educational needs mean they are unable to give informed consent), have you prepared an information sheet and consent form for parents/carers to seek permission in writing, or to give parents/carers the opportunity to decline consent?	✓		
10) If your research involves processing sensitive personal data[1], or if it involves audio/video recordings, have you obtained the explicit consent of participants/parents?			✓
11) If you are using a data processor to subcontract any part of your research, have you got a written contract with that contractor which (a) specifies that the contractor is required to act only on your instructions, and (b) provides for appropriate technical and organisational security measures to protect the data?			✓
12a) Does your research involve data collection outside the UK?		✓	
12b) If the answer to question 12a is "yes", does your research comply with the legal and ethical requirements for doing research in that country?			✓
13a) Does your research involve collecting data in a language other than English?		✓	
13b) If the answer to question 13a is "yes", please confirm that information sheets, consent forms, and research instruments, where appropriate, have been directly translated from the English versions submitted with this application.			✓
14a. Does the proposed research involve children under the age of 5?		✓	
14b. If the answer to question 14a is "yes": My Head of School (or authorised Head of Department) has given details of the proposed research to the University's insurance officer, and the research will not proceed until I have confirmation that insurance cover is in place.			✓
If you have answered YES to Question 3, please complete Section B below			✓

- Complete **either** Section A **or** Section B below with details of your research project.
- Complete a risk assessment.
- Sign the form in Section C.
- Append at the end of this form all relevant documents: information sheets, consent forms, tests, questionnaires, interview schedules, evidence that you have completed information security training (e.g. screen shot/copy of certificate).
- Email the completed form to the Institute's Ethics Committee for consideration.

Any missing information will result in the form being returned to you.

A: My research goes beyond the 'accepted custom and practice of teaching' but I consider that this project has **no** significant ethical implications. (Please tick the box.)	✓
Please state the total number of participants that will be involved in the project and give a breakdown of how many there are in each category e.g. teachers, parents, pupils etc. Reading record: 30 students Online survey: sent to approx. 5,000 students	
Give a brief description of the aims and the methods (participants, instruments and procedures) of the project in up to 200 words noting: 1. title of project 2. purpose of project and its academic rationale 3. brief description of methods and measurements 4. participants: recruitment methods, number, age, gender, exclusion/inclusion criteria 5. consent and participant information arrangements, debriefing (attach forms where necessary)	

[1] Sensitive personal data consists of information relating to the racial or ethnic origin of a data subject, their political opinions, religious beliefs, trade union membership, sexual life, physical or mental health or condition, or criminal offences or record.

6. a clear and concise statement of the ethical considerations raised by the project and how you intend to deal with then.
7. estimated start date and duration of project

1. Title: Adolescent reading practices: a corpus linguistics approach to defining success

2. The purpose of this project is to survey the reading habits of students in KS4 and KS5 in at least two educational institutions, one an 11-16 school and one a 6th form college. This is part of a PhD examining the types of reading practices that are most likely to expose students to the kinds of vocabulary present in English language GCSE reading texts. The exam texts are being analysed using corpus linguistics in the first part of this research. The student reading record and student surveys are the second part.

3. The survey will take place first. The survey link will be sent to students' school/college email accounts. For Key Stage 5 participants, that is those over 16, the survey link will also be circulated via social media in order to reach some students outside the participating schools/colleges. The survey will ask students about how much they read and what genres of texts they read. There will also be an author recognition test, where participants will identify the names of authors they know, as a measure of print exposure. Vocabulary knowledge will also be tested using multiple choice questions. A copy of the survey is included in the appendices.

   The next stage of the project will be to collect reading material from year 10 lessons. These could be photos of worksheets, pages from textbooks, pages from reading books or newspaper articles, anything that the student is asked to read in a lesson. The researcher will ask one teacher per subject to provide reading material from a week of year 10 lessons . The researcher will then code the texts and convert them into plain text files and save them on a password protected computer. Students reading at will be represented using the results from a year 10 mixed ability English class homework assignment set by the researcher in February 2022. Students were told to read independently over the half-term and then provide details of what they read. Using the list of books and other reading materials collected the researcher will collect samples of these texts. These texts will be used to create a corpus to represent students' actual reading materials. Corpus linguistic analysis tools will then be used to analyse and compare the corpus to reference corpora and the exam text corpus created earlier in this project.

4. The online survey will be sent via email to all KS4 and/or KS5 students in at least two institutions. It will also be circulated on social media to personal contacts of the researcher.

   For the reading record group, a mixed ability year 10 class or form (approx. 30 participants) will be selected to represent a range of prior attainment levels, a balance for gender and be representative for SEN and EAL.

5. There will be two separate consent and participation arrangements for the survey. The KS5 participants are over 16 and do not require parental consent. The link will be sent to their college email address or via social media inviting them to take part and they can choose whether or not to click on the link. At the beginning of the online survey there will be information about the survey, ethics and data protection. Participants will be asked to indicate that they understand the information and agree to take part. If they do not tick that they agree with these then the survey will not continue.

   To collect classroom reading material an information sheet and consent form will be given to each teacher. The researcher will explain the research during a year 10 English class and then parents and students will be emailed the information and the opt-out consent to use the list of reading material submitted for the homework assignment in February (emails on pages 11 & 13 below). Students and parents will be given at least one week within which to withdraw.

6. Ethical considerations are mainly around confidentiality, which is fully explained in the information leaflets. Whilst the researcher will be able to identify individual participants in the reading record group, all records will be anonymised at the point of data collection. Participants will be assigned a number and only that number will be used in anything that may be published. No identifiers linking participants or the school to the study will be included.

7. It is estimated that the reading record will be trialled in Spring 2021 with the full study taking place Summer 2021.

   It is also anticipated that the online survey will be trialled Summer 2021 and then will take place Summer and Autumn 2021.

B: I consider that this project **may** have ethical implications that should be brought before the Institute's Ethics Committee.	

Please state the total number of participants that will be involved in the project and give a breakdown of how many there are in each category e.g. teachers, parents, pupils etc.

Give a brief description of the aims and the methods (participants, instruments and procedures) of the project in up to 200 words.

1. title of project
2. purpose of project and its academic rationale
3. brief description of methods and measurements
4. participants: recruitment methods, number, age, gender, exclusion/inclusion criteria
5. consent and participant information arrangements, debriefing (attach forms where necessary)
6. a clear and concise statement of the ethical considerations raised by the project and how you intend to deal with then.
7. estimated start date and duration of project

C: SIGNATURE OF APPLICANT:

Note: a signature is required. Typed names are not acceptable.

I have declared all relevant information regarding my proposed project and confirm that ethical good practice will be followed within the project.

Signed:…… Print Name…B Jennings…. Date…25/3/2021…….

STATEMENT OF ETHICAL APPROVAL FOR PROPOSALS SUBMITTED TO THE INSTITUTE ETHICS COMMITTEE

This project has been considered using agreed Institute procedures and is now approved.

Signed: … Print Name…Holly Joseph…. Date…6/5/2021
(IoE Research Ethics Committee representative)*

* A decision to allow a project to proceed is not an expert assessment of its content or of the possible risks involved in the investigation, nor does it detract in any way from the ultimate responsibility which students/investigators must themselves have for these matters. Approval is granted on the basis of the information declared by the applicant.

University of Reading
Institute of Education
Risk Assessment Form for Research Activities February 2014

Select one:
Staff project: ☐ PGR project: ✓ MA/UG project: ☐

Name of applicant (s): **Beverley Jennings**

Title of project **Adolescent reading practices: a corpus linguistics approach to defining success**

Name of supervisor (for student projects): **Dr Holly Joseph & Dr Daisy Powell**

A: Please complete the form below

Brief outline of Work/activity:	Collection of reading records and an online survey.

Where will data be collected?	The reading records will be collected in a participating school and anonymised at the point of collection. Data will be kept electronically using the University storage (Onedrive) on a password protected computer and on the corpus tool Sketch Engine. The online survey will be created in the University REDCap software. A link will be sent to participants' educational institution email. The data will be kept securely in REDCap and analysed using SPSS software.

Significant hazards:	

Who might be exposed to hazards?	No-one

Existing control measures:	n/a

Are risks adequately controlled:	Yes ✓ No

If NO, list additional controls and actions required:	Additional controls	Action by:

B: SIGNATURE OF APPLICANT:

I have read the Heath and Safety booklet posted on Blackboard, and the guidelines overleaf.
I have declared all relevant information regarding my proposed project and confirm risks have been adequately assessed and will be minimized as far as possible during the course of the project.

Signed: Print Name…B Jennings ……. Date…25/3/2021

STATEMENT OF APPROVAL TO BE COMPLETED BY SUPERVISOR (FOR UG AND MA STUDENTS) **OR** BY IOE ETHICS COMMITTEE REPRESENTATIVE (FOR PGR AND STAFF RESEARCH).

This project has been considered using agreed Institute procedures and is now approved.

Signed: … Print Name Holly Joseph Date…6/5/21….

* A decision to allow a project to proceed is not an expert assessment of its content or of the possible risks involved in the investigation, nor does it detract in any way from the ultimate responsibility which students/investigators must themselves have for these matters. Approval is granted on the basis of the information declared by the applicant.