This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Declarations

1. I declare that this thesis, presented for the degree of Doctor of Philosophy, has been composed solely by myself, Elinor C. Bridges, unless otherwise acknowledged, and that it has not been submitted for any other degree or professional qualification.

2. I declare that work submitted here is my own, other than for the case of journal publications or submissions which have been jointly authored, which are clearly indicated throughout the thesis. The contributions by myself and co-authors are stated below.

3. I confirm that I have appropriately cited the relevant sources where the work of others has been referred to.

The work presented in Chapter 5 was previously published in *Twin Research and Human Genetics* as *Longitudinal Reading Measures and Genome Imputation in the National Child Development Study: Prospects for Future Reading Research*. All authors are as follows: myself, Elinor C. Bridges; Dr N. William Rayner; Dr Hayley S. Mountford; the secondary supervisor of this thesis, Dr Timothy C. Bates; and the primary supervisor of this thesis, Dr Michelle Luciano. Conception of this study was carried out by myself, Michelle Luciano, and Timothy C. Bates. N. William Rayner carried out part of the genetic data preparation for this analysis, including determining genome build of genotyped data samples, and advised on genomic quality control methods and provided technical guidance. Hayley S. Mountford advised on genomic quality control methods, provided technical guidance, and assisted with preparation of the X.
chromosome data for analysis. The research paper was written by myself, and I carried out all other analyses. All authors provided feedback on the manuscript before publication.

The work presented in Chapter 7 has been submitted for publication in *The Journal of Pain (US Version)* as *Childhood Reading Ability and Pain in Childhood through to Midlife*. All authors are as follows: myself, Elinor C. Bridges; Dr Carole Torsney; the secondary supervisor of this thesis, Dr Timothy C. Bates; and the primary supervisor of this thesis, Dr Michelle Luciano.

Conception of this study was carried out by all authors. The research paper was written by myself, and I carried out all analyses. Feedback on the manuscript was provided by all authors before initial submission, and Dr Michelle Luciano and Dr Carole Torsney provided further feedback on receipt of comments from the reviewers and editor.

*Please note that since examination of this thesis, the work presented in Chapter 7 has been accepted and is in press with the following citation:* Bridges, E. C., Torsney, C., Bates, T. C., & Luciano, M. (2024). Childhood reading ability and pain in childhood through to midlife. *The Journal of Pain (US Version)*. In press. https://doi.org/10.1016/j.jpain.2024.03.014

Signed,

Elinor C. Bridges
A Genetic and Phenotypic Approach to Understanding Reading Ability and Associated Health and Lifestyle Outcomes in the National Child Development Study

Elinor C. Bridges

Submitted for the degree of Doctor of Philosophy

Psychology
School of Philosophy, Psychology, and Language Sciences
The University of Edinburgh

2023
Acknowledgements

I never anticipated that I would be working on my PhD during a pandemic, and I have so many people to thank for their support. To start, I’d like to thank my supervisory team, Dr Michelle Luciano and Professor Tim Bates, for their academic guidance and always insightful feedback. I’d particularly like to thank Michelle for always being supportive and helpful during the twists and turns of the thesis, right up to the end. In addition, I give my huge thanks to every collaborator mentioned in this thesis; Dr Hayley Mountford, Dr Will Rayner, and Dr Carole Torsney. I have learnt something from all of you. I would also like to thank every fellow student who showed me encouragement and support during the production of this thesis. That sense of community is so important.

This work would not have been possible without my family. I want to thank my parents for always being on the end of the phone, for their optimistic support and encouragement, their love and patience throughout, and most of all, for spending so many hours instilling a love of reading in me when I was young. I would not have made it this far without you. I would also like to thank my granddad, who always believed in me.

I give my endless thanks and gratitude towards my partner-in-PhD and in life, the newly titled Dr Matthew Grundy. You have been my constant support and inspiration throughout this process, and no one could have done it better. Your love, care, and reminders that I can do it have more value than I can put into words. Finally, I must thank my best friend. You sat with me during the long days and late nights, took me out on walks which helped me gain new perspective, made me laugh when I didn’t feel like it, and gave me a paw to hold during the tough moments. You are an excellent home-office companion, and most importantly, a very good boy.
Abstract

Reading ability is a polygenic trait, which is also influenced by environment. Despite the global social and economic cost of poor reading skills, molecular genetic research regarding reading ability as a quantitative trait is still in its infancy. Modern molecular methods have the potential to reveal much about such complex traits, including their genetic basis and how they relate to other traits. Chapter 1 explores the literature around reading ability, including how reading skills are defined, and the predictors and outcomes associated with them. Chapter 2 reviews the current state of knowledge on the genetics of reading ability, including studies from family, twin, and molecular genetic designs. This thesis has utilised the National Child Development Study (NCDS) to investigate longitudinal reading, the genetics of reading, and associations with other traits in detail. Chapter 3 reviews the cohort in more detail, including the data available and a review of existing reading research in the dataset. Chapter 4 outlines the genetics methods that are used to carry out the empirical research in the remainder of the thesis, including genotype imputation, genome-wide association studies, and linkage disequilibrium score based genetic correlations.

The NCDS is a UK based birth cohort study, which has been running since 1958. A subsample of NCDS participants have provided genetic data, however this genomic dataset has been poorly documented. In Chapter 5, I present an updated version of the NCDS genomic dataset, which has been imputed to the Haplotype Reference Consortium v1.1 reference panel, for use in future research (n = 6431). This contribution will allow other researchers to carry out analysis using modern molecular methods on a dataset imputed to a reference panel which has been shown to be more effective than the previous reference panel the NCDS was imputed to, the 1000 Genomes reference panel. The NCDS contains a series of different reading measures,
including objective tests, teacher ratings, and self-report measures, from childhood though to midlife. In Chapter 5, these measures have been analysed to develop a series of reliable and valid age-specific indices which measure reading ability quantitatively, with potential to be used in future reading research.

Chapter 6 uses reading composites as the phenotype outcome for a series of genetic analyses, which replicate several previous findings, showing that the reading composites generated in the NCDS, while not gold standard reading assessments, are still appropriate measures of reading ability. Genome-wide association analysis was conducted on reading ability composites derived at different ages through childhood to young adulthood. Downstream analysis showed the reading measures at different ages to be strongly genetically correlated with each other, with reading and language variables from other samples, and moderately genetically correlated with a series of additional traits related to social, educational, occupational and health outcomes, replicating several previous findings. Positive genetic correlations were found between reading and cognitive performance, educational attainment, and employment status, whereas negative genetic correlations were found between reading and occupational conditions, wellbeing measures, and health outcomes. This study replicated an earlier intriguing finding of a genetic correlation between reading and pain, which was then explored further by interrogating genetic correlations between reading and additional pain related variables, such as painkiller use.

Given the richness of the NCDS longitudinal data, in Chapter 7 I further explored the association between reading ability and pain in an effort to disentangle potential causal relationships between the two. Logistic and linear regression analyses were used to analyse the association between childhood reading ability and experiencing pain in childhood and adulthood. Several pain types were tested, including abdominal pain and headache in childhood, and back
pain, eye pain, headache, migraine, and rheumatism in adulthood. Numerous significant associations were discovered at multiple age points, with all significant associations indicating that better readers are less likely to experience pain. Further analysis indicated that socioeconomic status (SES) may act as a mediator for the relationship between reading and some pain types. Evidence of full mediation was found for the relationship between reading and back pain, while evidence regarding a mediation effect between reading and rheumatism was unclear. SES, measured by occupation, did not mediate the relationship between reading and eye pain, and the results indicated that reading ability and SES acted together to influence headache. I conducted post-hoc analysis in an attempt to further untangle the relationship between reading and back pain, including physical demands of the participant’s job as a covariate. The mediation effect remained, indicating that mediation was due to other elements of SES.

Chapter 8 explores the further implications of this research. This thesis shows that non-validated reading measures can be combined in order to generate a reliable and valid quantitative reading composite that is suitable for further research in reading, including genetic research. This is significant, as the NCDS is an exceptionally rich longitudinal dataset, which has the potential to be used to study gene by environment interactions in future reading research using molecular methods, an area with potential which is still in its infancy. This work has added to a small amount of existing molecular evidence which suggests that reading is genetically stable over time, with high genetic correlations found between reading at ages seven, eleven, sixteen, and an overall reading composite including measures from childhood through to age 33. Additionally, this thesis replicated previously identified genetic correlations between reading ability and health, including pain. Phenotypic analysis has indicated that, in some cases, the relationship is not due to SES, and therefore it is possible that pain and reading ability share a biological
mechanism; future research using multivariate genetic methods (expanding to other neurodevelopmental traits) are needed to identify these common biological causes. Finally, this thesis has shown that the richness of the NCDS dataset, combined with genetic information, proves it valuable for behaviour genetic reading research. With much still to discover about the behaviour genetics of reading ability and its transactions with the environment, this thesis will hopefully act as a starting point for further work in this area, which should include a thorough investigation of gene by environment correlation and interaction in reading ability.
Reading ability differs between people; some people have high levels of reading ability, some have low levels of reading ability, and some are in the middle. Reading ability is influenced by a person’s genetics and environment. Reading ability is not caused by one gene – it is a polygenic trait, which means that many genes have a small impact on how well a person reads. We are in the early stages of understanding how DNA influences reading ability, and we can learn more about this by studying the DNA and reading ability of a large number of people. This helps to identify overall trends.

The National Child Development study is a long-running research study that began in the UK in the year 1958. Every person who gave birth during a particular week was invited to take part. If they agreed, information was collected about them and their baby, such as health and lifestyle information. The babies were followed up every few years, and more information was collected, creating a huge resource for researchers to understand how people develop over time. The National Child Development Study contains a lot of information about reading ability of the people who participated, and also has DNA samples for many of the participants. The National Child Development Study has not previously been used to carry out reading research with the most up to date DNA methods.

In this thesis, I have used the information in the study to develop reading measures that give participants a value for their reading ability on a scale. This allows us to understand all levels of reading ability, rather than simply comparing those who have dyslexia and those who do not. The way the DNA data in the study was structured was outdated; I have updated this genetic dataset to make it a better resource for other researchers.

Reading ability has links with many other traits, such as mental health, socioeconomic status, and employment. This thesis investigates whether there is a connection between reading ability in childhood and pain in adult. I have found evidence for a relationship between reading ability and several types of pain. Further tests showed that, in some cases, this is because poor reading ability can contribute to low
socioeconomic status, and this in turn can cause pain. However, this was not the case for all pain types, suggesting that there might be a biological connection between reading ability and pain.

To investigate further, I carried out tests using the DNA data in the National Child Development Study. I used this genetic data to test for three things. First, I successfully used it to show that, even though participants were asked different questions about reading at ages seven, eleven, and sixteen, reading ability is influenced by similar genes at all of these ages. I showed that the reading measures in the study were genetically similar to reading measures used in other, larger studies, which suggests that the measures I have developed are truly measuring reading, rather than other, related skills or attributes. Finally, I was able to conform previous research with different groups showing genetic similarities between reading ability and pain. Together, the results regarding reading and pain indicate that, for some pain types, it is possible that the same genes are responsible for influencing reading ability and pain. This information could be useful in the future for understanding reading difficulties, or developing new drugs to treat pain. Finally, I found evidence of genetic similarity between reading ability and the taking of painkillers. This further highlights the importance of understanding this relationship, given the rising problem of painkiller addiction in many counties.
# Table of Contents

Declaration .................................................................................................................................................. 1
Acknowledgements ................................................................................................................................. 4
Abstract .................................................................................................................................................. 5
Lay Summary ........................................................................................................................................... 9
Contents .................................................................................................................................................. 11
Tables ..................................................................................................................................................... 13
Figures .................................................................................................................................................... 15

## Section One: Introduction and Literature Review .............................................................................. 16

Chapter 1: Defining Reading Ability ..................................................................................................... 17
  1.1. Thesis Overview .......................................................................................................................... 17
  1.2. Definition of Reading Ability .................................................................................................... 21
  1.3. Associations with Reading Ability ........................................................................................... 31
  1.4. Outcomes Associated with Reading Ability ............................................................................. 35
  1.5. Literature Critique and Contribution of this Thesis ................................................................. 39
  1.6. References ................................................................................................................................ 43

Chapter 2: Genetics of Reading Ability ................................................................................................. 68
  2.1. Introducing Reading as a Heritable Concept ............................................................................. 68
  2.2. Heritability of Reading Ability .................................................................................................. 70
  2.3. History of Molecular Methods .................................................................................................. 84
  2.4. Gene-Environment Interplay .................................................................................................... 89
  2.5. Literature Critique and Contribution of this Thesis ................................................................. 90
  2.6. References ................................................................................................................................ 93

## Section Two: Data and Methods ......................................................................................................... 111

Chapter 3: The National Child Development Study .............................................................................. 112
  3.1. National Child Development Study Overview ........................................................................ 112
  3.2. Genetic Data Collection in the NCDS ..................................................................................... 117
  3.3. Phenotypic Reading Research in the NCDS ......................................................................... 118
  3.4. Limitations of the NCDS ........................................................................................................ 122
  3.5. Benefits of using the NCDS for Reading Research .............................................................. 123
  3.6. References ................................................................................................................................ 124

Chapter 4: Molecular Genetic Methods ................................................................................................. 132
  4.1. Genotype Imputation .............................................................................................................. 132
  4.2. Genome Wide Association Studies and Meta-Analysis ......................................................... 134
  4.3. Linkage Disequilibrium Score Regression Genetic Correlations ....................................... 139
  4.5. References ................................................................................................................................ 142

## Section Three: Empirical Studies ...................................................................................................... 150

Chapter 5: Longitudinal Reading Measures and Genome Imputation in the National Child Development Study: Prospects for Future Reading Research ........................................... 151
  5.1. Published Work ...................................................................................................................... 151
## Tables

### Chapter 5

**Table 1:** Breakdown of number of participants sampled on each array in the NCDS after removal of exclusions and duplications, and number of SNPs sequenced in each dataset .................................. 156

**Table 2:** Number and overall percentage of SNPs that were removed from each array due to failure to pass quality control steps .......................................................... 158

**Table 3:** Number and overall percentage of individuals removed from each array due to failure to pass quality control steps .................................................................................................................. 159

**Table 4:** Percentage and number of SNPs with genomewide imputation $R^2$ scores greater than 0.8 for each array ............................................................................................................................. 159

**Table 5:** Descriptive statistics of continuous reading variables in the full sample and matched genotyped subsample ........................................................................................................ 159

**Table 6:** Principal component loadings and communality from a principal components analysis ................................................................................................................................. 160

**Table 7:** Interfactor correlations of the three principal components generated by principal components analysis of the full NCDS dataset .................................................................................................. 160

**Table 8:** Cronbach’s alpha to show reliability of sale of overall and age-specific reading composites in the full sample and subsample ................................................................................ 161

### Chapter 6

**Table 1:** Sample information about external summary statistics for a set of reading and language related variables .......................................................................................................................... 183

**Table 2:** Descriptive statistics relating to four reading composites in the NCDS ........................................................................................................................................................................ 188

**Table 3:** Number of individuals who reported reading difficulties since leaving school in the NCDS ................................................................................................................................................... 188

**Table 4:** The number of arrays available for analysis on each array after removing low quality variants and low MAF variants ................................................................................................................. 189

**Table 5:** The number of individuals analysed for each GWAS analysis and meta-analysis by array and age ........................................................................................................................................ 190

**Table 6:** Genetic correlations between NCDS reading composites at different ages ................................................................................................................................. 193

**Table 7:** Genetic correlations between reading composites in the NCDS and reading and language measures from other cohort studies ........................................................................................................ 195
Table 8: Genetic correlations between reading ability composites in the NCDS and non-specific pain complaints…………………………………………………………………………………………………203
Table 9: Genetic correlations between reading composites in the NCDS and chest pain and angina………………………………………………………………………………………………205

Chapter 7
Table 1: Number of individuals experiencing pain used for each analysis………………………………………247
Table 2: Results of a logistic regression estimating the odds of regularly experiencing pain at age 7 as predicted by a composite reading score for age 7 and covariates…………………………249
Table 3: Results of a logistic regression estimating the odds of regularly experiencing pain at age 11 as predicted by a composite reading score for age 7 and covariates……………………250
Table 4: Results of a logistic regression estimating the odds of regularly experiencing pain at age 16 as predicted by a composite reading score for age 7 and covariates……………………252
Table 5. Results of a logistic regression estimating the odds of regularly experiencing pain at age 23 as predicted by a composite reading score from ages 7 to 16 and covariates………………254
Table 6. Results of a logistic regression estimating the odds of regularly experiencing pain at age 33 as predicted by a composite reading score from ages 7 to 16 and covariates………………255
Table 7. Results of a logistic regression estimating the odds of regularly experiencing pain at age 42 as predicted by a composite reading score from ages 7 to 16 and covariates…………………256
Table 8. Results of a logistic regression estimating the odds of regularly experiencing pain at age 50 as predicted by a composite reading score from ages 7 to 16 and covariates………………257
Table 9: Results of a set of causal mediation analyses testing a mediation effect of SES between childhood reading ability and different pain types……………………………………260
Figures

Chapter 5

Figure 1: correlation heatmap between all reading variables in the full sample (A) and subsample (B) .................................................................160

Figure 2: Eigenvalues generated from a parallel analysis of principal components in the full NCDS sample suggesting that three components should be retained..........................160

Figure 3: correlation heat map of age-specific reading composites and reading measures in adulthood in the full sample (A) and subsample (B).................................161

Chapter 6

Figure 1: Kernel density curves showing distribution of four reading composite scores in the NCDS.................................................................187

Figure 2: A series of QQ plots produced in relation to GWAS for reading ability at four age points.................................................................191

Figure 3: Manhattan plots showing the outcome of a series of GWAS on reading ability.......192

Figure 4: Significant genetic correlations (P ≤ 0.05) between reading ability composite at age seven and a series of cognitive, educational, mental health and wellbeing, and occupational measures.................................................................198

Figure 5: Significant genetic correlations (P ≤ 0.05) between reading ability composite at age seven and a series of health, pain, and SES measures.........................................199

Chapter 7

Figure 1. Probability of experiencing different types of pain at age 7 as predicted by logistic regression of childhood composite reading score, with confidence intervals...............251

Figure 2. Probability of experiencing different types of pain at age 42 as predicted by logistic regression of childhood composite reading score, with confidence intervals. (color to be used on this figure) .................................................................259
SECTION ONE:

INTRODUCTION AND LITERATURE REVIEW
Chapter 1: Defining Reading Ability

1.1 Thesis Overview

Good reading skills are required in developed societies to carry out many fundamental tasks. Poor reading skills have a high global cost, both socially and economically (Katiyar, 2021), but reading limitations still show high prevalence in the UK, affecting approximately 1 in six adults (National Literacy Trust, n.d.). It has been shown that quantitative reading ability is a highly heritable trait, with a recent meta-analysis of twin-studies estimating heritability at 66% (Andreola et al., 2021), and a recent large molecular genetic meta-analysis estimated a common genetic variant heritability of 20-22% for word-reading, calculated using the LDSC method (Eising et al., 2022). However, little is known about the genetic underpinnings of reading as a quantitative trait, with most existing research focusing on reading difficulties as a binary trait. This thesis aims to add to the understanding of the genetic architecture of reading ability, and its association with other traits, through using a moderately sized genetic sample (n = 6431) obtained from the 1958 British Birth Cohort study, also known as the National Child Development Study (NCDS).

To begin with, this thesis explores the properties of a series of phenotypic reading measures in the NCDS at ages seven, 11, 16, 23, and 33. The reading measures involved rely predominately on self- and teacher- report of reading ability, rather than standardised reading tests, which are more typical for the measurement of reading ability in behaviour genetic studies (eg. Gialluisi et al., 2014; Eising et al., 2022; Luciano et al., 2013). The NCDS dataset does include two reading tests; the Southgate Word Reading Test, administered at age seven, and age-
adjusted reading comprehension tests, which are administered at ages 11 and 16 (Shepherd, 2012). This thesis demonstrates that it is possible to form a valid and reliable reading composite measure at ages 7, eleven, and sixteen, and an overall reading composite measure using measures at all recorded age points between 7 and 33, which can be used to inform further reading research, in both a genetic (Chapter 6) and phenotypic (Chapter 7) capacity.

Chapters 5 and 6 of the thesis seek to demonstrate, first of all, that these composites can be used to carry out further research. Chapter 5 shows the creation of a series of longitudinal reading composites that encompass a combination of objective and subjective reading ability, in order to successfully capture the full spectrum of reading variation. Chapter 5 also demonstrates the preparation and imputation of a moderately sized genetic dataset to the highest quality reference panel available for use in European populations, the Haplotype Reference Consortium reference panel (McCarthy et al., 2016). Chapter 6 uses reading composites to carry out a series of genetic analyses which indicate that these composites can be used to make valid inferences about the genetics of reading ability and associated outcomes. This includes estimating heritability of reading in the NCDS dataset, and showing that genetics of reading are stable throughout childhood and into adulthood, along with showing that reading in the NCDS is genetically correlated with a series of reading-related measures from independent samples. The Chapter also shows that reading in the NCDS is genetically associated with a series of additional outcomes, including replicating previous findings relating to cognitive ability, educational attainment, SES, occupational factors, and health related factors. In Chapter 7, reading composites have been used to investigate a potential phenotypic relationship between reading ability and pain. This thesis has sought to investigate whether an association exists between these two phenotypes, and what the causal factors in this association may be by testing for potential
mediation effects of socioeconomic status (SES), measured by occupation. This tests the hypothesis that the relationship between reading ability and pain is mediated by occupational conditions, in other words, that poorer reading ability increases likelihood of working in conditions which are not conducive to good health, such as manual work, which in turn cases pain (Smart et al., 2017). Chapter 7 shows that this can only be confirmed for back pain, and that in other cases that this is due to some other reason, perhaps other elements of SES or a shared biological pathway.

This thesis makes a contribution to the field of the genetics of reading ability, and knowledge of the health and lifestyle outcomes associated with reading ability. It has long been known that reading ability has a genetic component, however most studies using modern molecular methods are underpowered, particularly in the case of quantitative reading, which has been researched less than cases of reading disability (Eising et al., 2022). One reason for this is that measuring reading ability in large cohort studies can be difficult and time consuming when doing so using psychometric style batteries of tests (Eising et al., 2022). This study shows that it is possible to conduct both genetic and phenotypic research using a combination of well-chosen measures that include self- and teacher-report, reducing the labour required to carry out studies of reading and potentially opening up opportunities for other cohort studies with more subjective measures of reading to be included in meta-analyses. This is particularly significant in the case of molecular genetic studies, where existing research has been hampered by small sample sizes (Eising et al., 2022).

Alongside this, this thesis has added to a small amount of both phenotypic and genetic evidence that suggests that reading ability is relatively stable with age, and that few new genetic influences being acting on reading ability once acquisition has begun (eg. Eising et al., 2022;
Betjemann et al., 2007; Logan et al., 2013). This is significant as this information may be used to inform (1) future studies of reading, for which only childhood or adulthood measures are available and (2) future intervention designs for reading ability, in combination with further research.

This thesis also includes one of a small number of studies examining the relationship between continuous reading ability and health outcomes, specifically, measures of pain. This is a novel finding of this thesis that replicated across age and pain types, indicating that those with poorer reading ability in childhood are more likely to suffer pain in adulthood. This has several potential clinical applications. Primarily, now that we know this association exists, it gives even more impetus for reading programmes to ensure that all young people are able to attain a good standard of reading ability. Failure to do this may lead to poorer life outcomes, further disadvantaging individuals who are already likely to be struggling with challenges such as low-status occupation (Smart et al., 2017) and low SES, which is itself associated with poorer health outcomes (Ritchie & Bates, 2013; Kivimäki et al., 2020), potentially exacerbating incidence of pain (Prego-Domínguez et al., 2020) and entrenching pre-existing inequalities. Secondly, these findings indicate to healthcare providers that they may wish to be alert to complaints of pain from those individuals who have low reading ability.

In short, the core focus of this thesis is to use a rich longitudinal birth cohort study to improve our understanding of reading ability, and its environmental and genetic correlates. The results of this thesis have several potential practical applications, which may be used to inform future preventative and intervention techniques, in order to provide improved support for this group of individuals, who are likely already facing multiple disadvantages.
1.2. Definition of Reading Ability

In order to give the reader a full understanding of the significance of the work presented in this thesis, this chapter will review the literature on reading ability to present the current state of knowledge, as well as the significance of understanding reading ability. This will begin with a discussion of the definition of reading ability. This will include what the different elements of it are, and how reading is acquired. This will then be used to inform a discussion of different methods of identifying reading difficulties, and why it may be more useful from a research perspective to operationalise reading ability as a quantitative variable. The section will then go on to examine the significance of reading ability, and why understanding it is important.

1.2.1 Defining Reading and Reading Difficulties

The ability to read is not a singular process; mastery of other processes is required in order to be able to read. Perhaps the most famous model of reading, the Simple View of Reading, claims that there are two key underlying elements to reading; decoding and linguistic comprehension (Gough & Tunmer, 1986). The model indicates that a deficit in either one, or both, leads to reading difficulties (Gough & Tunmer, 1986). The exact nature of the relationship between decoding, linguistic comprehension, and reading comprehension is yet to be settled, and multiple operationalisations of this relationship have been proposed (Joshi & Aaron, 2000; Hagtvet, 2003), although meta-analysis shows decoding and reading comprehension are highly correlated, with linguistic comprehension mediating this relationship (García & Cain, 2014). With a view to providing some background on the component processes of reading, this section will begin with a brief overview of decoding and comprehension.
Decoding is the process by which the reader is able to recognise parts of words or words (Gough & Tunmer, 1986). The ability to decode is built upon several different cognitive processes. To begin with, the reader must be able to recognise the letters that make up the word or word sound (Byrne & Fielding-Barnsley, 1989). Research indicates that the most important underlying processes of decoding are based in phonological awareness (eg. Stanovich, 1982; Patel et al., 2004). This is the process by which individuals are able to identify written text, such as letters or parts of words, with their corresponding sounds, known as phonemes (Henry, 1993). The phonemes may then be combined sequentially to form words (Caravolas, 2017). The successful decoder also requires easy access to their own stored library of words, or lexicon, in order quickly match the text with a meaningful word (Perfetti, 1984). Additional traits, such as executive function, have also been associated with the ability to decode successfully (Haft et al., 2019; Ober et al., 2020). The development of decoding skills based on underlying processes is a longitudinal event, with research indicating that early phonological awareness can predict later decoding abilities (Hogan et al., 2005).

Decoding is a key element that is required for reading comprehension (Jong & van der Leij, 2002), however before reading comprehension can be achieved, linguistic comprehension is also required, which is the ability to understand language which is heard rather than read (Gough & Tunmer, 1986; de Jong & van der Leij, 2002). Linguistic comprehension may become more important as the ability to decode improves, and may be the primary driver of differences in reading ability in older children and stronger readers (Vellutino et al., 2007), and there is some empirical evidence to support this (eg. Garcia & Cain, 2014). Linguistic comprehension itself is underpinned by language skills, including knowledge of grammar and vocabulary (Hjetland et al., 2017; Brimo et al., 2017; de Jong & van der Leij, 2002; Rogde et al., 2019). Several other
cognitive skills have also been suggested as drivers of both linguistic and reading comprehension, such as working memory (Daneman & Merikle, 1996; Hjetland et al., 2019; Seigneuric & Ehrlich, 2005) and the ability to make inferences by connecting presented language with content that is not explicitly stated (Lepola et al., 2012; Tompkins et al., 2013; Cain & Oakhill, 1999). Background knowledge has been proposed as a contributory factor to linguistic comprehension, as it may assist with interpreting content and making inferences (Hjetland et al., 2017), and has been shown to be a contributor to reading comprehension (Smith et al., 2021). The ability to apply that real-world knowledge to the text itself is a separate, but related, element of reading comprehension (Bowyer-Crane & Snowling, 2011). As with decoding, the development of comprehension skills appears to be a longitudinal process, as one study of a sample at-risk for dyslexia has shown that oral language skills in early childhood predict reading comprehension at age 8, even when controlling for decoding at the word level (Hulme et al., 2015).

Word-reading, or decoding, difficulties were first recognised as a specific disorder called ‘word-blindness’ in Germany in 1877 (Kaussmaul, 1877, cited by Anderson & Meier-Hedde, 2001). First coined in 1883 (Berlin, 1883, cited by Kirby, 2020), the term ‘dyslexia’ began to be commonly used to describe reading difficulties in Britain in the 1960s, however to this day there is an ongoing social and scientific debate as to the true meaning and usefulness of the term (Kirby, 2020; Elliott, 2020). It has been suggested that the term dyslexia has been applied under the assumption that those with the condition are different to other poor readers, despite a lack of evidence that this is the case (Stanovich, 1994), although it should be noted that some researchers, such as Gough and Tunmer (1986), have argued that dyslexia is specifically an issue of lack of decoding skill.
Previously, dyslexia was diagnosed using a definition that relied upon not just reading ability, but Intelligence Quotient (IQ), defined as an individual’s general cognitive ability (Snowling et al., 2020). Under this definition, a diagnosis of dyslexia was made when the individual in question had lower reading ability than what would be expected based on their IQ score and environment (Snowling et al., 2020). The IQ-discrepant method of diagnosis is associated with several issues. Under this definition it is possible that children with low IQ who also had specific reading difficulties were not diagnosed, and therefore unable to access support that may have been otherwise available to them (Tanaka et al., 2011). Additionally, reading skills and the results of IQ tests are innately intertwined. Many IQ tests require children to read and answer written questions correctly, which is more challenging for individuals who have reading difficulties, and this means that those with low reading skills often score lower on IQ tests at least partially because of their poor reading skills, meaning a dyslexia diagnosis cannot be made under the discrepancy model as both the reading skill and the IQ skill would be interpreted as below average (Stanovich, 1991; Snowling et al., 2020). Some IQ tests can be read aloud to the participant; however it has been shown that those with reading difficulties may also have difficulty in comprehending oral language (Georgiou et al., 2021), meaning that this problem cannot be completely avoided. Aside from these critiques, it has been shown individuals with IQ-discrepant dyslexia do not appear to have unique phenotypic traits that separate them from those with general low-reading ability; this includes the fact that high- and low-IQ individuals with reading difficulty show similar deficits in phonological processing, which is considered a core feature of dyslexia (Stanovich, 1996; Tanaka et al., 2011; Stuebing et al., 2002; López & González, 2000; Siegel, 1993; Ramus, 2001). An alternative approach to IQ-discrepant method of diagnosis is one in which dyslexia is diagnosed when poor readers fail to respond to
appropriate intervention (Fletcher and Vaughn, 2009; Snowling, 2012). This is, however, not a perfect solution; it has been argued that this method causes an unnecessary delay in diagnosis, compared to simply assessing the reading ability of the individual and determining whether they lie in the low end of a quantitative distribution (Snowling et al., 2020).

The challenge of defining reading difficulties through the low end of a normal distribution is that it does not reveal any information about the cause of the problem; the definition is vague, and reflects a current debate regarding whether dyslexia is simply the low end of a normal distribution of reading ability or whether it is characterised by deficits in specific underlying traits, such as a phonological processing deficit (Snowling et al., 2020). Evidence for a causal relationship between phonological processing and dyslexia has been accumulating since the 1970s (Richardson, 1992), and the link continues to be consistently replicated today (eg. Moura et al., 2014; Smith-Spark et al., 2017; Peters et al., 2020). Measures of phonological awareness differ (McBride-Chang, 1995), but the concept of phonological deficit may be described as, “problems in recognizing all levels of linguistic structure” (Nittrouer et al., 2011, p.763). While it is appears generally accepted by many (eg. Roitsch & Watson; 2019, Peterson & Pennington, 2015) that dyslexia is defined through evidence of phonological deficit, it should be noted that since the discrepancy-definition was largely discredited, diagnosis of dyslexia has been more broadly based on poor general reading skills, which is likely to include poor readers who struggle with component processes other than phonological sills in addition to those with phonological deficits (Snowling et al., 2020; Saksida et al., 2016; White et al., 2006). In some instances, an arbitrary cut-off such as the bottom 10% of readers may be applied to define cases (e.g., Olson, 2002). The debate over which component processes, if any, should be considered in dyslexia diagnosis is yet to be settled. Due to the inherently functional nature of reading ability,
i.e., it is only useful for as long as the society one resides in requires the individual to be able to read, the author of this thesis prefers a definition of reading difficulties that encompasses all those at the lower end of the reading ability distribution, regardless of specific cognitive deficits, as the life outcomes associated with the difficulty are likely to be similar no matter the cause.

1.2.2. Reading Ability as a Quantitative Trait

It has been suggested that defining dyslexia by phonological problems may be robust but inadequate (Snowling et al., 2020), as other evidence indicates that while phonological processing deficits may be a key contributor to poor reading skills, it is not the only contributor (Pennington, 2006). It has been argued by some that there is no need for dyslexia diagnosis, that dyslexia does not differ from any other form of poor reading ability (Stanovich, 1994; Siegel, 1992), and that dyslexia is ultimately the lower end of a normal distribution of reading ability (Shaywitz et al., 1992). As a result, additional information may be gained from studying reading ability from a quantitative perspective, including the full continuum of abilities (Stanovich, 1994).

A large amount of existing genetics research on reading ability has been focused on those individuals with reading difficulties, likely for two reasons: (1) collecting samples of individuals with specific traits, in this case, poor reading is more resource efficient than collecting general population samples (Miles, 2013), and (2) because the functional nature of reading ability means that those who struggle to obtain an acceptable level of skill for normal functioning in society are of more concern and interest than those who span the rest of the distribution, from below average but adequate, through average, to exceptional.
While this is understandable, it is possible that focusing attention only on those with dyslexia is causing reading researchers to miss out on key pieces of information about reading ability due to the amount of information that is lost through operationalising reading ability as a binary trait. In some studies, even when quantitative reading is measured, reading ability/disability is categorised with a cut-off at the extreme low end, representing reading difficulties (Olson, 2002; Gialluisi et al., 2019; Gialluisi et al., 2021). However, this does not tell us how the predictors or outcomes associated with very poor readers are relevant to below average readers, average readers, or exceptional readers. Additionally, large population-based studies such as the NCDS may include general measures of reading, without specific reference to dyslexia (Power & Elliot, 2005). Failure to use resources such as the NCDS to examine reading ability would constitute a significant oversight, given the richness of the dataset and array of other associations that may be explored.

To complicate matters further, a diagnosis of dyslexia is considered a medical diagnosis which requires assessment by a trained professional (The British Dyslexia Association, n.d.). It has been reported that individuals of lower SES, which is itself associated with low reading ability, may face barriers to accessing medical care (Cookson et al., 2016). As a result, it may be more difficult for those who are more likely to struggle with reading to get a formal dyslexia diagnosis. Research using a large cohort study in England and Wales found an association between attending a fee-paying school and being diagnosed with dyslexia, and also found that high parental SES is a significant predictor of dyslexia diagnosis (Knight & Crick, 2021). The reasons for this are unclear, however it has been proposed that it may be due to additional financial or social parental resources for obtaining a diagnosis (Knight & Crick, 2021).
In light of these complex social factors, it is arguably the case that quantitative reading measures may be a more accurate method for assessing an individual’s level of reading ability over a case-control definition of dyslexia, as it is less likely to be affected by issues of differential diagnosis due to external factors. The NCDS does not contain any dyslexia related variables, however it does contain several variables assessing reading ability with reference to decoding and comprehension, and self- and teacher-reported reading difficulties. Given the limitations of research carried out on using dyslexia alone as an outcome, this provides an opportunity for reading research to be more focussed on understanding the reading ability continuum as a whole, rather than just one element of it, which may not be representative of the whole picture.

1.2.3. Epidemiology of Reading Difficulties

In modern societies, reading is a skill that is required for many crucial interactions with the surrounding world. This includes education, employment, navigation, paying for goods and services, communication, and using the internet. Despite this, dyslexia diagnosis is relatively common. Rates of estimation vary, but a recent large meta-analysis has estimated global prevalence in primary age children to be approximately 7.1% (Yang et al., 2022), although this value may be as high as 17.5% (Shaywitz et al., 1994) or as low as 3.5% (Di Folco et al., 2021; Barbiero et al., 2019). The true prevalence of reading difficulties is difficult to determine for several reasons; these include differential presentation across languages, lack of large-scale testing, and non-typical cases (Miles, 2004), and lack of a clear, universally accepted definition (Barbiero et al., 2019). Problems with reading ability present in many different parts of the world, across languages and cultures (eg. Stevenson et al., 1982; Peterson & Pennington, 2012),
although research indicates that the key deficits may differ across different orthographies (Goswami, 2002; Hadzibeganovic et al., 2010; Seymour et al., 2003; Diamanti et al., 2017).

In the UK, reading difficulties often become identifiable after the child has started primary school, which is when they are formally instructed on how to read (NHS, n.d.), and cases of reading difficulty usually emerge before the fourth grade in the USA, at which time children are approximately 9-10 years old (Leach et al., 2003). There is a lack of longitudinal research in this field that has measured reading ability form childhood through to adulthood, but existing research indicates that reading difficulties are fairly stable through childhood and adolescence (Catts et al., 2012), particularly in cases where there is evidence of phonological deficit (Peterson et al., 2014; Lefèvre et al., 2023). Not all early reading difficulties persist (Catts et al., 2012). For example, it has been found in an Australian longitudinal sample that only 44% of those with reading difficulties, defined as the low end of a quantitative reading test, at age 7-8 years continued to show evidence of difficulties when retested at age 13-14 years (Smart et al., 2001). It should also be noted that not all reading difficulties are obvious from early childhood; in some cases, reading difficulties do not become clear until the child is older (Leach et al., 2003). A modelling-based study which measured reading ability at intervals between 2nd and 10th grade and used latent transition analysis to assess how reading difficulties change over time identified latent classes for late emerging poor word reading, late emerging poor reading comprehension, and late emerging combined difficulties (Catts et al., 2012). This study estimated that poor reading skills may emerge ‘late’ in as much as 13.4% of the population in this age group (Catts et al., 2012). It is unclear if, in these cases, reading difficulties have always been present and have only had a substantive effect when the child has started encountering more complex material, or if new difficulties have emerged that did not exist before (Leach et al., 2003).
Prevalence of reading disability and quantitative reading ability differs across sex, with females generally having stronger reading skills than males (see Krafnick & Evans, 2019 and Logan and Johnston, 2010 for reviews), and more males scoring at the extreme low end of the reading distribution (Baye & Monseur, 2016). In a large sample of six year olds it was found that boys were more likely to exhibit poor phonological awareness (Lundberg et al., 2012), and in addition, boys are more likely to be diagnosed with dyslexia (Knight & Crick, 2021) and acquire reading skills more slowly (Wolf & Gow, 1986). Most existing reading research has not drawn a distinction between sex and gender, with both terms being used to refer to males and females (Granocchio et al., 2021; McGeown et al., 2011); this an area where further research may be useful in order to help determine how much of this difference is due to societal factors associated with gender.

1.2.4. The Importance of Understanding Reading Difficulties

Individuals with poor reading ability are more likely to face a multitude of challenges compared to their peers; briefly, they are more likely to suffer from mental health problems (Hunn et al., 2023), low income (McLaughlin et al., 2012), and their children are more likely to suffer from reading difficulties (Bonifacci et al., 2013), perpetuating a cycle of low SES and other difficulties (Ritchie & Bates, 2013). Reading difficulties may also create additional difficulties regarding the leisure activities that an individual can partake in, such as travel (Lamont et al., 2013), shopping, or ordering food from a menu (Tanner, 2009). Alongside these personal challenges, the cost to society is also high (Katiyar, 2021). Low reading ability can cause low educational attainment and low employment, which is expected to cost the economy in lost productivity (Moll et al., 2022). Reading ability intervention has been shown to be cost-effective (Moll et al., 2022), further highlighting the importance of understanding reading ability.
Building on our current understanding of reading ability and disability could, in the long term, help to improve some of the negative outcomes identified. Early intervention is important (Snowling, 2012); for example, one study has found that children with, or at risk for, reading difficulties who experienced a reading intervention in first or second grade showed greater improvement in their skills than those who received the intervention in third grade (Lovett et al., 2017). It was also shown that the benefits persisted over time, with the early intervention group showing a continued increased rate of improvement at a follow up assessment up to three years later. The long-term benefits of successful intervention could be incalculable, with the potential to interrupt some of the pathways leading to negative outcomes associated with poor reading ability. Addressing reading difficulties during childhood is a long-term solution with long-term benefits, but in order to properly manage reading difficulties, an understanding of reading skill across the whole continuum could add value.

1.3. Associations with Reading Ability

Given that reading ability has such a large effect on daily life, it is unsurprising that there is a vast number of traits that are associated with reading ability. This ranges from individual level traits such as personality and cognitive traits, all the way through to socioeconomic factors (Beaujean et al., 2011; Ritchie & Bates, 2013). Phenotypic research has uncovered a wealth of information covering associations with reading ability from infancy to adulthood. This section will review the current existing literature on these associations. Many factors associated with reading ability function as part of a life-cycle, with parents influencing their children, early life influencing adulthood, and so on. This section will work with a chronological pattern through a person’s life. It will begin with an overview of known phenotypic predictors of reading ability and reading ability acquisition. For information on genetic predictors of reading ability, see
Chapter 2. This will then move onto a review of traits which, while not necessarily causal of poor reading ability, are often found to co-occur with reading difficulties. Understanding these traits and how they associate with reading ability may reveal important links in how these traits are similar and how they differ.

1.3.1. Predictors of Reading Ability

Excluding genetics, which will be covered in Chapter 2, here I will present an overview of the most well-known predictors of reading ability and reading difficulties. Predictors of reading ability are wide ranging. Environmental prediction of reading ability begins during pregnancy. An association has been found between maternal cigarette smoking during pregnancy and decoding and comprehension skills (Micalizzi et al., 2021), although this may be explained by other confounding factors, including genetics (Micalizzi et al., 2021; Peixinho et al., 2022). Additional research from a large UK cohort study has indicated that maternal exposure to nicotine during pregnancy has a negative association with 6 out 7 tested reading-related outcomes, including reading speed, spelling, and real word reading (Cho et al., 2013). Risk of miscarriage, defined by hospitalization during pregnancy due to risk, has also been associated with dyslexia in the child (Mascheretti et al., 2013), and premature birth shows an association with low reading ability (Bowen et al., 2002; Kovachy et al., 2014; Aarnoudse-Moens et al., 2009). Early life factors also show a significant association with reading ability, with those of higher birth weight showing higher reading skill (Aarnoudse-Moens et al., 2009), and being breastfed in early life is associated with higher quantitative reading scores at age 10 (Oddy et al., 2011). It should be noted that all of these pre- and post-natal factors are also associated with SES, and so it is possible that the association is confounded by this (Norsker et al., 2012; Glinianaia et al., 2013; Oakley et al., 2013). There is some research to indicate being the first-
born sibling is associated with higher letter and word recognition-related reading skills (Lehmann et al., 2016). This extends to other cognitive traits, and while the mechanism is unclear, recent research suggests that the cause is likely to be found in the post-natal environment, and is unlikely due to biological differences (Isungset et al., 2022). Being right-handed is also associated with higher reading ability, possibly due to the effects of hemispheric lateralization (Abbondanza et al., 2023). In addition, a genetic correlation has been identified between dyslexia and showing no preference for left- or right-hand use (Doust et al., 2022).

As alluded to earlier, one consistent early predictor of reading ability is early speech and language skills, with children with poor speech also showing poorer scores on a reading comprehension test in the NCDS (Calnan & Richardson, 1977). A separate UK based cohort study has identified that speech and language concerns at age five are associated with poorer quantitative word-reading reading ability at age seven (Russell et al., 2016). Other longitudinal studies have shown that children with language impairments in kindergarten are more likely to score lower than one standard deviation below the mean on a quantitative reading measure in 4th Grade (Catts et al., 2000), and a twin study found that moderate positive correlation is present between language ability at age four and reading abilities at age seven (Harlaar et al., 2008). It should be noted, however, that in the latter study a substantial proportion of this covariance was found to be due to shared environmental effects, that is environmental elements that are common to both twins. This indicates that the association may be confounded by other, environmental variables. Difficulties with language have also been identified in those who read words normally, but have difficulties with comprehension (Nation et al., 2010). As these studies show, it has been consistently replicated that reading ability is associated with early language skills, and one review has found that there is a particular association between the roles of phoneme and
morpheme awareness and reading ability, with both skills influencing the development of good reading skills, and good reading skills in turn improving phoneme awareness (Duncan, 2018). Early vocabulary knowledge is also associated with later reading ability (Duff et al., 2015). SES and the home environment have also been consistently associated with reading ability; see Chapter 5 for more details.

Despite the existing evidence regarding the phenotypic predictors of reading ability, reading difficulties remain a large-scale challenge, and reading difficulties continue to persist. It has been shown that quantitative reading ability has a heritable component (e.g., Bates et al., 2006; Wadsworth et al., 2015; Eising et al., 2022), and so exploring the genetics of reading ability further will give us the ability to better understand the phenotypic associations we see, and to develop new solutions.

1.3.2. Co-Occurrence of Reading Ability with Other Traits

Dyslexia frequently co-occurs with other notable psychiatric and developmental traits. The most commonly reported co-occurring trait with dyslexia is ADHD, with co-occurrence rates estimated to be between 25-40% (Boada et al., 2012). There is also evidence to suggest that attention deficit hyperactivity disorder (ADHD) is associated with quantitative reading ability; one study has found that severity of ADHD symptoms was associated with lower scores on quantitative reading tests for decoding and comprehension between Grades 1 and 4 (Ehm et al., 2016), and an association between ADHD and poor reading comprehension has also been reported elsewhere (Parks et al., 2021). While ADHD and dyslexia are distinct disorders and causality is difficult to attribute, due to their nature, emerging research indicates that co-occurrence is likely due to shared genetics rather than one disorder acting to cause another (van Bergen et al., 2023, preprint). Research has shown that inattentive type ADHD shows greater
association with reading difficulties than hyperactive type ADHD (Willcutt & Pennington, 2004). While the shared aetiology of these two conditions is still somewhat unclear, it appears that genetics plays a substantial role, and this will be discussed further in Chapters 2 and 6.

Research has also suggested that Autism Spectrum Disorder (ASD) is associated with reading difficulties, with one study estimating that 11.7% of dyslexic individuals in their sample also showed signs of ASD (Brimo et al., 2021). In a sample of approximately 100 teenagers with ASD, it was found that ~10% showed word-reading skills that were significantly lower than what would be expected based on their IQ, and ~37% showed a comprehension discrepancy (Jones et al., 2009). One literature review concluded that autistic individuals are likely to perform better on word reading tasks than comprehension, and that comprehension may be low even if word-reading is not (Fernandes et al., 2015). The co-occurrence between ASD and reading difficulties may be due to difficulty with language, which is fairly common in autistic individuals (Nally et al., 2018). Alongside ADHD and ASD, poor reading abilities may also co-occur with mathematics difficulties (Landerl & Moll, 2010) and Developmental Coordination Disorder (DCD) (Brimo et al., 2021). Reading and maths difficulties have been found to co-occur more than would be expected by chance in more than one sample (Landerl & Moll, 2010; Dirks et al., 2008), and recent research has indicated that part of this co-occurrence may also be due to difficulties with language that contribute to both reading and maths difficulties (Snowling et al., 2021). The reasons for co-occurrence of dyslexia and DCD are less clear, however some research has implicated brain differences in this association (Nemmi et al., 2023).

1.4. Outcomes Associated with Reading Ability

The previous section covered common predictors and traits which co-occur with poor reading ability. This section will now cover the literature that explores life outcomes related to
reading ability and reading difficulties. This will cover important life outcomes such as educational attainment and SES outcomes. Finally, this section will explore the health and lifestyle outcomes that have been associated with reading ability, including healthy and risk-taking behaviours, mental health outcomes, physical health outcomes, and life expectancy. This section will also include a brief overview of health literacy, and how it may be related to reading ability, along with the differences between the two concepts.

1.4.1. Reading Ability and Education, Occupation, and SES

As may be assumed, it has been shown repeatedly across countries and cultures that stronger reading ability is associated with better educational outcomes and better performance in school (Nyarko et al., 2018; Massing & Schneider, 2017; Levlin et al., 2022). This is unsurprising, given that good reading skills are required in order to fully participate in most educational systems without adjustments. One longitudinal study has indicated that children identified as having reading difficulties at age 7-8 were less likely to complete their secondary school education, but no difference was found for likelihood of completing further qualifications beyond secondary school unless co-occurring with behavioural problems (Smart et al., 2017). Similarly, in a Canadian sample, reading ability at age 15 predicted whether students finished their high school education by 19 (Knighton & Bussière, 2006). Childhood reading comprehension also predicts later academic achievement; a small study found that those with poor comprehension skills at age nine were less likely to meet government targets on standardised testing at age eleven, and tentative evidence suggested that poor comprehenders may also struggle to meet government targets at age 16 (Ricketts et al., 2014). UK Students with dyslexia who attend university are less likely to complete their degree, and those who do
complete are less likely to achieve the highest grades, although it was noted in the same study that dyslexic students are able to succeed if appropriate support is provided (Richardson & Wydell, 2003).

In addition to this, reading problems have been associated with unemployment, however this has not always been replicated. In a small Norwegian sample of dyslexic individuals, rates of unemployment were higher than the national average (Undheim, 2003), and a longitudinal study using a New Zealand based sample indicated that poor childhood word-reading ability is associated with higher levels of unemployment in early adulthood (Caspi et al., 1998). Another study found that those with a reading disability were less likely to be receiving unemployment support than those with a maths disability (Aro et al., 2018). However, in a longitudinal Australian study, no association was found between childhood reading difficulties and unemployment in early adulthood (Smart et al., 2017). Although the relationship between reading and unemployment is not clear, research has shown that those with low reading skills are more likely to be in low-wage or low status jobs (Currie & Thomas, 1999; Smart et al., 2017). In addition to this, childhood reading ability is associated with SES more generally in adulthood; in the NCDS, it has been shown that low reading ability at age seven is associated with lower SES at age 42, indicating that outcomes related to reading ability can endure for many years (Ritchie & Bates, 2013). This series of disadvantage provides strong impetus for understanding how those with early reading difficulties can go on to achieve educational goals despite their additional challenges, as the consequences of failing to do so could be lifelong.

1.4.2. Reading Ability and Health and Lifestyle Outcomes

Reading ability has been associated with several mental health and psychiatric outcomes. For example, studies have shown an association between poor reading ability or dyslexia and
sub-clinical symptoms of depression and depression diagnosis in children (Maughan et al., 2003), in adolescents (Kiuru et al., 2011; Fröjd et al., 2007), and in adults (Undheim et al., 2011; Boetsch et al., 1996). Further to this, there is evidence that poor reading ability is also associated with anxiety symptoms (Undheim, 2002; Grills-Taquechel et al., 2011) and suicidal ideation (Boetsch et al., 1996), and self-harm and suicide attempts in a sample of homeless dyslexic individuals (Macdonald et al., 2016). The mechanism through which these associations act is not always clear, however there is evidence that learning difficulties can lead to symptoms of depression in students via feelings of inadequacy (Kiuru et al., 2011). Some evidence has been found to indicate that the relationship between reading ability and anxiety symptoms is bidirectional, and additionally, dependent on the type of reading skill tested and the type of anxiety symptoms (Grills-Taquechel et al., 2011). Others have argued that poor reading skill can be a contributor to anxiety (Carroll et al., 2005). Additionally, there is some evidence that genetics may be significant in some of these associations; see Chapter 2 for more details.

Poor reading ability has also been associated with a series of poor physical health outcomes, although it should be noted that most research in this area focuses on ability to read health-based information, which may be referred to as health literacy. Health literacy is a heterogeneous concept which may be measured in multiple ways, such as through simple quantitative reading tests, reading tests based on medical terms, or more complex tests which include other content, such as comprehension (DeWalt et al., 2004). Due to this, it is difficult to state whether it is word reading ability specifically which is associated with health outcomes, and where other factors may be at play. Reviews have found that, generally speaking, poor health literacy is associated with several poor outcomes, including lack of health knowledge, participating in routine preventative health interventions such as vaccinations and screenings,
and adherence to treatment (see DeWalt et al., 2004). Those with poor health literacy are 2.19 times more likely to report poor general health (Baker et al., 1997). It has been found that low reading ability, tested through a quantitative based reading test of medical-related terms, puts adults at higher risk of non-attendance at medical appointments (Miller-Matero et al., 2016).

1.5. Literature Critique and Contribution of this Thesis

This literature review has broadly covered many of the key elements of our knowledge of reading ability, including its relation to dyslexia, predictors of reading ability, outcomes of reading ability, and co-occurring traits. However, this review is not able to provide a completely comprehensive assessment of what reading ability means, both literally and in terms of value to the individual, as there is still so much left to be discovered about this key trait.

A key limitation of our knowledge of reading ability is a lack of whole-life studies that study reading ability and associated traits in detail from birth to old age. In order to more fully understand the predictors and consequences of reading ability, research must be conducted on high quality, large, longitudinal birth cohorts which contain a wide array of properly documented variables. This literature review has demonstrated that the range of traits associated with reading ability is large, and the range of known associations is only likely to keep increasing as more research is carried out. Many studies, even those with samples that include quantitative reading measures, focus on a case-control definition of reading difficulties, meaning that we lack knowledge of what the difference, if any, may be between the predictors and outcomes of an average reader versus an exceptional reader. There are two ways to rectify this: (1) to increase the collection of high quality reading data from current and future cohort studies, and (2) to make better use of existing cohort studies in an attempt to understand these associations more fully.
Full life course longitudinal studies are the only methodology that can allow us to truly understand the lifelong longitudinal associations between reading ability and other variables. Reading skill must be measured regularly, as while this review has indicated that reading ability tends to crystallise in early adulthood, there are exceptions to this rule and these cases can be used to help us understand how important age of resolution for reading difficulties is when considering further life outcomes.

While it is evident that there are many traits, environmental conditions, and developmental disorders associated with reading ability, causality has often not been properly established. This means that, while we know associations exist, we do not truly understand them. This is another limitation which can be addressed through the use of large, population based cohort studies with varied repeated measures across ages. It is clear that the full extent of the association between reading abilities and other traits, even well researched traits, is not understood.

This thesis aims to build on both the phenotypic and genetic understanding we currently hold of reading ability, and its associations with health and lifestyle outcomes. As mentioned previously, studies of reading ability are limited by the data available, and in particular due to the labour-intensive practice of measuring reading ability more objectively through test batteries (Eising et al., 2022). This thesis uses a longitudinal birth cohort study, the National Child Development Study, to carry out reading research. The NCDS contains only one psychometric measure of word reading ability, and one reading comprehension test, which is age-adjusted for the two age-points at which it is delivered (Shepherd, 2012). The remaining variables are a series of self- and teacher-report variables, which are available at ages seven, 11, 16, 23, and 33. This thesis shows that it is possible to use subjective measures of reading ability, alongside more
objective tests, to measure reading ability accurately from early childhood through to adulthood. In a field of limited large scale data, this itself is a valuable contribution to the field, as it provides proof of concept for use of the NCDS in future longitudinal reading research. Despite most researchers working off the assumption that reading ability is stable, there is little empirical evidence to support this viewpoint, particularly beyond childhood and into adolescence and adulthood. This thesis will show that reading ability remains stable from ages seven to 33, using a series of reading measures from the NCDS.

As mentioned above, the full range of variables that are associated with reading ability are not known. There is some evidence to suggest relationships between reading ability and health and health related behaviours, but there is a shortage of research into the possible reasoning behind this association, and whether this applies in cases of word-reading ability rather than the broader concept of health literacy. This thesis presents the first known study testing for an association between childhood reading ability and pain in adulthood, and goes one step further to test potential causal mechanisms by testing for a mediation effect of SES. This has the potential to alert primary care physicians to potential additional health issues that may be experienced by poor readers, and additionally provides even further impetus for ensuring that all young people are given the support required to help them to achieve an acceptable standard of reading. This finding has the potential to prevent unnecessary hardship, if further research is able to confirm the causal mechanisms in action. This would have the added benefit of potentially avoiding unnecessary prescription of opioids, which has multiple associated risks (Sullivan & Howe, 2013).

This chapter has given an overview about what we know about reading ability, including predictors, outcomes, and co-occurring traits. While this information is valuable in and of itself,
there is something missing which has been alluded to throughout this chapter; it is not possible to truly understand the cause of variation in reading ability without considering its genetic component. It has been repeatedly shown that quantitative reading ability is a heritable trait, (e.g., Eising et al., 2022), therefore an understanding of the genetics of the trait is required in order to understand some of the variation that we see. In addition to this, some of the associations between reading ability and other traits that have been discussed in this chapter have a genetic component, for example, ADHD (Daucourt et al., 2019). There is a reasonably large body of research covering the genetics of reading, however there are relatively few studies covering the molecular genetics of reading using modern methods and adequate sample sizes (Eising et al., 2022). The next chapter will explore the current state of knowledge of the genetics of reading ability, in order to place the genetics research included in this thesis into its broader context.
1.6. References


https://doi.org/10.3758/bf03214546

https://doi.org/10.1111/j.1525-1497.2004.40153.x

https://doi.org/10.1080/10888438.2021.1998067

https://doi.org/10.1080/10888438.2017.1338291

https://doi.org/10.1177/0022219408321128

42 genome-wide significant loci associated with dyslexia. *Nature Genetics, 54*(11), 1621–1629. https://doi.org/10.1038/s41588-022-01192-y


Levlin, M., Okonski, L., & Sullivan, K. P. H. (2022). Educational attainment in Swedish (L1) and English (L2) for students with reading difficulties: a longitudinal case study from


https://doi.org/10.1037/dev0000184


Tanner, K. (2009). Adult dyslexia and the ‘conundrum of failure.’ *Disability & Society, 24*(6), 785–797. [https://doi.org/10.1080/09687590903160274](https://doi.org/10.1080/09687590903160274)


Chapter 2: Genetics of Reading Ability

2.1. Introducing Reading as a Heritable Concept

Chapter 1 gave an overview of the key existing phenotypic research regarding reading ability, its definition, and how it is related to life outcomes. In Chapter 2, the key focus will be the contribution of genetics to reading ability. This chapter will open with a brief summary of the current understanding of the genetics of reading ability, in other words, the assumptions about the genetics of reading ability that will underlie this thesis. This will be followed by a history of genetic reading research, starting with twin and family studies, from which most early insights into the genetics of reading were gained. More recent findings which have been reported through the use of molecular genetic methods will then be discussed. This includes heritability of reading, a brief discussion of what genetics studies have been able to reveal about the genetic stability of reading ability, and what is known about the genetic overlap between reading ability and other traits. There will be a brief overview covering the contribution that genetics of reading research has made, including key findings from the two largest molecular genetics studies of reading to date, and an introduction to gene-environment interactions in reading research. Finally, this section will discuss how this thesis contributes to the overall body of literature in this area.

2.1.1. Reading as a Polygenic Trait

Historically, it was believed that most inherited traits followed a Mendelian pattern of inheritance, meaning one or two genes would express different alleles. While some disorders do follow this pattern, genetic transmission of many common traits is more complex than this (Chakravarti, 2021), and it was not clear to scientists in the late 19th and early 20th century how
family resemblance in continuous traits could be accounted for under a model of Mendelian inheritance (Olby, 1987). In response to this apparent impasse, a theoretical basis for complex inheritance of traits was developed by Fisher (1918). In 1918, Fisher proposed the *infinitesimal model*, under which multiple genes could follow a pattern of Mendelian inheritance with additive effects, allowing the expression of continuous phenotypic traits. Whereas genetics was initially considered in the context of observable (i.e., physical traits), skills, cognitive abilities, and behaviours can also be inherited (Plomin et al., 1994). Statistical genetics studies have confirmed that many psychiatric traits, including cognitive traits, follow a mode of inheritance known as *polygenic inheritance* (Purcell et al., 2009; Davies et al., 2011). Polygenic inheritance occurs when many variants across the genome have a small influence on a trait, and collectively this influence is additive (Fullerton et al., 2019), just as was proposed by Fisher (1918) long before molecular methods were able to show this. Modern statistical genetics methods have been used to confirm that reading ability is one example of a polygenic trait (e.g., Meaburn et al., 2007). This means that that there are not one or two ‘reading genes’; instead, we can expect to see many genetic variants hold a small additive influence on overall reading ability. This is also the case for several related traits, including general cognitive ability (Davies et al., 2011), ADHD (Groen-Blokhuis et al., 2014), and Specific Language Impairment (Reader et al., 2014). It should be noted that, while for the majority of cases reading appears to be highly polygenic with influential variants of small effect size, there is evidence of exceptions to this, in that some small family pedigree studies indicate the presence of a larger causal genetic effect for dyslexia which may be family specific (Bishop, 2015; Paracchini, 2011). The majority of modern research indicates that, for almost all variants influencing reading, effect sizes are small (Paracchini, 2011).
2.1.2. Reading as a Heritable Trait

Reading ability is heritable, regardless of the measure used; cases of high ability in word-reading (Friend et al., 2009), of dyslexia (Erbeli et al., 2022), and the distribution of quantitative reading ability in decoding and comprehension (Andreola et al., 2021) all show a substantial heritable component. Heritability is a sample or population level concept, and describes the proportion of variance in a trait, in this case reading ability, which is due to genetic factors (van Dijk et al., 2022). Genetics of reading ability are indicative rather than deterministic; an individual who has a low genetic propensity to read well may be able to overcome those challenges and still become a strong reader, whereas a child with a genetic propensity to read well may have poor instruction or substandard living conditions and struggle with their reading (van Dijk et al., 2022).

2.2. Heritability of Reading Ability

Studies of the families of those with reading difficulties showed that relatives of those with reading problems were likely to face reading difficulties themselves (Foch et al., 1961; Finucci et al., 1976; Wolff & Melngailis, 1994). Modern research continues to replicate these findings, with one study showing that 42% of a sample of children selected for family history of dyslexia met the threshold for reading difficulties themselves, on a composite score of word reading and spelling (Snowling et al., 2007). In addition to this, those children with a family history of dyslexia who did not meet the threshold for reading difficulties performed worse on reading related skills, such as reading fluency, at age 8 than a control sample, indicating that sub-clinical reading difficulties are heritable (Snowling et al., 2007) and displaying the value of measuring quantitative reading ability. Another longitudinal study found that those in a sample
with family history of reading problems achieved lower scores on reading fluency and spelling tests than those in a control group (Lohvansuu et al., 2021).

Early behaviour genetic research into reading ability was predominately conducted using various forms of family studies. The majority of these studies take advantage of the fact that twins offer a unique opportunity for studying heritable traits (Boomsma et al., 2002). Twins provide the opportunity to perform several types of natural experiments to help us understand heritability of traits, including reading. This section will cover the existing research that has demonstrated that reading ability is a heritable trait, beginning with twin and family studies before moving onto molecular methods.

2.2.1. Twin and Family Studies of Reading Ability

There are several types of family-based study which have been used to demonstrate that reading has a heritable component; one of the most commonly used is the twin study. Twin based samples offer a unique natural experiment which allows researchers to assess whether a trait is influenced by heritable factors (Boomsma et al., 2002). This is because monozygotic (MZ) twins are genetically identical, meaning that, theoretically, any differences between the two individuals in a twin pair must be due to environmental factors. Dizygotic (DZ) twins share approximately half of their segregating genes with each other, so when MZ twins are more similar to each other on the trait than DZ twins, it indicates that the trait has heritable influences (Boomsma et al., 2002). This can be taken advantage of in different ways. The classical twin design, also known as the ACE model, refers to the three variance components that the ACE model measures: (A) additive genetic variance, (C) common environment, and (E) unique environment (Zyphur et al., 2013). This can be used to assess the proportion of variance of a trait which is due to each of the three elements above.
Many twin and family studies have been conducted on reading and reading related traits; it is outside the scope of this thesis to conduct a full literature review of all such studies, however here I will lay out some key findings of interest from this research area.

Early twin studies generally had modest sample sizes, and began by establishing that reading ability is heritable (eg. DeFries et al., 1987). For example, an early twin study of quantitative reading ability in a sample of adolescents estimate a heritability of 18-33%, depending on the specific trait measured, with the highest heritability for a reading comprehension measure (Stevenson et al., 1987). There are some key longitudinal twin samples which have been used in order to demonstrate much of the current knowledge on the behaviour genetic influences on reading, and these will be referred to frequently in the remainder of the chapter. As such, the following section will provide a brief overview of three of these key studies.

**Colorado Learning Disabilities Research Center (CLDRC).** One of the most significant twin samples to be used for reading research is the Colorado-based twin sample. The Colorado Learning Disabilities Research Center (CLDRC) sample of twins were identified though a combination of school and birth records in Colorado, USA (Wadsworth et al., 2015). Twins pairs were invited to join the sample if either one or both twins showed evidence of reading difficulties or ADHD based on school records (Wadsworth et al., 2015; Gayán & Olson, 2003). Tests conducted at the first wave included measures of word-reading, comprehension, and spelling, from which a discriminant composite score was generated and applied (Wadsworth et al., 2015). Specific underlying reading skills, such as phoneme awareness, phonological decoding, and orthographic coding were also tested (Gayán & Olson, 2003). The age range for this sample is large, with individuals aged approximately 8 to 18 being invited to participate over

72
the course of several years in the 1990s and 2000s (Wadsworth et al., 2015; Willcutt et al., 2019; Wadsworth et al., 2007) It is a large sample, with one study using the data including approximately 770 pairs of twins for which data was collected between 1996 and 2009 (Wadsworth et al., 2015). A control sample of twin pairs without formal reading problems was also selected (Wadsworth et al., 2015; Gayán & Olson, 2003). Subsamples of the group have been followed up approximately 5 years after their initial involvement (Wadsworth et al., 2007; Wadsworth et al., 2015). Similar reading measures were used at follow-up (Wadsworth et al., 2007; Wadsworth et al., 2015).

**Twins Early Development Study (TEDS).** The Twins Early Development Study (TEDS) is a longitudinal study of twins born in England and Wales between 1994-1996 (Lockhart et al., 2023). Twins were identified through birth records, and the first wave of data collection occurred at 18 months old, with several waves of data collection since (Lockhart et al., 2023). A total of 12,939 twin pairs participated in the first wave (33% MZ), and sample sizes remain high, with data for approximately 4500 twin pairs collected during the most recent waves (Lockhart et al., 2023). The focus of TEDS has shifted over the years, with early childhood waves focusing on cognition, including reading and language variables, and adulthood data collection emphasising collection of mental health related measures (Haworth et al, 2012; Lockhart et al., 2023). Language-related variables were measured at ages two, three and four, and in a sub-sample at age four and a half years (Haworth et al., 2012). Language variables measured include vocabulary and grammar (Dale et al., 1998; Spinath et al., 2004; Asbury et al., 2005). More in-depth language data was collected on the sub-sample, including phonological awareness (Hayiou-Thomas et al., 2006). Language skill data was also collected at ages 14 and 16 (Haworth et al., 2012). Reading data was collected at ages four, seven, nine, 10, 12, and 16.
Early reading measures included word-reading (Arden & Plomin, 2006), and as children moved through school general, curriculum-based measures of reading ability were used (Harlaar et al., 2007). These focus on decoding in the earlier years, moving to comprehension as the children get older and higher demands are placed on their ability to read and understand in school (Harlaar et al., 2007).

**International Longitudinal Twin Study (ILTS).** The International Longitudinal Twin Study (ILTS) began when the participants, all sets of twins, were four years old, with the intention of capturing and measuring reading-related processes before the start of formal reading instruction (Byrne et al., 2002). The initial round of data collection consisted of 250 twin pairs, and included twin pairs from Australia (46 MZ pairs, 27 DZ pairs), Norway (16 MZ pairs, 19 DZ pairs), and the USA (63 MZ pairs, 79 DZ pairs). USA and Australian samples were collected from local twin registries, with the USA sample being derived from the Colorado Twin Registry, while the Norwegian sample was collected from a local birth registry. Twin pairs from Sweden, acquired through a local birth registry, were later added to the sample (Samuelsson et al., 2005).

Initial data collection included a wide range of reading-related skills, including phonological awareness, verbal working memory, letter knowledge, and word recognition (Byrne et al., 2002). Information about the HLE was also collected. The sample have been the subject of multiple follow up studies, including at the end of pre-school (Samuelsson et al., 2005), kindergarten (Byrne et al., 2005), first grade (Byrne et al., 2007), and second grade (Byrne et al., 2009), with reading-related variables measuring more complex skills, such as word-reading fluency and comprehension, as the participants aged (Byrne et al., 2009). By age 7-8, or Grade 2, the sample contained over 600 twin pairs across the four countries (303 MZ twins, 312 DZ twins) (Byrne et al., 2009).
The above twin studies, along with several others, have continued to replicate the finding that reading related skills are heritable. Data from CLDRC has shown that decoding (Gayán & Olson, 2001; Gayán & Olson, 2003) and comprehension (Betjemann et al., 2007; Keenan et al., 2006) are heritable. Similarly, ILTS data has confirmed that measures relating to phonological awareness are heritable at the pre-school age (Byrne et al., 2002), and measures relating to both decoding and comprehension are heritable at Grades 2 and 4 (Olson et al., 2011; Byrne et al., 2009). Work in other samples has indicated a genetic influence in dyslexia and comprehension (Hensler et al., 2010). Modern twin studies have confirmed that results can differ depending on the type of reading measure used, and data from CLDRC indicates that the reading measure used in a twin study can influence the estimated proportions of variance explained by the three elements of the ACE model (Christopher et al., 2013). Christopher and colleagues (2013) also note the importance that the individual sample may play in this variation, arguing that more twin samples are needed to confirm results.

Twin studies have been used to compare the heritability of different reading related skills. A study using the CLDRC sample found that heritability differed depending on the deficit; a phonological deficit was more heritable (h = 0.67) than difficulties in reading irregular words which reflect lexical processes (h = 0.31) (Castles et al., 1999). Similarly, another study with an unspecified sample found that phonological deficits are more heritable than orthographic problems, measured by asking participants to distinguish between homophones (Olson et al., 1989). Another CLDRC-based study of 19 reading related traits showed that genetics was a significant contributor to all but one trait, named as silent phonological decoding latency, with.
genetic heritability of the remaining traits showing a large range, from 46% to 72% (Gayán & Olson, 2001). This implies that it is important to explore as many reading phenotypes as possible, as it may be that the genetics of one measure of reading ability or disability differ from another. Interestingly, one Colorado-based twin study found a difference in heritability of a composite of quantitative reading ability, based on comprehension, recognition, and spelling, between individuals with and without reading difficulties, with reading being more heritable in those from the reading difficulties group (DeFries & Alarcón, 1996). The authors of the study argue that these findings are indicative of at least one gene with a large effect size that specifically influences reading difficulties, rather than general ability (DeFries & Alarcón, 1996). However, genomic studies to date do not support this hypothesis, instead suggesting many variants of small effect sizes across the distribution (Doust et al., 2022; Eising et al., 2022).

Specifically, there is some research which as able to shed some light on the genetic aetiologies of word-reading and comprehension, and how they relate to one another, although twin research exploring this specific research question is relatively scarce as many studies using data from both decoding and comprehension combine this data into one measure (Keenan et al., 2006, p.78). Data from ILTS indicates that the genetic similarity between word-reading and reading comprehension is very high in Grades 1 and 2, and it has been suggested that this is because word-reading is required for reading comprehension (Byrne et al., 2007; Byrne et al., 2009). Research in the same data has also indicated that the genetics of earlier phonological awareness show a direct contribution to the genetics of later reading comprehension (Byrne et al., 2009). Additional research using the Colorado based twin sample has also indicated that the genetics of word-reading contribute to the genetics of reading comprehension and to the genetics of listening comprehension (Keenan et al., 2006). In addition to this, the same study
used Cholesky modelling to show that there are also independent genetic effects on reading comprehension aside from those caused by word-reading, indicating that while the two phenotypes are related and have some genetic overlap, there is also some level of distinction between them (Keenan et al., 2006).

Aside from twin studies, family-based experimental designs can be extended by considering cases of adoption. Adopted children are unique in that they do not share DNA with their parents or siblings, therefore any similarities on cognitive traits must be environmental in nature (Cadoret, 1995). One combined twin and adoption study of early readers used the ACE modelling method to assess the heritability of four phonology and decoding related measures in the adoptive sample; all showed a significant heritable component, with estimated heritability ranging from 25% to 69%, depending on trait (Petrill et al., 2006). Some of the results in the study regarding shared environment supported earlier results conducted in a the ILTS sample, however, others did not (Petrill et al., 2006; Byrne et al., 2002). The adoptive study found that, for phonological awareness, word attack, and letter identification, shared environment and non-shared environment were both significant contributors of variance, and this was replicated in their own twin sample (Petrill et al., 2006). However, as the authors point out, this was not the case in ILTS, for which the effect of shared environment did not reach the significance level (Petrill et al., 2006; Byrne et al., 2002). The authors suggest that this is due not to differences in study design, as their own twin study replicated their adoption results, but because of sampling error. Another adoption study found that there was no association between the word-reading ability of parents and adoptive children, whereas quantitative word-recognition ability of parents and their biological children was correlated, providing further evidence that reading as a fairly strong heritable component (Wadsworth et al., 2002). While adoption studies can be a useful tool
for understanding the genetics of reading and other heritable quantitative traits, they also have limitations; these are explored in Section 2.4.2.

2.2.2. Genetic Architecture of Reading Ability

Multivariate twin studies have shown that there is likely a shared genetic architecture between reading ability and several cognition-based traits. For example, one study using the TEDS sample used multivariate twin modelling to show that language difficulties at age four are predictive of later difficulties in comprehension and decoding, and that genetics are a partial mediator of this relationship (Hayiou-Thomas, 2008). This is in line with phenotypic evidence, as discussed in Chapter One. Multivariate twin studies conducted in TEDS and CLDRC, among others, have also indicated a substantial genetic overlap between intelligence and both word-reading ability and decoding skills (Harlaar, Hayiou-Thomas, et al., 2005; Alarcón & DeFries, 1997; Brooks et al., 1990; Haworth et al., 2009; Cardon et al., 1990); TEDS data indicates a genetic correlation of 0.50 between a reading composite, based on decoding and national curriculum based measures, and intelligence, although it should be noted that the same study identified that there are also unique genetic influences that are specific to reading, with just over half of the variance in reading estimated to be caused by independent genetic effects (Harlaar, Hayiou-Thomas, et al., 2005). Of the total calculated heritability of reading, approximately three quarters was estimated to be independent of the genetics of cognitive ability (Harlaar, Hayiou-Thomas et al., 2005). Research conducted in the CLDRC sample has supported this further, through findings indicating that while there is genetic similarity between cognitive ability and reading comprehension, listening comprehension, and word reading, independent genetic effects on reading variables remain when controlling for cognitive ability (Keenan et al., 2006). This includes shared variance between the genetics of listening and reading comprehension that exists
independently of intelligence, indicating that while the genetics of cognitive ability and those of reading show some commonality, there are also genetic elements that are specific to comprehension processes outside of general cognitive ability (Keenan et al., 2006). CLDRC data has indicated that quantitative reading ability, comprising of decoding, comprehension, and spelling, and intelligence have greater shared heritability in those with reading difficulties than those without, suggesting that there are perhaps a small number of variants with larger effect sizes which are common to both reading and intelligence (Alarcón & DeFries, 1997). In summary, then, twin data has revealed that there is substantial genetic overlap between the traits of general cognitive ability and reading ability, however it has also shown that there are genetic effects on reading ability that are independent of intelligence. TEDS data has indicated that phenotypic reading skills, measured using composites of different reading measures including both decoding and comprehension, may contribute to phenotypic intelligence (Harlaar, Hayiou-Thomas et al., 2005; Ritchie et al., 2014). If this extends to the genetics of these traits, we may see an example of vertical pleiotropy, in which genetics contribute to trait one, and trait one influences trait two, meaning that the genetics of trait one are detected in trait two (Harlaar, Hayiou-Thomas et al.; Verbanck et al., 2018, see Section 2.3 in this chapter for more). As a result, research on the genetics of reading which does not specifically model for the genetics of general cognitive ability must be interpreted in the knowledge that the genetics of reading are also likely to be capturing some of that of cognitive ability.

In addition to general cognitive ability, genetic similarities have also been found regarding reading and maths difficulties. Several multivariate twin studies, including TEDS, have suggested that reading and maths disorders have a shared genetic heritability, with bivariate heritability estimates ranging from 0.26 to 0.76, and consistent moderate to strong estimated
genetic correlations between the traits have been reported (Light & DeFries, 1995; Hart et al., 2009; Markowitz et al., 2005; Kovas et al., 2005; Kovas et al., 2007; Haworth et al., 2009). Genetic similarities have also been reported between the quantitatively measured reading and maths traits in the CLDRC sample, including both word reading and comprehension (Willcutt et al., 2019). Interestingly, TEDS data also indicates that the shared environmental influence on reading and maths is highly correlated (Kovas et al., 2005; Kovas et al., 2007), and it has been suggested that this may be due to highly influential overall environmental factor, such as SES (Haworth & Plomin, 2010). Further information on shared genetic architecture is available in Chapter 6.

Similarly, there is a rather large body of work indicating that reading problems show a shared genetic basis with Attention Deficit Hyperactivity Disorder (ADHD). Early twin study research on this topic indicated that ADHD and case-control reading difficulties did not share a genetic basis, except in the instance where both traits were present together (Gilger et al., 1992). However, more recent research has suggested an association. For example, one meta-analysis of 38 multivariate twin studies with various different reading measures has indicated that quantitative reading and ADHD show an average genetic correlation of 0.42 (see Daucourt et al., 2019 for a review). A genetic association has also been explicitly reported between ADHD and reading comprehension and ADHD and decoding (Plourde et al., 2015). Due to the high genetic overlap and co-occurrence of the conditions, it has been suggested that dyslexia and ADHD are not two entirely distinct conditions despite their differing symptomology; instead, they may both be characterised by shared underlying aetiology (Pennington, 2006). Some evidence, including from the CLDRC sample, suggests that the shared aetiology may be due to a shared cognitive trait, such as processing speed, which contributes to both conditions, although the CLDRC work
does not distinguish between comprehension and decoding, and instead uses a cut-off of an overall reading score based on multiple skills to define reading difficulties (Shanahan et al., 2006; Pennington, 2006). It should be noted that one twin study has shown that different reading comprehension tests that are designed to test the same skill can produce different results regarding genetic covariance with other traits, even when administered in the same sample, in this case, the CLDRC (Betjemann et al., 2011). Given this, it is important to consider that the specific reading measures used may impact results when studying shared genetic architecture through twin studies.

While very long-term longitudinal genetic studies of reading are limited, some longitudinal twin research indicates that the genetics of reading are stable over time. For example, TEDS research has indicated that word reading is relatively stable between ages seven and 12, and that a large portion of the stability is accounted for by genetic stability (Harlaar et al., 2014) The study used both ACE twin methodology and supplemented this work with modern statistical genetics methodology, and both methods returned similar results. Similar results were found in follow-up CLDRC data using a composite based on both decoding and comprehension, in which ~75% of the stability in reading ability between two time points was found to be due to genetics, however it should be noted that in this study there was greater variation of ages and time points one and two, which may have affected the results (Astrom et al., 2007). Other family studies have produced similar results; the Colorado Adoption Study, which contains a sample of genetically related and adopted siblings, found that a high proportion of the stability in recognition-based reading ability between ages eight and 16 is due to the stability of genetic influences, and additionally heritability of reading in this sample increased with age (Wadsworth et al., 2001). For a more detailed account of the genetic stability of reading ability see Chapter 5.
2.2.3. Shared Environmental Influences

When looking at the results of twin studies, it must be considered that the ‘common environment’ element of the twin study can differ across countries and cultures, and that this may have an impact on the results. For example, some results indicate that heritability of word-reading in boys may be higher than that of girls in the UK TEDS sample, but not a US sample which includes participants from CLDRC and used a more general reading measure comprising both word-reading and comprehension (Hawke et al., 2006; Harlaar, Spinath, et al., 2005; Schumacher et al., 2007). Hawke and colleagues (2006) argue that this may be due to age or testing differences between the samples, and while this would be a valid explanation, it is also possible that this finding is due to cultural differences between the UK and US.

It has been specifically suggested that heritability may differ as a result of educational differences across countries, and a demonstration of this can be seen in ILTS. One study using ILTS data investigated the behaviour genetics of pre-reading, reading, and spelling in three geographical areas: the US, Scandinavia, and Australia, and ACE estimates for pre-reading skills were similar across all three samples (Samuelsson et al., 2006). Scandinavia was excluded from comparative follow-up analyses at the end of kindergarten due to large differences in schooling, however comparison between the US and Australian twin samples showed that heritability for later decoding-related measures and spelling was much higher in the Australian sample, and this difference approached statistical significance for spelling (Samuelsson et al., 2006). It is hypothesised that this could be due to differential schooling, as the Australian system places more emphasis on teaching reading skills at this age (Samuelsson et al., 2006). Additional research, which extended from pre-kindergarten to the end of 2nd grade, indicated that shared environmental influence decreased in favour of genetic influences in 1st grade in the
Scandinavian sample for decoding, supporting the hypothesis that differential schooling may influence heritability (Christopher et al., 2013). The authors also argue that differential orthography across the three areas may have contributed to this difference. Interestingly, Christopher and colleagues (2013) found that the contribution of genetics to the rate of change in skill was consistent across the three geographical sub-samples, indicating that genetics are important for learning to read, rather than being able to read, regardless of local context.

2.4.2. Limitations of Family Studies

While twin and adoption studies have been a valuable method for beginning to understand the heritability of reading ability, they hold several limitations that mean that they are, alone, not enough for us to truly understand the genetics of reading ability. First, the samples used are not representative of the general population, and therefore it has been debated whether the results of twin studies can be generalised outside of twins; for example, twins experience a birth and early life experience that is unique to them (Eaves et al., 1978).

Twin studies are bound by several underlying assumptions, a key one of which is the Equal Environments Assumption (EEA). This states that DZ pairs of twins are, on average, subject to no more or less similarity in their environments and treatment by others than MZ twins (Richardson & Norgate, 2005). Research investigating whether or not this assumption is likely to hold has been relatively limited, however there is evidence to suggest that it is not always the case, creating potential for inaccurate results, including in the case of cognitive traits (Richardson & Norgate, 2005). Richardson and Norgate (2005) argue that ‘socio-cognitive interactions’ may invalidate twin studies of cognitive abilities; this theory suggests that if MZ and DZ twin types are treated differently from each other, their own way of thinking may shift in response to that differential treatment, potentially influencing other cognitive traits. Another key
assumption that can introduce bias if not met is that of assortative mating, as parents who choose mates that are genetically similar to themselves results in DZ twin genetic similarity above 50% (Sahu & Prasuna, 2016). Concerns have also been noted surrounding the ascertainment of twin samples, in that selecting twin samples for the trait of interest can create bias (Bundey, 1991), providing another reason why use of samples which measure quantitative reading ability across the full distribution are useful for studying reading ability, rather than focusing only instances of cases and controls which are selected for reading difficulties.

A different type of ascertainment bias can also be an issue in adoption studies, as families with adoptive children tend to have higher quality environments, with a reduced distribution in environmental quality, than other samples (Stoolmiller, 1999). This is likely because it tends to be more educated and higher earning parents that adopt (Stoolmiller, 1999). As a consequence, that the results from adoption study cannot be generalised as they often do not represent the full variation of environmental circumstances that we might expect to see in other samples (Cadoret, 1995). In addition to this, pre-adoption circumstances and their potential impact on results must also be considered; in reading, an example of this is pre-adoption exposure to language (Petrill et al., 2006).

While twin and adoption studies can broadly tell us about the heritability of a trait, they are not able to direct us to specific genes which are influential (Friedman et al., 2021). In other words, twin studies are useful for obtaining a broad picture, but are not able to provide any information on specific genes, or any information on through which mechanisms those causal genes may be acting. In order to do this, molecular methods are required; these will be explored more in the section below, and in Chapter 4.

2.3. History of Molecular Methods
In 2001, the Human Genome Project was completed, which had large implications for the study of genetics (Lander et al., 2001; Venter et al., 2001; Gates et al., 2021). This meant that it was now possible to sequence the DNA from the whole human genome, which opened up many possibilities for molecular genetic research. Initially, DNA sequencing was time consuming and expensive, however, over the time the cost and labour required has reduced substantially, and sequencing a human genome is a relatively cheap and fast process (McCombie & Mcpherson, 2018). With the increasing access that researchers had to genetic data came a series of new methods that may be used to study the genetics of reading.

Prior to the availability of genome-wide sequencing data, molecular genetics studies of reading focused on linkage analysis. These attempted to isolate genomic regions that may be influential on a trait, using a small selection of polymorphic markers, often in family-based samples (Teare & Barrett, 2005; Skiba et al., 2011). Several of these studies were undertaken, particularly on the dyslexia or reading difficulty phenotype (e.g., Anthoni et al., 2007; Kaplan et al., 2002; Turic et al., 2003), but also on quantitative, decoding and phonology based reading skills (Grigorenko et al., 1997) and some success was found in terms of replication of genes associated with dyslexia across studies (Kere, 2014), with repeated replication found for a region on chromosome 6 (see Olson, 2002). Despite this, linkage studies have several disadvantages. Linkage studies do not tend to work well for traits which are highly polygenic, like reading (Erbeli et al., 2022), and they do not provide the capability to fine-map regions in order to determine the causal locus or loci (Skiba et al., 2011). Linkage studies have gradually fallen out of favour with the introduction of a newer method, the Genome Wide Association Study (GWAS).
GWAS take a full genome of SNPs and test against the phenotype of interest for significant associations between each SNP and the phenotype (Uffelmann et al., 2021). Several GWAS have been conducted on dyslexia and other reading related phenotypes; a full literature review can be found in Chapter 4. There are two key studies in this area that are particularly significant in the field, as they represent the largest GWAS meta-analyses that have taken place on dyslexia (51800 cases, 1087070 controls) (Doust et al., 2022) and five quantitative reading and language related skills: word reading, nonword reading, spelling, phoneme awareness, and nonword repetition (n = up to ~34000) (Eising et al., 2022), and large sample sizes are essential for GWAS studies to have sufficient power to detect causal variants of small effect (Uffelmann et al., 2021). These GWAS have failed to replicate many earlier findings (Erbeli et al., 2022).

Both of these meta-analyses have indicated that reading is heritable, supporting the work of twin studies, however in both cases heritability estimates which are based only on common SNPs are lower than those estimated from twin studies which include all types of genetic variation (Doust et al., 2022; Eising et al., 2022). This lower SNP heritability is a common finding across GWAS studies known as missing heritability (Young, 2019). The true cause of missing heritability is unknown, but a range of hypotheses have been proposed, although a full exploration of these is outside the scope of this thesis (Young, 2019).

Aside from confirming that molecular data shows that reading and reading difficulties are heritable, downstream analysis of the two key studies mentioned above has provided a wealth of additional information about the genetics of reading (Doust et al., 2022; Eising et al., 2002). In both studies, genome-wide significant variants were identified, indicating that those variants have a significant effect on the outcome trait. In the dyslexia study, 42 significant variants were identified, and 27 of these variants had not previously been identified, whereas the remaining 15
had previously been identified as potential causal variants for measures of cognitive or education
related traits (Doust et al., 2022). Despite the large number of significant variants identified, no
previous dyslexia-specific associations were replicated (Doust et al., 2022). By contrast, only one
variant achieved genome wide significance in the quantitative word reading ability GWAS
(Eising et al., 2022). No significant variants were identified for the other traits, or in a
multivariate GWAS of all five traits. There are several possible reasons for this difference. To
begin with, the dyslexia GWAS included a substantially larger sample size than the quantitative
word-reading GWAS, although both samples were very large. In addition to this, it is possible
that these results are in support of previous evidence suggesting that variants that contribute to
low reading ability have larger effect sizes than those which contribute to other variation in
reading ability (DeFries & Alarcón, 1996).

In both cases, downstream analysis was able to provide some indication of the potential
biological pathways through which variants may be acting. In the dyslexia study, it was found
that genetics of dyslexia do not correlate with genetics of brain differences associated with
language, indicating that brain structure is not likely to be a cause of dyslexia (Doust et al.,
2022). Despite this, gene-based analysis showed that three significant genes in close proximity to
significant SNPs were predominately expressed in the brain (Doust et al., 2022). The
multivariate GWAS output of quantitative reading traits showed significant genetic correlation
with one measure of brain structure that had been previously associated with language, out of 58
variables tested (Eising et al., 2022), which is broadly consistent with the results reported by
Doust and colleagues (2022). In addition to this, both studies investigated genetic correlations in
range of cognitive, psychiatric, and health traits. In both studies, findings broadly indicated that
low reading ability is genetically associated with lower SES, poorer educational outcomes, and a
range of poor health outcomes, although there are differences on individual traits between the two studies (Doust et al., 2022; Eising et al., 2022). More information on this can be found below, and in Chapter 6.

Along with showing that reading is heritable, molecular genetics studies have also provided additional support to twin-based findings that there are genetic similarities between reading ability and other traits. For example, a TEDS based study has identified similarity between reading comprehension and decoding and educational attainment (Selzam et al., 2017), and a separate study found associations between decoding-based skills and comprehension and ADHD (Verhoef et al., 2019), through the calculation and application of polygenic scores (PGS). PGS are estimates of the genetic propensity an individual has for a trait, and are calculated based on GWAS output of additive variance per SNP and the individual’s genetic data (Sugrue & Desikan, 2019) (See Chapter 6 for more details). In addition, the calculation of genetic correlations using the Linkage Disequilibrium Score (LDSC) regression method has found significant genetic correlations between dyslexia and ADHD (Doust et al., 2022). Early molecular research has indicated support for the findings regarding the genetics of reading ability and general cognitive ability reported above. For example, one molecular genetic study applied a PGS for overall reading in children and young people to a large sample of adults, and this predicted some, but not all, variance in adulthood verbal-numerical reasoning (Luciano et al., 2017). The association remained even when controlling for a PGS of cognitive ability, providing further evidence that reading may be partly causal for intelligence (Luciano et al., 2017), and that genetics of reading research must be interpreted with this in mind.

It is possible for a genetic variant to act on two or more phenotypes, either through directly influencing a biological pathway that is relevant to multiple traits (horizontal pleiotropy).
or by influencing one phenotype, which in turn then influences another phenotype (vertical pleiotropy) (Verbanck et al., 2018). The LDSC method has been used more generally to start to develop a picture of pleiotropy across quantitative human traits. One study which carried out LDSC regression analysis on 558 existing GWAS of a wide variety of complex human traits has indicated that shared genetic architecture is high, with 15% of correlational pairs showing a significant genetic correlation at the Bonferroni level (Watanabe et al., 2019). In addition to this, in this study gene-level expression data was used to show that pleiotropic genes enact less-specific functions by being expressed in more tissue types, therefore leading to a more general effect (Watanabe et al., 2019). This is in line with the ‘Generalist Genes Hypothesis’, which proposes that variants that are influential for one cognitive trait, such as reading, are more likely to also have influence on other traits (Plomin et al., 2007).

2.4. Gene-Environment Interplay

As has been established throughout Chapters 1 and 2, both genetics and environment have important roles to play in predicting reading ability. However, the two do not operate in isolation; it is possible for genetics and environment to interact or correlate with each other. Gene-environment correlation occurs when a genetic propensity for a trait results in selection of environments that then continue to reinforce that behaviour (Jaffee & Price, 2007). Several examples of this have been identified in relation to reading ability, for example, a study in a small sample found that children with a family history of dyslexia who are themselves are poor readers spend less time reading than children with a family history of dyslexia who do not have reading difficulties (Snowling et al., 2007). Gene-environment correlation has also been identified in studies that have had a specific focus on genetics. One longitudinal twin study that utilised genomic data for GCTA analysis has found that reading fluency and print exposure at
age twelve are very highly genetically correlated, with an estimated value of 0.89 using GCTA and 0.58 using more traditional twin modelling methods (Harlaar et al., 2014).

Genetics and environment can interact in other ways, such as where one moderates the other. An example of this can be seen in a twin study of reading difficulty (measured as poor performance on a reading composite comprising multiple skills), in which the education of the parents was also measured (Friend et al., 2008). In this study, it was found that the genetic element of the twin model was more influential for individuals whose parents had higher levels of education (Friend et al., 2008). The authors offer several hypotheses to explain their results, but favour the bioecological model of gene-environment interaction, which states that genetics become more influential in favourable environments (Bronfenbrenner & Ceci, 1994). A similar study did not replicate this finding, with no moderation effect found with parental education or reading ability (Kirkpatrick et al., 2011).

**2.5. Literature Critique and Contribution of this Thesis**

Several limitations in the current literature have already been identified throughout this piece, notably the lack of replication that has been observed and the fact that our knowledge of the relationship between the genetics of reading and other traits is limited at the molecular level. As with phenotypic research, a key limitation in the research on the genetics of reading is a lack of molecular genetics studies on the full life course, with most research projects focusing on early years and childhood only. As with phenotypic research, a key limitation is the availability of this data. Additionally, very few large, longitudinal, population-based cohort studies exist that have both genetic data and information on reading ability at multiple ages (This is explored in more detail in Chapter 3).
Another area where further research is needed is in the understanding of the genetic architecture of reading in relation to other traits, particularly neurodevelopmental disorders and physical and mental health problems. While research has begun to study these issues in more depth, using advances in molecular genetics methods, little is known about the pleiotropic effects that variants associated with reading ability may have. This is work that is required alongside continued efforts to better understand the SNP-based heritability of reading ability, as it may yield useful insights into the underlying aetiology of one or more of the traits involved.

As has been mentioned previously, existing behaviour genetic research on reading ability has been limited by the small sample sizes available. As with phenotypic research, data collection in this area must be prioritised. However, this is a slow process, with data collection starting at birth not being useable for studies on reading ability for several years. In order to move the process along more quickly and to make efficient use of existing resources, current research must seek to make the best possible use of existing datasets in this area, even where they are suboptimal for the research question. The NCDS dataset provides a rare opportunity to better utilise an existing, rich longitudinal dataset for reading research to build on our current knowledge.

This thesis aims to close the gap in several ways. To begin with, this thesis will present an updated version of a moderately sized, longitudinal genetic dataset for use in quantitative reading research. While the sample size of the NCDS genetic dataset alone is not adequate to detect significant variants, genetic reading research in this dataset can help to address many of the other questions we have surroundings the genetics of reading ability, including the longitudinal influence of genetics. It is also possible that this dataset could be used in future meta-analyses of reading ability, as this thesis has demonstrated that the reading measures
included are valid and reliable. In addition, Chapter 6 of this thesis presents LDSC based genetic correlations to help us to better understand the stability of the genetics of reading, how NCDS reading measures compare to psychometrically assessed reading measures in other samples, how the genetics of reading are associated with the genetics of other traits, including some traits mentioned in this review, such as ADHD, and some traits that have very rarely been studied in this context, as occupational outcomes and pain.
2.6. References


throughout the school years. *Scientific Studies of Reading*, 21(4), 334–349.

https://doi.org/10.1080/10888438.2017.1299152


https://doi.org/10.1038/s41588-018-0099-7


https://doi.org/10.1023/a:1012218301437

https://doi.org/10.1111/1469-7610.00085

https://doi.org/10.1007/s11881-007-0009-7

https://doi.org/10.1017/thg.2015.66


SECTION TWO:

DATA AND METHODS
Chapter 3: The National Child Development Study

3.1. National Child Development Study Overview

The National Child Development Study (NCDS) is a long-running, UK based birth cohort study. The NCDS dataset is the core data used throughout this thesis, and part of the contribution that this thesis makes is to show that the NCDS, despite its limitations, is a valuable resource for conducting reading research. In order to properly set out the significance of this dataset, this chapter will provide an overview of the NCDS. It will start with a description of the survey itself, including how it began, when and how data has been collected, and an example of a range of variables that were included. This thesis only uses data from ages birth to 50, and so this will be the core focus here. This description will include some information on the characteristics and representativeness of the collected sample. Following on from this, this chapter will include an overview of the genetic information that has been collected from participants. This includes information on how the sample was selected, and how the genetic data was collected.

Due its status as a large-scale dataset with several reading related variables, there is already a body of reading research that has been undertaken in the NCDS. This is primarily phenotypic in nature, but there is also a small number of genetic studies of reading which have been conducted. This chapter will include a brief review of the reading research that has already been carried out in the NCDS. Largely due to its age, the NCDS data has several limitations, in relation to both the genetic and phenotypic data. This section will describe these limitations, then go on to show how some of these can be overcome to make the NCDS a valuable resource for future reading research.
3.1.1. National Child Development Study Origins and Data Collection

The NCDS is a population-based, British cohort study for which data collection began in the year 1958. During one week in March in 1958, all birth parents of new births in England, Wales, and Scotland were invited to participate in a survey (Power & Elliott, 2005). At this time, the NCDS was known as the Perinatal Mortality Survey, and it was intended to be a ‘one-off’ survey with the aim of collecting evidence to help reduce cases of stillbirth (n = 17416) (Power & Elliot, 2005). However, follow up surveys were later commissioned and the resource became known as the National Child Development Study, or alternatively, the 1958 British Birth Cohort, with participants re-contacted at semi-regular intervals for additional rounds of data collection (Davie & Butler, 1972; Power & Elliott, 2005). During the childhood years, classed as from ages seven to 16, any immigrants to England, Wales, or Scotland who were born during the week that initial data collection began were also invited to take part (Power & Elliot, 2005). Below is a summary of the survey at each wave, including method of data collection, age and number of participants, and a brief selection of variables included.

**Year 1958: Birth.** The first point of data collection, sometimes referred to as Wave 0, occurred in 1958 and 17416 people took part (Power & Elliot, 2005). This point of data collection primarily focused on factors relating to the health and lifestyle of the mother, such as illness, smoking, and SES, along with questions about the birth and health factors related to the new baby, such as birthweight and any health conditions that had already been identified (Power & Elliot, 2005; Centre for Longitudinal Studies UCL, n.d.-l). Data was collected through face-to-face interviews with the mother of the child and the midwife, along with using data from medical records (Centre for Longitudinal Studies UCL, n.d.-l).
**Year 1965: Age Seven.** The first follow-up occurred when the subjects were seven years old (n = 15425), and participants were traced through their school records (Power & Elliot, 2005). Data collection at age seven was intensive; data was collected on a wide range of variables through face-to-face means (Centre for Longitudinal Studies UCL, n.d.-a). Information was collected on the child’s home and family life through interviews with parents, and on their school achievement and behavior from interviews with their teacher (Centre for Longitudinal Studies UCL, n.d.-a). The children completed a series of ability tests, including reading, maths, and general cognition, and underwent a thorough medical assessment (Centre for Longitudinal Studies UCL, n.d.-a).

**Year 1969: Age Eleven.** The next follow-up occurred when participants were eleven (n = 15337) (Power & Elliot, 2005). Data collection was similar to age seven, however along with completing tests, participants also completed a physical survey form about their leisure time and wrote an essay about what they believed their lives would be like at age 26 (Centre for Longitudinal Studies UCL, n.d.-b).

**Year 1974: Age 16.** Age 16 was the final year through which respondents could be traced through their school records (n = 14647) (Power and & Elliot, 2005), and the last time point at which parents and teachers were interviewed on behalf of the participant (Centre for Longitudinal Studies UCL, n.d.-c). Data collection was similar to age eleven, with some adjustments. There was no cognition test at age 16, but participants were asked to complete a physical survey about attitudes, aspirations, spare time, and how they felt they performed in school subjects (Centre for Longitudinal Studies UCL, n.d.-c). Teachers were asked about what exams the participants would be entered for (Centre for Longitudinal Studies UCL, n.d.-c).
Year 1981-2: Age 23. Age 23 is the first adulthood measurement of NCDS data. The sample size dropped substantially at this wave due to difficulties in contacting participants who had moved on from their addresses (n = 12537) (Power & Elliot, 2005). Subjects participated in a face-to-face interview covering topics such as education, employment, skills, income, relationships, children, lifestyle, and health (Centre for Longitudinal Studies UCL, n.d.-d).

Year 1991: Age 33. At age 33 (n = 11407), participants underwent a similar interview to age 23 (Power & Elliot, 2005; Centre for Longitudinal Studies UCL, n.d.-e). In addition, they filled out physical surveys about attitudes and employment history; their partners were also asked to fill out the employment history questions (Centre for Longitudinal Studies UCL, n.d.-e). A third were asked to fill in a physical survey about their children, including their schooling and behaviour, and the children completed cognitive tests (Centre for Longitudinal Studies UCL, n.d.-e). This was the last year for which the interviewers recorded answers without a computer.

Year 1999-2000: Age 41/42. At age 42, 11419 individuals were interviewed (Power & Elliot, 2005). For this wave, questions were aligned to another birth cohort study, the 1970 British Cohort study, and topics covered housing, income, employment, leisure, healthy lifestyle behaviours, attitudes, and criminal convictions (Centre for Longitudinal Studies UCL, n.d.-f).

Year 2005-2005: Age 45/46. At this age, a short telephone interview was conducted (n = 9536) (Centre for Longitudinal Studies UCL, n.d.-h). This included questions about substantial life changes, income, employment, skills, health behaviours, and health, and using a computer (Centre for Longitudinal Studies UCL, n.d.-h). It should be noted that research conducted at this age has identified that those with poor reading scores are more likely to have left the study (Atherton et al., 2008).
**Year 2008-9: Age 50/51.** For the age 50 wave, a mixture of interviews and paper surveys were used (n = 9790) (Centre for Longitudinal Studies UCL, n.d.-i). Before the interview, a paper survey was sent out to the participants for them to complete; this included questions about free time, attitudes, and personality (Centre for Longitudinal Studies UCL, n.d.-i). The interview was administered by an interviewer using an electronic device, and a portion of the interview required the subject to fill out answers on the device themselves (Centre for Longitudinal Studies UCL, n.d.-i). The interview topics included employment, health, health behaviours, and cognitive tests (Centre for Longitudinal Studies UCL, n.d.-i). Data from this wave was linked to NHS records.

**Year 2013/14: Age 55.** Similar to age 44/45, this wave of data collection consisted of a shorter survey (n = 9137) (Centre for Longitudinal Studies UCL, n.d.-j). The participants were contacted and asked to complete a questionnaire online, and those who failed to do so were invited to complete a short telephone interview. Topics included family, housing, income, and health (Centre for Longitudinal Studies UCL, n.d.-j).

**Year 2020-2024: Age 62-64.** This, most recent, wave of data collection has been disrupted by the COVID-19 pandemic, and so data collection was paused part way through (Centre for Longitudinal Studies UCL, n.d.-k). Data collection has recommenced, and it is anticipated that data will be available in Autumn 2024 (Centre for Longitudinal Studies UCL, n.d.-k).

**Sub-Studies.** Over the years, several ‘sub-studies’ have been carried out on NCDS participants (Shepherd, 1995). These have normally involved additional waves of data collection for individuals with the traits of interest (Shepherd, 1995). Examples of topics studied include
adoption, ‘gifted’ children, successful disadvantaged individuals, those with basic skills problems, and those with epilepsy (Shepherd, 1995)

3.2. Genetic Data Collection in the NCDS

At age 44, a subsample of participants in the NCDS underwent a biomedical survey (Centre for Longitudinal Studies UCL, n.d.-g). This sample was collected differently to those above, in that a more strategic approach was taken to ascertainment. 12034 NCDS participants were asked to take part in the biomedical survey, and a total of 9349 participated (Fuller at al., 2006). NCDS cohort members who had not taken part since age 16 were not invited to participate, and neither were those who had not participated in the most recent wave at age 42, or those for whom current contact details were not held (Fuller et al., 2006). The biomedical survey was conducted with the intention of gathering objective health information in adults (Fuller et al., 2006).

A wide range of health-related information was collected during the biomedical survey. This included paper surveys, an interview and health assessment by a nurse, and a self-completion computer assisted questionnaire (Fuller et al., 2006). The most significant element of the biomedical survey for this thesis is the collection of blood samples from which DNA was extracted according to the procedure outlined in Fuller and colleagues. 8404 participants agreed to extraction, storage, and use of DNA (Fuller et al., 2006).

DNA was genotyped on eleven arrays (Newcastle University, n.d.), although it was advised in conversation with researchers familiar with the data that only seven of those arrays were appropriate for GWAS analysis, due to issues of coverage, and so only seven of those arrays are used in this thesis. The genetic data from the NCDS has been used as control data for
several other studies: the Affymetrix 500K, Infinium HumanHap 550k v1.1 and Illumina 15k Custom Chip have all been used as part of the Wellcome Trust Case Control Consortium 1; the Illumina 1.2M and Affymetrix v6 have been used as part of the Wellcome Trust Case Control Consortium 2; the Infinium HumanHap 550k v3 has been used as control data in the Type 1 Diabetes Genetics Consortium; and the Illumina Human 660-Quad array was used as control data for the Large Scale Genome-Wide Association Study of Asthma (Newcastle University, n.d.). There was significant overlap between the samples collected on different arrays; the approach taken to this is explained in detail in Chapter 5.

3.3. Phenotypic Reading Research in the NCDS

A body of phenotypic reading research in the NCDS already exists, primarily studying predictors and outcomes of reading ability. Here I will present a short review of this existing research, focusing on areas which replicate the more general findings reported in Chapter 1. Many predictors of low reading ability discussed in Chapter 1 are replicated by studies that have been carried out on NCDS data. One such study has found a positive correlation between birth weight and reading ability at ages seven, 11, and 16 (Jefferis et al., 2002). Another study found similar results, in that birth weight predicted reading comprehension at age 16, however it was also found that most of this association was mediated by verbal IQ at age 11 (Da Silva et al., 2019). This is unsurprising given that, as covered in detail in Chapter 1, reading skill and cognitive ability are closely intertwined. In the NCDS, maternal smoking during pregnancy is associated with lower word-reading scores at age seven (Butler & Goldstein, 1973; Sellers et al., 2020), and lower reading comprehension scores at 11 (Butler & Goldstein, 1973). In addition, parental nurturing behaviors up to and including age 11, such timing of first prenatal visit, taking an interest in the child’s school activities, and having high aspirations for the child, were
associated with better reading comprehension scores at age 11 (Michael, 2011). In addition, associations between reading ability and behavioural traits have also been identified, for example, one study has found an association between low attendance at age seven and reading comprehension at age 16, and low attendance at age 15 and reading comprehension at 16, and once the latter is controlled for, the relationship between attendance at age 7 and reading comprehension attenuates, indicating that contemporary attendance is the more important factor (Fogelman, 1978).

Existing NCDS research provides evidence for some of the co-occurring traits mentioned in Chapter 1. One existing study has found that reading ability in the NCDS is strongly associated with other neurodevelopmental difficulties (Addicoat et al., 2019). Using structural equation modeling, the authors showed that word reading, measured using the 30 item Southgate Test, loads onto a general factor of childhood neurodevelopmental abilities, along with other traits including maths, general cognitive ability, motor coordination, hyperactivity, and restlessness. Another has found that non-right handedness at age 11 is associated with low reading comprehension (Crow et al., 1998; Björk et al., 2012). Speech problems at age 11 are associated with poor comprehension scores at the same age (Calnan & Richardson, 1977).

Several studies of reading in the NCDS emphasise how important childhood reading skills are as a predictor of outcomes much later in life. The general neurodevelopmental factor mentioned above was able to predict mood problems at each adulthood wave, up to and including age 50 (Addicoat et al., 2019). This corroborates phenotypic and genetic evidence discussed in Chapters 1 and 2 that indicate that reading ability covaries with neurodevelopmental traits, and phenotypic evidence that low reading ability is associated with depressive symptoms. Another study found that reading ability in childhood, measured as a latent factor with observed
variables including objective and subjective measures of reading at ages seven and 11, mediates the association between early- and pre-term birth and qualifications achieved by age 33 (Basten et al., 2015). High reading score at age seven is associated with greater SES at age 33 (Currie & Thomas, 1999), and a structural equation modelling study found that reading score at age seven is associated with SES at age 42 (Ritchie & Bates, 2013). For women, this study reported that there was an effect of reading ability on SES that was independent of all confounding variables, indicating that reading ability hold specific predictive value alongside cognitive ability and years of education. It has also been found that the reading comprehension test at age 16 predicts earnings of full-time workers throughout adulthood to age 50 (Watts, 2020).

NCDS reading research also indicates how reading ability can be ‘transmitted’ across generations; one study found that the word-reading scores of NCDS participants at age seven correlate with the reading scores achieved by their own children on a different reading test (Brown et al., 2009). In addition, NCDS research shows how parental activities can influence the reading ability of their children; one example shows that the extent of socialising carried out by the NCDS participants is positively associated with the reading ability of their children (mean age eight years) (Brown & Taylor, 2009). The authors suggest that this is an element of social capital, once again showing the significance of familial SES in reading ability. While all of these studies are informative, they all fail to control for genetic confounding, in other words, the potential effect that parental genes may have on the outcome of the participant.

Research in the NCDS has indicated a range of health outcomes associated with reading ability in this dataset, including mental health outcomes. One study has found that participants who had died from suicide by age 54 showed a different trajectory of reading ability to those who were still alive. Reading test scores at ages seven, 11, and 16 were standardized, and it was
found that those who had died by suicide showed a less pronounced increase in the standardized reading score between the ages of seven and 16 than their peers who were still alive (Richard-Devantoy et al., 2019). This indicates that it is not only reading skill that can predict life outcomes, but development of reading skill (Richard-Devantoy et al., 2019). Another study has found that childhood reading skills are associated with cortisol levels in adulthood, but only in males; it was hypothesised that this may be a response to long-term stress caused by lack of achievement (Power et al., 2008). Reading comprehension at age 16 has been associated with diagnosis of Type 2 diabetes by age 42, even when adjusting for Body Mass Index (BMI) (Olsson et al., 2008).

Almost all existing work on the genetics of reading in the NCDS uses the participants as controls for reading disabilities rather than studying quantitative reading as a trait, as there is no variable regarding dyslexia or reading disability in the NCDS (Pagnamenta et al., 2010; Gialluisi et al., 2021). However, one study has used the NCDS to investigate the genetics of quantitative reading (Davies et al., 2015). This study conducted GWAS on multiple cognitive related traits, using a version of the genomic dataset that was imputed to the older 1000 Genomes reference panel, rather than the more recent Haplotype Reference Consortium panel that is utilized in this thesis (Auton et al., 2015; McCarthy et al., 2016). The traits tested included word-reading on the Southgate test at age seven and reading comprehension at age 11, estimating heritability for these traits as 0.15 and 0.28, respectively, using the GREML method. The GCTA method was used to estimate genetic correlations between the traits with an educational attainment measure, and each showed a moderate genetic correlation with the number of O-Levels achieved ($r_G = 0.20$ at age seven, $r_G = 0.61$ at age 11). At age 16, heritability of reading comprehension was estimated at 0.26, and genetic correlation with O-Levels achieved was moderate at 0.53 (Davies et al., 2015).
3.4. Limitations of the National Child Development Study

Despite its clear breadth and depth of information, there are several challenges associated with properly utilizing the genetic data associated with the NCDS. The key challenge is that the funding and data holders of the NCDS have changed several times since its inception (Power & Elliot, 2005), and this applies to genetic data as well as phenotypic data. As a result, centralized documentation of the genetic data available in the NCDS is limited, with very little information available regarding specific elements of the data, including genome-build, sample overlap, and which, if any, quality control and filtering of individuals has already been carried out. As a result, the genetic data is difficult to work with. Additionally, existing research (e.g., Davies et al., 2015) has been conducted with a version of the NCDS genetic data which has been imputed to the 1000 Genomes reference panel (Auton et al., 2015). This has now been superseded by the Haplotype Reference Consortium reference panel, which offers higher quality imputation, meaning that the version of the NCDS genetic data held by the data managers prior to the work contributed in this thesis was outdated and not optimal (McCarthy et al., 2016). Updating this is challenging, due to the lack of documentation regarding the central version of the dataset.

In addition to this, the NCDS has several limitations regarding the phenotypes that are available. The long-running, longitudinal nature of the data means that measurements are not consistent and often change from wave to wave, including for health and SES related variables. Reading variables are not consistent across the study, meaning that they cannot be easily monitored in a longitudinal fashion. Despite holding information on reading ability at ages seven, 11, 16, 23, and 33, the NCDS does not have a consistent reading test. Instead, the Southgate word-reading test is used at age seven, which is designed to identify poor readers, and comprehension tests are delivered at ages 11 and 16 (Shepherd, 2012). Alongside this are a series
of subjective teacher- and self-report reading measures, which while informative, do not fulfill the criteria required for inclusion in meta-analyses of genetics of reading studies (e.g., Eising et al., 2022), however this does not mean that they are not useful sources of information that could be used.

3.5. Benefits of using the NCDS for Reading Research

Although there are several challenges associated with it, the NCDS is still a valuable resource for carrying out reading research in the UK. It is one of the longest continuously running population-based cohort studies in Great Britain, surpassed only by The MRC National Survey of Health and Development Cohort of 1946 (Wadsworth et al., 2005). In addition to this, the NCDS has a unique combination of characteristics that make it well suited to reading research. The cohort is long-running, meaning that predictors and outcomes of reading ability can be studied across the lifespan from birth to old age. Many participants now have children and grandchildren, making it possible to study these influences across generations. Additionally, the genetic data available means that the genetic interplay between reading and environmental factors can be untangled. The NCDS offers an underutilized opportunity to explore both the phenotypic and genetic correlates of quantitative reading ability in a population-based across the lifespan.
3.6. References


Newcastle University. (n.d.). *GWA data | Data to Knowledge (D2K) Research Group | Newcastle University.*
https://research.ncl.ac.uk/d2k/ourresearch/58forwards/geneticdatatypes/gwadata/


Chapter 4: Molecular Genetic Methods

In order to explore genetic correlates of reading, this thesis uses three key molecular methods. As way of background to the next three chapters of this thesis, here I will provide an overview of each of these three methods: (1) Genotype Imputation, (2) Genome Wide Association, (3) Linkage Disequilibrium Score Regression based genetic correlation. This chapter will provide a brief explanation of how each of these method works, and why they were utilised in this thesis.

4.1. Genotype Imputation

The first method I will cover here is that of genotype imputation. Genotype imputation is carried out to ensure that as many SNPs as possible can be tested in molecular genetic analyses, by using statistical probabilities to impute missing variants (Marchini & Howie, 2010). This is required because most arrays do not collect data for all genotypes; most arrays collect hundreds of thousands of SNPs (Das et al., 2016), and it’s estimated that arrays tend to sample around 1% of the human genome (Das et al., 2018). Imputation of missing variants allows this to be extended so that millions of variants may be analysed (McCarthy et al., 2016). The human genome contains regions of DNA made up of variants which are usually inherited together in chunks, called haplotypes (Li et al., 2009). Variants which commonly occur together are said to be in Linkage Disequilibrium (LD) (Belmont et al., 2005). The genotype imputation process capitalises on the existence of these haplotypes in order to estimate missing variants based on a reference panel, for which all individuals are extensively genotyped or fully sequenced (Li et al., 2009; McCarthy et al., 2016). Genotype imputation differs from usual statistical imputation techniques, as it is designed to accommodate the unique characteristics of genetic data, such as mutations, linkage, and areas of genetic recombination (Das et al., 2018).
Modern genotype imputation is carried out in two steps; pre-phasing is a process that occurs before the imputation itself (Van Leeuwen et al., 2015). Pre-phasing is the process in which the haplotypes in the data sample are estimated based on the SNPs that have been genotyped, and the imputation process itself occurs when the individual variants within those haplotypes are estimated (Visscher et al., 2017; Howie et al., 2012). Use of pre-phasing was initially recommended as it is more computationally efficient than pre-existing alternatives, such as the haplotype sampling method, which struggled to keep up with increases in the size of reference panels (Howie et al., 2012). Use of the pre-phasing method resulted in only a minor decrease in accuracy compared to the haplotype sampling method, with large gains in efficiency (Howie et al., 2012), and accuracy has improved since early pre-phasing programmes were developed (Loh et al., 2016).

This thesis uses the Eagle 2 phasing engine, followed by the minimac4 pipeline for imputation, both delivered through the Michigan Imputation Server (Das et al., 2016). Using the Michigan Imputation Server has several benefits. It is efficient without a drop in accuracy, and it is user-friendly (Das et al., 2016). Use of the Michigan Server allows imputation to the Haplotype Reference Consortium (HRC) reference panel, a large panel that provides good imputation quality compared to other reference panels (Das et al., 2016; McCarthy et al., 2016). The HRC reference panel sample was drawn from 20 studies of individuals who had, in most, cases, undergone low-coverage whole-genome sequencing (McCarthy et al., 2016). The HRC reference panel included data from a range of sources, including the 1000 Genomes reference panel (Auton et al., 2015), the UK10 project on rare variants in health, (Walter et al., 2015), and the Genome of the Netherlands project (Francioli et al., 2014), amongst others.
There are several reasons why genotype imputation was used in this thesis. Even though genotype sequencing technology continues to improve, it is not cost-effective to sequence a whole genome, therefore SNP arrays only collect data on a pre-selected set of SNPs (Goodwin et al., 2016). It has been suggested that there is little to gain from sequencing a whole-genome for GWAS studies, as common variants are imputed with high accuracy (Visscher et al., 2017). Imputation increases number of variants available for analysis across all arrays, which provides greater power to the GWAS analysis (Das et al., 2018), and has even contributed to discovery of causal variants which were not previously identified in non-imputed data (Li et al., 2009). In addition to this, The NCDS genetic data was genotyped across several different arrays, each of which included a different set of SNPs. Genotype imputation, using the same procedure each time, ensures that the different genetic datasets collected on different arrays have the same sets of variants and so can be meta-analysed together (Van Leeuwen et al., 2015).

There are several large reference panels that can be used for European populations, however in this thesis I will be using the HRC reference panel, which is the most recently released large-scale reference panel for European populations (McCarthy et al., 2016). The HRC panel combines data from multiple studies, including samples with a range of high and low genome coverage, and offers high quality imputation (McCarthy et al., 2016). One study has shown that using the HRC reference panel for imputation, rather than the 1000 Genomes Reference panel, can improve the quality of GWAS by producing smaller P-values for suggestive loci when testing the same trait in the same sample (Iglesias et al., 2017).

4.2. Genome Wide Association Studies and Meta-Analysis

The second molecular genetic method that will be explored here is that of the Genome Wide Association Study (GWAS), which was introduced in Chapter 2. The GWAS is based on
genomic data in the form of Single Nucleotide Polymorphisms (SNPs), individual loci where an allele may differ. GWAS is based on regression analysis; the trait of interest is regressed onto each SNP on the autosomes, in order to determine whether there is an association between each genetic variant and the trait (Friedman et al., 2021). Sex chromosome data is treated slightly differently; see below for details of how this was handled in this thesis. The trait of interest may be any trait with a heritable component, and may be defined continuously or dichotomously (Dehghan, 2018). Reading ability has been operationalised both ways in previous research. It is standard to control for potential confounders when carrying out GWAS, including sex and ancestry principal components (Friedman et al., 2021). A GWAS may involve hundreds of thousands, or even millions, of SNPs, and so Bonferroni correction for multiple testing is required in order to ensure that associations found are not false positives, meaning that a commonly used significance threshold is $P < 5 \times 10^{-8}$ (Friedman et al., 2021), although suggestive thresholds may be considered at $P < 5 \times 10^{-4}$ or $P < 5 \times 10^{-5}$ (Hammond et al., 2021). In addition to this, because effect sizes are so small, GWAS requires very large sample sizes in order to have adequate power to detect true associations (Uffelmann et al., 2021), and this can be difficult given the challenges of collecting phenotypic data from samples of this size (Friedman et al., 2021; Eising et al., 2022). It is typical for GWAS to focus on common variants, for two reasons: (1) it has been hypothesised that common variants are likely to be impactful for common conditions and (2) rare variants may not have enough power to detect significant associations (Dehghan, 2018; Uffelmann et al., 2021).

GWAS typically measure additive genetic variance, meaning that the effects of each influential SNP combine additively to produce the effect on the trait. In the case of complex human traits, large sample sizes are needed to achieve genome-wide significant effects, as the
effect sizes of SNPs can be extremely small (Uffelmann et al., 2021). It should be noted that GWAS are not always able to identify the specific causal variant that is causing the effect, and it is possible instead for significant variants to be in LD with the true causal variant (Visscher et al., 2017; Friedman et al., 2021).

Given the advantages that large sample sizes provide, it is common for multiple samples to be meta-analysed together in order to increase the statistical power of GWAS analyses (Begum et al., 2012). The data used in this thesis was collected on several different arrays with different levels of coverage, and therefore the decision was made to treat the sample for each array as a separate sample, which could then be brought together and meta-analysed. The first round of GWAS, conducted on each array, was carried out using RVTESTS, which produces summary level results which can be easily meta-analysed (Zhan et al., 2016). RVTESTS handles male sex chromosome data by multiplying dosage values by two, and in the meta-score analysis, certain statistics, such as the p-value, Hardy-Weinberg value, and sample size of the three allele combinations for each locus are calculated using data only from females (Zhan, n.d.). The ensuing meta-analysis was carried out using METAL, a specialist GWAS meta-analysis software, using the classical standard errors based approach (Willer et al., 2010).

In order to demonstrate how GWAS methods have been used on quantitative reading ability to date, I will present here a summary of existing GWAS on reading, excluding the recent large meta-analysis by Eising and colleagues (2022), which is reported in Chapter 2. The intent of this section is to demonstrate how GWAS of reading have developed over time. This will conclude with examples of success stories from other cognitive traits, in order to demonstrate that given the right sample and measurements, GWAS can be an effective method for understanding the genetics of cognitive traits.
Several GWAS have been carried out on quantitative reading ability, although sample sizes have usually been relatively low. Early GWAS included a twin-based design, which measured a reading composite comprised of a word-reading test and a teacher rating at age seven in a moderately sized UK sample (N = 5760) (Meaburn et al., 2007). Several SNPs were identified as showing a possible association with reading ability, however, effect sizes were small, and it was from this study that it became clear that individual effect sizes on quantitative reading ability were likely to be small (Meaburn et al., 2007). In 2013, GWAS was conducted on quantitative reading ability in two cohorts (maximum total N = 6649), one based in the UK and the other a twin- and sibling-based Australian cohort (Luciano et al., 2013). Reading was measured through a series of tests, which measured a variety of component skills, such as word-reading, non-word reading, and spelling. In this study, a small number of SNPs showed suggestive association with reading and spelling measures, and several SNPs were implicated in more than one reading skill, suggesting that these variants may be influencing overall reading skill rather than any one component process (Luciano et al., 2013). This study also found overlap in associations for reading and language; an early example of molecular genetic evidence of genetic pleiotropy across traits. A similar study was carried out by Gialluisi and colleagues (2014) in a much smaller sample size (N = 1862), examining a series of different quantitative reading-related measures across three samples. In this case, the first principal component of the reading measures was used. Two previously unreported SNPs were implicated in this analysis (Gialluisi et al., 2014). In 2016, however, a separate study evaluated a series of SNPs which had shown significant associations in previous reading- and language-related GWAS, and all failed to replicate at the Bonferroni-corrected level of significance (Carrion-Castillo et al., 2016). This was a clear indication that perhaps the results of earlier GWAS of reading were not reliable, and
it was suggested that future reading and language GWAS projects should make use of larger sample sizes through meta-analyses (Carrion-Castillo et al., 2016).

Following this, there have been a series of meta-analyses of quantitative reading-related traits in an attempt to overcome the previous problems with lack of replication, however in some cases sample sizes remained small. For example, one cross-country meta-analysis of a series of reading related traits had a total maximal sample size of 3468 (Gialluisi et al., 2019). Only one suggestive SNP was identified, and this was for the rapid atomised naming of letters phenotype. Perhaps more interestingly, this study included further molecular genetic analysis; PGS for several traits, including ADHD and years of education, were found to be associated with some of the reading related measures, adding to the emerging picture of genetic pleiotropy (Gialluisi et al., 2019).

In the same year, a separate GWAS study was conducted on three reading-related traits, rapid automatised naming in letter and numbers and rapid alternating stimulus, in a relatively small Hispanic American and African American sample (N = 1331) (Truong et al., 2019). This represents one of the earliest studies of the molecular genetics of reading which has not been conducted in a sample with predominately European ancestry (Truong et al., 2019). Multivariate GWAS of these traits revealed one SNP of significance, which replicated in a small European sample (Truong et al., 2019).

A GWAS meta-analysis of quantitative word-reading ability in two North American samples did not return any significant SNPs, however several suggestive SNPs were identified (total N = 5054) (Price et al., 2020). Similarly to the Gialluisi et al. paper published in 2019, the most interesting contribution of this GWAS was the downstream analysis; a series of PGS analyses replicated and expanded upon those earlier findings, with PGS for ADHD, educational
attainment, and intelligence significantly predicting word-reading in one of the two samples (Price et al., 2020). Modifications of traditional GWAS approaches have also been used to study quantitative reading ability, including a study which used hypothesis-driven GWAS (Price et al., 2022).

By far the largest GWAS of quantitative reading ability was a meta-analysis of up to 34000 individuals, published in 2022 (Eising et al., 2022) (See Chapter 2 for more details). This study far surpasses previous efforts in terms of sample size, and interestingly, failed to replicate previous associations discovered (Eising et al., 2022). This indicates that while molecular genetic research on quantitative reading ability has been ongoing for some time, there are still fundamental questions about the genetics of reading uncovered by GWAS and the impact that different measures and sample sizes can have on results.

4.3. Linkage Disequilibrium Score Regression Genetic Correlations

The final method to be discussed in this chapter is that of Linkage Disequilibrium Score Regression (LDSC) based genetic correlations, also known as cross-trait LD score regression, which use the output from GWAS analysis to calculate genetic correlations (Bulik-Sullivan, Finucane, et al., 2015). The LD Score is a population-based measure of linkage disequilibrium for a given SNP, and is measured by summing the correlations of the target SNP and all other measured SNPs (Bulik-Sullivan, Loh, et al., 2015; Lee et al., 2018). The other parameter required to calculate LDSC based genetic correlations is the product of the z-scores (standardised effect size per SNP) for each of the two traits in a correlation pair (Bulik-Sullivan, Finucane, et al., 2015). The z-score product is regressed onto the LD score, and the slope of the regression provides a covariation estimate, from which the correlation coefficient is calculated (Bulik-Sullivan, Finucane, et al., 2015). The same foundation can also be used to estimate the SNP-
based heritability of a trait by including only values from one study (Bulik-Sullivan, Finucane, et al., 2015).

LDSC genetic correlations have several benefits; they do not require genetic data from individuals, and they are able to use information from the whole genome, which is useful when effect sizes are very small (Bulik-Sullivan, Finucane, et al., 2015). This makes it a good option for studies of reading ability, which appears to be influenced by many variants with small effect sizes (Meaburn et al., 2007). Additionally, the LDSC-based method can be used successfully and without bias in cases of sample overlap, as the statistical properties of the technique mean that any false associations caused by overlap would affect the LD Score intercept, but not the slope i.e., it would not affect the value of the correlation, which is represented by the slope (Bulik-Sullivan, Finucane, et al., 2015). Additionally, if sample overlap is known, certain parameters of the model can be adjusted to reflect this (Bulik-Sullivan, Finucane, et al., 2015). Genetic correlation estimates generated through LDSC are comparable to those produced by REML, an alternative method which requires raw genetic data for each trait (Bulik-Sullivan, Finucane, et al., 2015). Other work has indicated that LDSC may be slightly less accurate than GREML, however given that GREML requires full genotype data, the LDSC approach to regression was most appropriate in this case (Ni et al., 2018).

LDSC based genetic correlations have been used to identify genetic commonalities among other traits, including years of education and verbal-numerical reasoning (Hagenaars et al., 2016), and years of education and anorexia nervosa (Duncan et al., 2017). Strong genetic correlations may be indicative of pleiotropy (Hagenaars et al., 2016), and so the combination of this method and the data available provides a good opportunity to explore the genetic correlates of reading ability. Of relevance to this thesis, LDSC based genetic correlations have been used to
identify genetic overlap in the largest existing GWAS meta-analysis of language measures, and a range of traits measured in independent cohorts (Eising et al., 2022). Eising and colleagues (2022) reported consistent positive genetic correlations between reading ability and a range of education measures, including years of schooling, qualifications obtained, and age completed full time education. Alongside this, some negative genetic correlations were found between reading and wellbeing measures, including mood swings, fed-up feelings, and loneliness. A range of other correlations were found in the domains of cognitive ability, health and exercise, and pain; see Chapter 6 for more details.
4.4. References


reading in a novel independent dataset. *Genes, Brain and Behavior, 15*(6), 531–541.
https://doi.org/10.1111/gbb.12299

https://doi.org/10.1146/annurev-genom-083117-021602


SECTION THREE:
EMPIRICAL STUDIES
Chapter 5: Longitudinal Reading Measures and Genome Imputation in the National Child Development Study: Prospects for Future Reading Research

This section of the thesis presents three empirical research chapters which explore the genetic and phenotypic correlated of reading ability and associated health outcomes in the NCDS. The first of these chapters is a published paper which provides preparatory work for the remaining two. This chapter details the preparation of the NCDS genetic dataset for later genetic analysis. This includes imputation of the NCDS genetic data to the Haplotype Reference Consortium (HRC) r1.1 reference panel (McCarthy et al., 2016), work which has not been previously documented or published. This is significant as it allows the subsequent genetic analysis in this paper to be conducted on genetic data which has been imputed to a higher quality reference panel than the previous iteration of the data, which was imputed to the 1000 Genomes reference panel (Auton et al., 2015; Davies et al., 2015).

In addition, this paper presents a documentation and analysis of the reading related variables available in the NCDS dataset. This is significant as the NCDS contains a variety of variables across several ages, however variables are often not repeated across ages, making longitudinal research challenging. The paper presents methodology for a series of valid composites that can be used to measure reading ability at ages seven, 11, 16, and an overall composite showing reading ability using variables from ages seven, 11, 16, 23, and 33. This forms the basis of reading composites which will be used to explore the phenotypic and genetic associations between reading ability and other outcomes, with a focus on health outcomes, in Chapters 6 and 7.
The paper presented here was published in the journal Twin Research and Human Genetics in 2023 and is an Open Access article.
5.1. Published Work
Longitudinal Reading Measures and Genome Imputation in the National Child Development Study: Prospects for Future Reading Research

Elinor C. Bridges1, N. William Rayner2,3,4, Hayley S. Mountford2, Timothy C. Bates1 and Michelle Luciano1

1School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom; 2Institute of Translational Genomics, Helmholtz Zentrum München — German Research Center for Environmental Health, Neuherberg, Germany; 3Oxford Centre for Diabetes, Endocrinology and Metabolism, 4Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom; 5Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

Abstract

Reading difficulties are prevalent worldwide, including in economically developed countries, and are associated with low academic achievement and unemployment. Longitudinal studies have identified several early childhood predictors of reading ability, but studies frequently lack genotype data that would enable testing of predictors with heritable influences. The National Child Development Study (NCDS) is a UK birth cohort study containing direct reading skill variables at every data collection wave from age 7 years through to adulthood with a subsample (final n = 6431) for whom modern genotype data are available. It is one of the longest running UK cohort studies for which genotyped data are currently available and is a rich dataset with excellent potential for future phenotypic and gene-by-environment interaction studies in reading. Here, we carry out imputation of the genotype data to the Haplotype Reference Panel, an updated reference panel that offers greater imputation quality. Guiding phenotype choice, we report a principal components analysis of nine reading variables, yielding a composite measure of reading ability in the genotyped sample. We include recommendations for use of composite scores and the most reliable variables for use during childhood when conducting longitudinal, genetically sensitive analyses of reading ability.

Keywords: Reading; genotype imputation; cohort study; principal components analysis; longitudinal measures

(Received 9 September 2022; revised 30 December 2022; accepted 17 January 2023; First Published online 10 March 2023)

Poor functional reading skill is a global problem with high prevalence, even in economically developed countries. For example, 21% of the population of the United States have poor English literacy skills (Mamedova & Pawlovski, 2019), and 16.4% of the adult population in England struggle with reading unfamiliar material (National Literacy Trust, ca. 2017). The annual global cost of poor literacy skills is estimated to be £800 billion (World Literacy Foundation, 2018). It is imperative to identify the causes of reading skill variation so that early identification and intervention in cases of potential reading disorder is possible, which can improve both word-reading and comprehension skills (Snowling & Hulme, 2011). This could reduce related negative outcomes, such as lower educational attainment and unemployment (Carrie & Thomas, 1999). Cohort studies are particularly valuable for determining the causes of reading ability if they have extensive longitudinal data from birth and linked genetic data. Here, we outline a UK birth cohort study, the National Child Development Study (NCDS), that has a wide range of reading measures in childhood and measures of functional reading ability in adulthood, along with genotypic data in a large subsample. Data collection began in 1958, and there is a rich variety of environmental variables collected at all age points. We present a series of carefully selected reading measures that we suggest for future studies, along with the protocol for imputation of the genotype data. The imputed data is available for researchers to access from the Centre for Longitudinal Studies.

Longitudinal studies of reading are required to build a complete picture of predictors of reading ability from childhood through to adulthood. The majority of existing longitudinal reading studies have been conducted across short time periods in early childhood (Pyridou et al., 2020). This approach overlooks two reading trajectories: those who read well in early childhood and begin to develop reading problems later, and those who have difficulties in acquiring reading skills in early childhood but go on to read normally (Catts et al., 2013). This provides clear incentive for use of longitudinal reading datasets that extend beyond early childhood. Additionally, several studies have identified that environmental circumstances in early childhood, such as socioeconomic status (SES) and maternal education, are associated with reading outcomes years later (e.g., Russell et al., 2016; Senechal, 2006; Williams & Silva, 1993). While these existing studies have generated insight into early predictors, they do not reach adulthood and...
potential lifespan sequelae of reading difficulty. In general, reading fluency appears to stabilize around adolescence (Lohvansuu et al., 2021) with little or no changes during middle age (Reder, 2012). However, small to moderate gains can still be made in adulthood with targeted intervention (Sabatini et al., 2011), and individual differences in rate of change have been identified (Lechner et al., 2021), so it is important to track the trajectory of reading skill over the lifespan. The NCDS contains reading skill data for a large, genotyped subsample up to and including adulthood, allowing longer term longitudinal studies to be conducted.

Comprehensive reading studies should use environmental data to build a full picture of predictors. Some identified predictors of reading ability are physiological in nature (e.g., Leppänen et al., 2010; Lohvansuu et al., 2021), but the majority of early known predictors involve either the Home Literacy Environment (HLE) or home and family circumstances more broadly. For example, association of low SES with poorer reading related skills has replicated across studies and cultures (e.g., Fernald et al., 2012; Fung & Chung, 2019; Molfese et al., 2003; Russell et al., 2016; Williams & Silva, 1985), and literacy activities and experiences gained in the home have repeatedly predicted reading related skills (e.g., Hamilton et al., 2016; Russell et al., 2016; Senechal, 2006). Many predictors of reading ability, including housing tenure, single parenthood, maternal education, and reading to the child (Russell et al., 2016) are present in the NCDS.

While these studies each go some way to untangling the nature of the predictors of reading ability, each fails to control for potential genetic confounding. This prevents us from understanding whether the aforementioned environmental variables are causal (Kendler & Baker, 2006). Researchers have provided clear evidence that genetically informed reading studies are required to complete the picture. For example, one study has shown that maternal reading and language skills are positively correlated with storybook exposure in the home, and once these skills are controlled for, storybook exposure no longer acts as a significant predictor of childhood reading, spelling and language skill (Puglisi et al., 2017). This suggests a heritable component acting on childhood reading skill, demonstrated by the maintained association of the mother’s skill level with the skills of the child. The nature of these associations can be clarified by including genotype data for genetically informed reading studies, which can be used to control for genetic confounding.

Twin studies confirm that reading ability has a substantial heritable component (e.g., see Bates et al., 2007; Byrne et al., 2005; Wadsworth et al., 2007) with an estimated heritability of approximately .54-.73 in adolescents (Bates et al., 2004; Wadsworth et al., 2015). In recent years, genomewide association studies (GWAS) have been conducted to identify specific genetic variants that are associated with reading skill (Luciano & Bates, 2019). The largest of these was in ~34,000 individuals and identified one locus of genomewide significance associated with quantitative word-reading, along with a significant SNP-based heritability estimate (Eising et al., 2022). The largest GWAS of dyslexia, which might be considered the low extreme of reading skill (Hulme & Snowling, 2016), has identified 42 associated independent variants (Doust et al., 2022). The NCDS dataset has the potential to contribute to GWAS meta-analysis, increasing sample sizes and allowing for greater power (Panagiotou et al., 2013).

The influence of genetics on reading ability is an area that has been explored using twin cohort studies. Many of these studies point towards stability in the genetics of reading ability in childhood and adolescence. This is a finding that has been replicated with different methods. For example, use of Cholesky decomposition models have shown high genetic correlations across late childhood and adolescence for reading ability (Betjemann et al., 2007; Wadsworth et al., 2001). Use of DeFries-Fulkner regression has indicated that persisting reading difficulties are due to the same genetic components, and between 60–75% of stability in reading difficulties may be due to genetic influence (Astrom et al., 2007; Wadsworth et al., 2015, 2016). Similarly, latent growth models in twins have shown that genetic influence on reading ability at the age of 6 is related to reading performance through to age 8 (Petrelli et al., 2010), and has also indicated that no new genetic factors appear for reading ability between the ages of 6 and 12 (Logan et al., 2013), although a longitudinal genetic model from Erb et al. (2017) showed a new genetic factor appearing after kindergarten, which influences word reading at first grade and comprehension at seventh grade. Similarly, Ebeyer et al. (2010) showed that a second genetic factor becomes active on reading ability in Grades 1 and 2, compared to kindergarten. A similar result was demonstrated by Samuelsson et al. (2008) in samples from Scandinavia and the United States; however, no additional genetic factor was found in an Australian sample. The authors suggest this may be due to international differences in schooling. These discrepancies show that while we are starting to build a picture of the changing influence of genes on reading over the life-course, most studies have focused on childhood, with a few moving into adolescence, and they have not utilised modern molecular methods.

Data from genotyped individuals can be used to ask and answer more complex questions about prediction of reading ability. For example, the data can be used for polygenic prediction, in which the additive variants of an individual’s genome can be combined to indicate their genetic propensity for high reading ability (Luciano, 2017). A polygenic score (PGS) based on dyslexia explained up to 6% of variance in quantitative reading measures in samples both enriched and not enriched for poor reading (Doust et al., 2022). Selzam et al. (2017) showed that the proportion of variance in reading ability explained by an educational attainment PGS increased with the age of the child, highlighting the importance of longitudinal data in genetic studies. In another study, PGSs calculated for intelligence, educational attainment, attention deficit hyperactivity disorder (ADHD) and bipolar disorder were found to be significantly correlated with word reading ability (Price et al., 2020). PGS can further be used for gene by environment interaction (G × E) research over the life course. Whereas generation of PGS for reading ability for gene-environment interplay research has not yet been conducted, the phenotype of educational attainment provides an example; one study has found that environmental factors mediated approximately 40% of the impact of an educational attainment PGS (Allegrini et al., 2020).

In sum, the NCDS 1958 Birth Cohort contains a range of reading and reading related measures throughout childhood, and follow-up variables focusing on functional literacy in adulthood, providing a valuable resource for reading studies. The sizeable subsample for which genetic data is available, along with the richness of other variables collected through the life-course, makes it a potentially valuable resource for genetically sensitive longitudinal reading studies that has been underutilised thus far. The imputed genotyped subsample has, to date, only been available with the 1000 Genomes reference panel (Davies et al., 2015). In this article, we present the NCDS’s genetic dataset imputed using the more recent Haploype Reference Consortium r1.1 panel, which offers greater imputation quality (Haploype Reference Consortium, 2016). Conducting genetic research with NCDS genotype data is...
challenging, due to the genotyping of data on multiple chips and a lack of centralised documentation. To aid in future studies, we clearly demonstrate the quality control and processes that took place prior to imputation, so that future researchers may use this updated resource in their own work. In addition to preparation of genetic data, this article presents an overview of reading and reading-related measures available in this dataset from age 7 to age 33. Matched genetic and phenotypic data must be requested from the data holders (Centre for Longitudinal Studies, n.d.-a). The full sample of NCDS phenotypic data is openly available through the UK Data Service (UK Data Service, n.d.).

Material and Methods

About the NCDS Dataset

The NCDS, also known as the 1958 British Birth Cohort, is a national cohort study that surveyed the parents of babies born during one week in the year 1958 (Power & Elliot, 2005). There have been several follow-ups, which are still ongoing (Centre for Longitudinal Studies, n.d.-c), and the length of time for which this cohort study has been running makes this an excellent dataset for tracing longitudinal patterns and associations. Data for 17,416 births were collected, with 9137 respondents at the last completed wave at age 55 (Centre for Longitudinal Studies, n.d.-b). For a subsample of the participants, a biomedical survey was also conducted at age 44, which included the collection of DNA samples (Power & Elliot, 2005). This article presents the preparation and imputation of 13,738 overlapping genotyped samples (resulting in a final total of 6431 unique individuals) collected from the NCDS participants on seven different arrays (Table 1). This imputed dataset is available for researchers to access from the Centre for Longitudinal Studies. Access to linked phenotype and genetic data is dependent on research proposal approval.

Genotyped Data

Quality Control Procedure

All quality control was conducted using Plink v1.90b4, R v3.3.2 and RStudio v4.1.2. Code used to carry out these processes, along with further information regarding externally developed scripts and resources, is publicly available as a GitHub repository (Bridges, 2022). Quality control was carried out on each of the seven genetic datasets according to the following steps. SNPs with a call-rate of less than 98% were removed (Turner et al., 2011), and SNPs that were not in Hardy-Weinberg equilibrium were removed at a threshold of $p < 1 \times 10^{-6}$. Individuals failing quality control were also removed. Those with a genotyping call-rate of less than 97% were then removed (Wellcome Trust Case Control Consortium, 2007). Individuals showing unexpected levels of heterozygosity (+/- 3 SD from the mean) were removed.

Each dataset was updated to GRCh37 build for consistency, using up-to-date strand files and a series of commands collated in the script Update Build (Robertson, 2012; see Supplementary Information A for further detail). Heterozygous haplotype errors were present in six of the seven chips. They were removed by first ensuring that all variants in the pseudo-autosomal region had the correct chromosome code for the Plink format, and any remaining errors were set to missing. Any individuals with discrepant or ambiguous sex data were removed, excluding the Affymetrix 500K chip, for which X chromosome data was not available. A small number ($n = 8$) of related samples (representing 6 individuals) were identified using a relatedness threshold of 0.1875, and the sample with the most missing data from each related pair was removed. The computational burden to control for genomic relatedness of six cases was deemed too high to warrant their inclusion, and has the advantage that users will not need to check for relatedness prior to their own analysis.

To identify genetic ancestry outliers, the data was merged with 1000 Genomes reference data (The 1000 Genomes Project Consortium, 2015), and Principal Components Analysis (PCA) was conducted in Plink to identify outliers ($n = 17$). Outliers were removed according to the procedure documented by Meyer (2021a, 2021b), using a theta value of 3.

Data Preparation and Imputation

The cleaned data were checked against the HRC reference panel r1.1 site list (HaploType Reference Consortium, 2016) for strand issues, using a specifically developed Perl script named HRC-1000G-check-bim.pl, v.4.3.0 (Rayner, 2020). This script checks for strand inconsistencies between the data and the reference panel, and generates a series of Plink commands that can then be used to remove or update any problematic loci. Perl v5.24.0 was used, followed by Plink to make the aforementioned necessary changes. Sorted VCF files were generated for each chromosome using BCFtools (Danecek et al., 2021). A final set of quality control checks were carried out on the VCF files post- conversion, using a specifically developed Python script (Zhan & Liu, 2016). These checks include identifying duplicated sites, invalid genotypes and NonSNP sites. Python v2.7.10 was used. No issues requiring further action were identified. The data were uploaded to the Michigan Imputation Server for quality control and imputation (Das et al., 2016). Imputation was carried out against the European population of the HRC r1.1 2016 reference panel, using the Minimac4 pipeline. Eagle v2 phasing was used.

Post Imputation Quality Control

Following imputation, quality control checks were carried out using the ‘ic’ script developed by Rayner (2016), which incorporates a series of Perl commands to check alternate allele frequencies and imputation quality (Rayner, 2016). This output can be used to evaluate the quality of an imputed dataset, including alternate allele frequency counts and frequency counts of the $R^2$ imputation quality score by chromosome.

There was considerable overlap of individuals between arrays, with many individuals having been genotyped on multiple arrays. After imputation, we determined the number of unique samples.

### Table 1. Breakdown of number of participants sampled on each array in the NCDS after removal of exclusions and duplications, and number of SNPs sequenced in each dataset

<table>
<thead>
<tr>
<th>Array</th>
<th>N</th>
<th>SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illumina 1.2M</td>
<td>2908</td>
<td>1157986</td>
</tr>
<tr>
<td>Illumina 15k Custom Chip</td>
<td>1475</td>
<td>9803</td>
</tr>
<tr>
<td>Illumina Human 660 Quad</td>
<td>871</td>
<td>582892</td>
</tr>
<tr>
<td>Infinium HumanHap 550K v1.1</td>
<td>1436</td>
<td>555174</td>
</tr>
<tr>
<td>Infinium HumanHap 550K v3</td>
<td>2592</td>
<td>561303</td>
</tr>
<tr>
<td>Affymetrix 500k</td>
<td>1477</td>
<td>490032</td>
</tr>
<tr>
<td>Affymetrix v6</td>
<td>2979</td>
<td>934967</td>
</tr>
</tbody>
</table>

[https://doi.org/10.1017/978203.2 Published online by Cambridge University Press](https://doi.org/10.1017/978203.2)
that should be retained on each array by retaining the highest quality sample for each individual (Table S16). This was determined by listing all samples in order of chip quality, determined by the number of SNPs with an information score >0.8 after imputation (Al-Soufi et al., 2021). The values in Table S16 represent the number of individuals that were retained on each chip after removal of lower quality duplications. It should be noted that duplicates were not removed in the final imputed dataset that is available for researchers.

Phenotype Data

Measures

Age 7. Three reading measures are available at age 7 in the NCDS.

1. The Southgate Group Reading Test is a 30-item word reading test that assists in the identification of participants with poor reading skills (Shepherd, 2012). The test was distributed in school by the participant’s teacher, and one point was awarded for each correct answer (Shepherd, 2012). Of the 30 questions, 16 required the child to correctly circle the word from a list that corresponded to an image, and fourteen required the child to correctly circle the word from a list that corresponded to a word read aloud by their teacher (Shepherd, 2012). Little further information is available about the test; however, a small amount of research on reliability has been conducted. The Southgate Test manual reports parallel reliability of 0.95 (Southgate, 1958, as cited in Tizard et al., 2002). One study has found that the Southgate reading test shows relatively high concurrent validity, as shown by a correlation of .72, with the Concepts About Print test, which measures several early reading-related skills such as book orientation and the relationship between written and oral language (Sultmann et al., 1985).

2. The child’s usual teacher was asked to rate the child’s reading ability compared to other children of their own age on a 5-point scale (Teacher Rating). The available options were Avid reader. Reads fluently and widely in relation to his age; Above average ability. Comprehends well what he reads; Average reader. Poor reader. Limited comprehension; Non-reader, or recognises very few words. (3) The child’s usual teacher was asked to identify which level of book the child had reached in a reading scheme (Reading Level). Response options included Don’t know or inapplicable; On prereading activities only; At present on Book 1 or introductory book; At present on Book 2; At present on Book 3; At present on Book 4; Beyond basic reading scheme. The option Don’t know or inapplicable was coded as missing (full sample N = 148, subsample n = 66), leaving a 6-point scale. No further information was available on this variable.

Age 11. Four reading related measures are available at age 11.

1. A 35-item reading comprehension test was administered by the participant’s school, based on the 1947 Watts-Vernon test of reading ability (Shepherd, 2012). For each item, the participant was asked to choose a word from a selection of five in order to complete a sentence, and one point was awarded for correct answer (Shepherd, 2012). (2) The child’s teacher was asked to rate the child’s use of books compared to other children their age on a 5-point scale (Book Use). Response options included Exceptional. Reads very widely for pleasure and information; Above average. Turns to books very readily; Average. Skill and comprehension satisfactory for school requirements; Below average. Still learning the skill of reading, not inclined to turn spontaneously to books for pleasure or information; Very poor or non-reader. Recognises few words, very limited use of books because of poor skill. (3) The participant was asked how often they read books outside of school work, and was given the option to respond with often (nearly every day), sometimes, or never or hardly ever (Reads Books). (4) The participant was asked how often they read magazines, newspapers and comics. The response options were often (nearly every day), sometimes, and never or hardly ever (Reads Other).

Age 16. There are five reading related measures available at age 16.

1. (1) The same reading comprehension test was administered as at age 11 (Shepherd, 2012). (2) The participant’s teacher was asked if the participant could read well enough to cope with everyday needs (Can Cope). This was presented with the options yes, no, and uncertain. Uncertain responses were removed from analysis in order to create a binary variable (full sample N = 20, subsample n = 8). (3) The participant was asked whether they often read books outside of school work, and was provided with the categorical response options often, sometimes, never or hardly ever, or like to but no chance. Like to but no chance (full sample N = 388, subsample n = 152) was removed from analysis due to its categorical nature, leaving three ordinal categories (Reads Books). (4) The participant’s teacher was asked to rate the English ability of the participant from the options Capable of obtaining an A-level or Higher-grade pass in this subject; Above average. Capable of obtaining O-level or O-grade or CSE grade one; Of average ability in this subject. Capable of obtaining a CSE pass, grades 2–4; Below average. A possible CSE entrant; Little, if any, ability in this subject; Don’t know (English Ability). Those who responded Don’t know were set to missing to ensure an ordinal set of responses (full sample N = 32, subsample n = 17). (5) The participant was asked to rate their ability in English compared to other people of their age, from the options never studied, below average, average and above average (English Rating). Never studied responses were removed (full sample N = 61, subsample n = 24).

Age 23. At age 23 all participants were asked if they had had problems with reading since they left school (Reading Problems). Possible response options were yes, no and don’t know. Don’t know responses were removed from analysis (full sample N = 27, subsample n = 17). The follow-up question was a binary response variable asking whether their problems made things difficult in everyday life. This question was excluded from analysis because data were limited by the screening question. An open text question allowed participants to elaborate on what difficulties they faced. Participants were also asked whether they had attended any courses to improve their skills. This variable was also excluded in this analysis, for two reasons; first, limited data due to the screening question, and second, attendance at a course may not be representative of need as there are multiple practical barriers to attendance, including lack of temporal and financial resources (BSA, 2000, as cited by Melrose, 2014).

Age 33. At age 33, participants were again asked whether they had had problems with reading since they left school. If they responded yes, they were asked a series of follow-up questions regarding which common activities they had difficulty with due to their reading problems. Respondents were asked whether they could usually read and understand what is written in a newspaper or magazine; a letter; and paperwork or forms. Respondents were also asked whether they could read aloud to a child from a children’s book. Response options included yes, easily; yes, with difficulty; and no. Respondents were asked about which aspects of reading they found difficult, and whether they had attended any courses to improve their skills. As for age 23, these follow-up questions were excluded from analysis for the same reasons.

Age 42 and Beyond. Reading phenotypes are available beyond age 33 in this dataset. Our analysis did not include variables beyond.
age 33 because maximal reading skill is achieved by early adulthood, with research showing that mean literacy scores in population samples are unlikely to change significantly beyond the age of 34 (Lechner et al., 2021). For completeness, we briefly describe the reading variables available. At age 42, participants were asked whether they had done any courses to improve their reading since their last NCDS interview, and, if so, how many. Age 42 reading variables also included several binary response variables, such as Are you currently on a course to improve reading?, Have you ever wanted to improve your reading?, and Do you feel confident about helping your children (en) with reading? In addition to this, participants were asked whether they could usually read and understand what is written in a magazine or newspaper, whether they can read aloud to a child from a child’s storybook, and whether they can usually read and understand any paperwork or forms you would have to deal with in a job? For each of these questions, the respondent had the option of responding yes or no. Each time the respondent answered yes, they were asked whether they could read this easily or with difficulty. Participants were also asked whether their ability to read any paperwork or forms you must deal with has improved or got worse over the last 10 years? Response options included improved, got worse, and stayed the same. Other reading-related variables exist at age 42; however, we chose not to include them here due to lack of specificity, with many of the variables also encompassing writing and mathematics skills. No reading variables were collected with the biomedical subsample at age 44.

Several reading variables were collected at age 46, including how many reading courses, if any, the participant had attended since their last NCDS interview, and a binary response question asking whether they would like to improve their reading skills. Participants were asked how often they read magazines or newspapers for enjoyment, and given six ordinal response options, ranging from never to every day. Respondents were asked the same question with regard to how often they read books. No reading measures were available at ages 50 or 55.

Analysis

All phenotypic analysis was carried out in RStudio v4.1.2. Analysis was carried out separately in both the full data sample from UK Data Service (N = 18,558), and the quality controlled subsample from the biological survey after removal of duplicated individual samples (n = 6431), to allow for comparison. All data for which valid phenotypic values were present were used. Descriptive statistics were calculated for selected variables, as described above. Correlations were generated for all continuous, ordinal and binary variables using the hetcor function in the polycor R package, which is able to determine which pairs of variables are to be correlated using Pearson, polyserial or polychoric methods based on variable type (Fox & Dusa, 2022). All complete pairs were used for analysis. It has been previously stated that correlations across time are a valid method for assessing stability of reading ability (Hulslander et al., 2010). Horn’s parallel analysis of principal components was carried out to identify the most appropriate number of components for these variables (Horn, 1965; Franklin et al., 1995). Five thousand iterations were used. We opted to conduct the analysis with three PCs as recommended due to the range of reading variables, as we expected all variables to be related to a broader reading construct, but not necessarily to reading skill. Use of multiple PCs would allow us to assess which variables had high scores on the first PC to ensure that we were, in fact, capturing reading ability. Following this, PCA with oblimin rotation was conducted to assess which reading variables could be combined to form general and age-specific composites (Song et al., 2013). PCA was carried out using the principal function in the psych R Package ( Revelle, 2022), using the correlation matrices generated previously as input. All aforementioned reading variables between ages 7 to 33 were included in the correlations and therefore the PCA.

Following inspection of PCA results, variables were selected for retention in a composite at a threshold of > .45 (Comrey & Lee, 1992, as cited in Finch et al., 2017, p. 1364). The variables retained in the full sample were used to calculate weighted reading composite scores for each individual. This included calculation of an overall composite, comprising of all retained measures, and a composite for use at age 7, 11 and 16. The scores calculated from the full PCA were used in order to weight the individual measures by their contribution to the reading ability component that spanned different measures and times. To achieve this, a z score was calculated for each data point, and these values were multiplied by the corresponding loading, before being summed to provide the composite (Bridges, 2022). The polycor package was then used to correlate the age specific childhood reading composites with adulthood reading variables, in order to assess their validity.

Results

Tables 2 and 3 show the number and percentage of SNPs and individuals respectively that were removed from each array during the quality control process. The percentage of SNPs removed from each chip was low (maximum 4.88%), as was the percentage of individuals removed, which ranged from 0.15% to 9.0%. Imputation was completed successfully for all seven arrays. The number and percentage of variants in each chip with a R2 score > .8 can be found in Table 4 (Al-Soufi et al., 2021; see Supplementary Material B for full breakdown).

Basic descriptive statistics were generated for each of the key variables in both the full sample and the subsample. Breakdown of all noncontinuous variables is available in Supplementary Information C. The proportion of respondents selecting each binary and ordinal response option was similar in both the full sample and the subsample for all questions. Rates of teacher-reported difficulty in coping with reading were low in both the full sample (n = 203) and the subsample (n = 40; Figure S7). A similar trend was found in adulthood self-reported reading difficulties; the number of self-reported reading difficulties at age 23 was low in the full sample (n = 497) and subsample (n = 169). This

Table 2. Number and overall percentage of SNPs that were removed from each array due to failure to pass quality control steps

<table>
<thead>
<tr>
<th>Array</th>
<th>Low call rate SNPs</th>
<th>Failed HWE</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illumina 1.2M</td>
<td>11893</td>
<td>1.03</td>
<td>11893</td>
</tr>
<tr>
<td>Illumina 15k Custom Chip</td>
<td>220</td>
<td>2.24</td>
<td>0</td>
</tr>
<tr>
<td>Illumina Human 660 Quad</td>
<td>5024</td>
<td>0.86</td>
<td>680</td>
</tr>
<tr>
<td>Infinium 550k v1.1</td>
<td>15549</td>
<td>2.80</td>
<td>850</td>
</tr>
<tr>
<td>Infinium 550k v3</td>
<td>2291</td>
<td>0.41</td>
<td>2340</td>
</tr>
<tr>
<td>Affymetrix 500</td>
<td>21415</td>
<td>4.37</td>
<td>2503</td>
</tr>
<tr>
<td>Affymetrix v6</td>
<td>12638</td>
<td>1.35</td>
<td>0</td>
</tr>
</tbody>
</table>

https://doi.org/10.1017/thg.2023.2 Published online by Cambridge University Press
20. A chi-square goodness-of-fit test can be found in Supplementary Table S15. All three reading tests showed similar distributions, with slightly higher means and medians in the subsample (Table 5). A chi-square goodness-of-fit test using the Southgate Group Reading test scores showed that the difference in distribution between the full sample (mean = 23.34, SD = 7.14) and subsample (mean = 24.20, SD = 6.48) was significant \( (p = 6.216 \times 10^{-12}, df = 30); \) however, the large sample size meant that very slight differences could be detected. Correlations between all variables ranged from very weak to very strong, with clear clusters emerging of similar variables (Figure 1). Correlations for variables assessing frequency of book reading at different ages, and English Rating showed the weakest overall correlations with the remaining variables, suggesting that these may not be appropriate for inclusion in the composite. Southgate, Teacher Rating, Reading Level, Book Use, Comprehension at both ages, English Ability, and Can Cope correlated particularly strongly with each other. All of these variables show moderate-to-strong correlations with self-reported reading difficulties at age 23 and 33, confirming the validity of these measures.

Results were similar for the PCA conducted in the full sample and the subsample (Table 6). Horn’s parallel analysis indicated that a three PC solution was the most appropriate (Figure 2). Intercorrelation coefficients were low (Table 7). While loadings differed between the full and subsample, the overall trends were consistent. The full sample results were used to generate composite scores to reduce any selection bias present in the subsample.

PC1 appears to be representative of reading skill level in the participant. In the full sample, PC1 explained 66% of the variance. Ten variables were retained (Table 6). Cronbach’s alpha was calculated on the standardized variables weighed by PCA scores that were used to calculate the overall reading composite, and composites at age 7, 11 and 16 (Table 8). All composites were in the acceptable range of \( >0.7 \) (Taber, 2017), except for the composite at age 16 comprised of Comprehension, Can Cope and English Ability, which fell below this value. Removal of each variable in turn showed that an acceptable Cronbach’s alpha could be achieved by removing Can Cope from the data.

Following this, the three age-specific composites were correlated with each other and with Can Cope, and reading difficulties at age 23 and 33 in both samples (Figure 3). All correlations were moderate to strong, again confirming the validity of these composites and these measures over time. It should be noted that Can Cope had stronger correlations with adulthood reading difficulties (Reading Problems 23 and Reading Problems 33) than the age 16 composite did; however, the age 16 composite showed greater correlations with childhood reading difficulties. Correlation coefficients were similar in all cases, suggesting either measure may be used. However, in the subsample, only 40 participants were believed not to be able to read enough to cope, making this a very small sample size for statistical analysis. As a result, we recommend

---

**Table 3.** Number and overall percentage of individuals removed from each array due to failure to pass quality control steps.

<table>
<thead>
<tr>
<th>Array</th>
<th>Missing data</th>
<th>Unexpected heterozygosity</th>
<th>Sex errors</th>
<th>Related individuals</th>
<th>Ancestry outliers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>Illumina 1.2M</td>
<td>74</td>
<td>2.54</td>
<td>55</td>
<td>1.89</td>
<td>1</td>
<td>0.10</td>
</tr>
<tr>
<td>Illumina 15K Custom Chip</td>
<td>10</td>
<td>0.68</td>
<td>10</td>
<td>0.68</td>
<td>122</td>
<td>8.27</td>
</tr>
<tr>
<td>Illumina Human 660 Quad</td>
<td>8</td>
<td>0.92</td>
<td>7</td>
<td>0.80</td>
<td>7</td>
<td>0.80</td>
</tr>
<tr>
<td>Infinium 550k v1.1</td>
<td>54</td>
<td>3.79</td>
<td>25</td>
<td>1.74</td>
<td>6</td>
<td>0.42</td>
</tr>
<tr>
<td>Infinium 550k v3</td>
<td>5</td>
<td>0.39</td>
<td>26</td>
<td>1.69</td>
<td>3</td>
<td>0.12</td>
</tr>
<tr>
<td>Affymetrix 500K</td>
<td>0</td>
<td>0.00</td>
<td>14</td>
<td>0.95</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Affymetrix v6</td>
<td>61</td>
<td>2.05</td>
<td>65</td>
<td>2.18</td>
<td>1</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Note: * No sex chromosome data was available for the Affymetrix 500K.

**Table 4.** Percentage and number of SNPs with genomewide imputation R² scores greater than 0.8 for each array

<table>
<thead>
<tr>
<th>Array</th>
<th>% variants &gt; 0.8</th>
<th>Number of SNPs &gt; 0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illumina 1.2M</td>
<td>35.33</td>
<td>14,276,058</td>
</tr>
<tr>
<td>Illumina 15K Custom Chip</td>
<td>0.21</td>
<td>82,702</td>
</tr>
<tr>
<td>Illumina Human 660 Quad</td>
<td>29.77</td>
<td>120,266,652</td>
</tr>
<tr>
<td>Infinium 550k v1.1</td>
<td>31.28</td>
<td>126,354,492</td>
</tr>
<tr>
<td>Infinium 550k v3</td>
<td>33.13</td>
<td>133,865,322</td>
</tr>
<tr>
<td>Affymetrix 500K</td>
<td>26.40</td>
<td>103,265,722</td>
</tr>
<tr>
<td>Affymetrix v6</td>
<td>31.83</td>
<td>128,598,081</td>
</tr>
</tbody>
</table>

**Table 5.** Descriptive statistics of continuous reading variables in the full sample and matched genotyped subsample

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southgate Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sample</td>
<td>23.34</td>
<td>26.14</td>
<td>7.14</td>
<td>14929</td>
</tr>
<tr>
<td>Subsample</td>
<td>24.20</td>
<td>27.64</td>
<td>6.48</td>
<td>5821</td>
</tr>
<tr>
<td>Comprehension Test Age 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sample</td>
<td>15.98</td>
<td>16.29</td>
<td>6.29</td>
<td>14130</td>
</tr>
<tr>
<td>Subsample</td>
<td>16.71</td>
<td>17.56</td>
<td>5.96</td>
<td>5608</td>
</tr>
<tr>
<td>Comprehension Test Age 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sample</td>
<td>25.31</td>
<td>27.09</td>
<td>7.09</td>
<td>13986</td>
</tr>
<tr>
<td>Subsample</td>
<td>26.26</td>
<td>28.35</td>
<td>6.35</td>
<td>5000</td>
</tr>
</tbody>
</table>
components analysis
by principal components analysis of the full NCDS dataset

Table 6. Principal component loadings and communality from a principal components analysis

<table>
<thead>
<tr>
<th>Age</th>
<th>Variable</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>h²</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Southgate Test</td>
<td>.91</td>
<td>.07</td>
<td>.19</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Teacher Rating</td>
<td>.89</td>
<td>.94</td>
<td>.87</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Book Level</td>
<td>.90</td>
<td>.12</td>
<td>.20</td>
<td>69</td>
</tr>
<tr>
<td>11</td>
<td>Comprehension Test</td>
<td>.76</td>
<td>.17</td>
<td>.06</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Book Use</td>
<td>.77</td>
<td>.16</td>
<td>.09</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Reads Books</td>
<td>.09</td>
<td>.72</td>
<td>.02</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Reads Other</td>
<td>.10</td>
<td>.04</td>
<td>.81</td>
<td>60</td>
</tr>
<tr>
<td>16</td>
<td>Comprehension Test</td>
<td>.76</td>
<td>.15</td>
<td>.01</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Reads Books</td>
<td>.04</td>
<td>.87</td>
<td>.09</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Can Cope</td>
<td>.89</td>
<td>.00</td>
<td>.30</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>English Ability</td>
<td>.76</td>
<td>.19</td>
<td>.05</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>English Rating</td>
<td>.21</td>
<td>.39</td>
<td>.31</td>
<td>43</td>
</tr>
<tr>
<td>23</td>
<td>Reading Problems</td>
<td>.64</td>
<td>.02</td>
<td>.45</td>
<td>80</td>
</tr>
<tr>
<td>33</td>
<td>Reading Problems</td>
<td>.58</td>
<td>.01</td>
<td>.48</td>
<td>73</td>
</tr>
</tbody>
</table>

Note: *These variables passed the 0.45 threshold, and so were included in composite calculation.

Table 7. Intercorrelations of the three principal components generated by principal components analysis of the full NCDS dataset

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>1.00</td>
<td>.36</td>
<td>.30</td>
</tr>
<tr>
<td>PC2</td>
<td>.36</td>
<td>1.00</td>
<td>.18</td>
</tr>
<tr>
<td>PC3</td>
<td>.30</td>
<td>.18</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Fig. 1. Correlation Heatmap Between All Reading Variables in the Full Sample (A) and Subsample (B).

Fig. 2. Eigenvalues generated from a parallel analysis of principal components in the full NCDS sample suggesting that three components should be retained.

carefully weighing these options before choosing which one is right for a particular study.

Discussion and Conclusion

In this study, we have presented the quality control and data preparation process of seven arrays that form a large genetic dataset, which was matched to a subsample of a UK birth cohort study who have reading skill data in childhood and adulthood. We have imputed missing genotypes for all arrays and carried out post-imputation quality control checks. We have also identified and
variables (Figure S11) suggest that there is a high degree of reliability for further analysis. The high correlations between almost all varying-related variables we have investigated offer multiple options to be used to conduct further association analysis.

...show a much higher proportion of high quality variants, which can be noted that most of the samples collected on the Illumina 15k chip are also present on other, higher quality chips, meaning that failure to use this chip would only result in the loss of 16 samples (see Supplementary Information D). The remaining six chips have been shown that a high proportion of unknown variants reduces imputation quality, and so the low number of genotyped loci may be responsible for the small proportion of high quality variants observed (Gao et al., 2011). It should be noted that most of the samples collected on the Illumina 15k chip are also present on other, higher quality chips, meaning that failure to use this chip would only result in the loss of 16 samples (see Supplementary Information D). The remaining six chips show a much higher proportion of high quality variants, which can be used to conduct further association analysis.

The strong relationships between many of the reading and reading-related variables we have investigated offer multiple options for further analysis. The high correlations between almost all variables (Figure S11) suggest that there is a high degree of reliability between adulthood self-report and childhood teacher report and objective tests when it comes to reading ability. This has two benefits; first, it allows for data reduction when considering reading ability as a composite, avoiding multicollinearity issues and reducing noise due to measurement error; and second, it means that it is feasible to conduct longitudinal reading studies with this data up to adulthood. Additionally, the strength of correlations across the whole studied period provides further evidence that reading ability remains relatively constant as age increases (Reder, 2012). It should be noted that the reading composites we have generated are more informative of poorer reading, given that several of our variables, including the Southgate Test, Can Cope and adulthood self-report measures are intended to capture reading difficulties. While the Southgate Test results show scores from across the distribution, there is a ceiling effect.

The PCA revealed that one PC explained a substantial proportion of variance, again indicating a strong relationship between the different variables across different ages. The genotyped sample, which only included NCDS members still active in the study at age 44, has previously been reported to be less representative of the general population, including fewer poor childhood readers (Atherton et al., 2008). This likely explains the differences we found between PCA results on the full sample and subsample, and for this reason, we recommend projection of the PCA analysis from the full sample, which is more population representative, onto the subsample, rather than using subsample-specific analysis.

The 3PC solution was the most appropriate fit to the data. Upon further inspection, these PCs can be interpreted as three different reading-related constructs. The strongest loadings in PC1 are objective measures of reading ability, teacher ratings, and self-report at age 23 and 33. This suggests that these are the variables that should be included in a measure of reading ability, as they are the strongest indicators of skill. The strongest loadings in PC2 relate to how often the participant reads books at different ages, and this might reflect shared variance related to interest or enthusiasm for reading, a variable in itself that may be of interest...
to reporting of reading difficulties with nonbook reading.

This work has shown that the NCDS is a valuable resource for conducting genetic studies of reading ability, particularly those with a longitudinal element. The imputed arrays can be analyzed separately for GWAS meta-analysis, or alternatively, the data can be combined and the array can be used as a covariate for further analysis as a single dataset. The range of reading measures available in this dataset allow for exploration of genetic influence across the full continuum of reading ability in childhood and adolescence, and for reading difficulties over time. The NCDS is currently the longest running UK cohort study for which genotyped data is available, meaning that long-term outcomes of participants, including retirement and life expectancy, can be investigated in this dataset in relation to their phenotypic reading ability and genome. Additionally, the NCDS is an incredibly rich resource in many other domains, including physical and mental health, employment, leisure activities, life in the home and financial information. This matched data provides opportunities for gene-by-environment interplay studies, including from a longitudinal perspective. In particular, this data allows the investigation of potential G × E interactions that provide a more nuanced understanding of how environmental circumstances can influence reading skill across a spectrum of abilities, and across the life course.

To conclude, the NCDS is rich potential resource for future genetic studies of reading ability. An imputed and cleaned, matched genomic dataset is available, and there is a reliable and valid set of reading variables available, as evidenced by the consistency of the different measures of reading ability across types of variable and with increasing age. Finally, the incredible detail contained in the NCDS dataset will help to make it a valuable resource for genomic studies of reading ability in the years to come.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1077/thg.2023.2

Data availability. The full phenotypic dataset is openly available to researchers at the UK Data Service. The genotyped subsample and matched phenotypic data is available from the Centre for Longitudinal Studies at University College London, upon submission of a successful research proposal.

Acknowledgments. Thanks to Gill Davies for advising on quality control techniques for pre-imputation genetic data, and Vicki Jackson for providing a pre-imputation protocol that informed part of this work.

Financial support. This work was funded by a PhD studentship from the Economic and Social Research Council with the Scottish Graduate School of Social Sciences. Data governance was provided by the METADAC data access committee, funded by ESRC, Wellcome, and MRC (2015–2018; Grant Number MR/N01014X/1 2018–2020; Grant Number ES/S008349/1). This work made use of data and samples generated by the 1958 Birth Cohort (NCDS), which is managed by the Centre for Longitudinal Studies at the UCL Institute of Education, funded by the Economic and Social Research Council (grant number ES/M001660/1). Access to these resources was enabled via the Wellcome Trust and MRC. SфорВЯдRIS grant [108439Z/11/Z (2) (The 1958 Birth Cohort: Fostering new Opportunities for Research via Wider Access to Reliable Data and Samples). Before 2015 biomedical resources were maintained under the Wellcome Trust and Medical Research Council 58READIE Project (grant numbers WT095219M and G1101799). HSM was supported by the Biotechnology and Biological Sciences Research Council [BB/T000813/1].

Conflict of interest. We have no conflict of interest to disclose.

Ethical statement. Ethical approval for this project was granted by METADAC.

References


Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature, 447, 661 678. https://doi.org/10.1038/nature05911


PUBLISHED WORK ENDS
5.2. Discussion and Conclusion

The published paper presented in this chapter has shown that reading measures are relatively stable across age in the NCDS, and additionally, that a selection of reading related measures can be combined into weighted composite scores in order to capture variation of quantitative functional reading ability across the full continuum of reading skills. In addition, it has presented a fully updated version of the NCDS genetic dataset, imputed to the Haplotype Reference Consortium reference panel (McCarthy et al., 2016). Together, the genetic and phenotypic elements presented in this paper provide an excellent starting point for future genetic studies of reading research in this sample, of which there has only been one to date, and which relied on a previously imputed version of the genetic data with a lower quality reference panel (Davies et al., 2015; McCarthy et al., 2016).

There are several reasons why this is significant. To begin with, as detailed elsewhere in this thesis, there are many studies demonstrating early predictors of childhood reading ability, such as early language skills and SES (Russell et al., 2016). However, almost all of these studies have failed to control for genetic confounding, and given that the heritability of quantitative reading is significant (eg. Eising et al., 2022), this means that it is not possible to gain a full understanding of the predictors of reading ability without considering genetic influence. The genetic and reading data in the NDS that has been explored in this paper may be used to retest reported phenotypic predictors of reading while controlling for genetics, through the use of polygenic scores, which has been done for other psychological and psychiatric traits (eg. Cuevas et al., 2020; Plomin et al., 2022).

Latent factors for reading have previously been used in the NCDS, using some of the same variables reported here (eg. Ritchie & Bates, 2013; Basten et al., 2015). Neither of these
studies used reading variables from the complete age range from childhood (age seven) to adulthood (age 33), meaning that they were only able to capture reading variance up to a maximum of age 11. The NCDS is a rare resource due to the collection of reading ability through adolescence and into adulthood, and so the composites generated in this thesis take better advantage of this unique opportunity to explore reading variation across the developmental trajectory. In addition to this, the previously constructed latent factors did not compare reading score longitudinally, whereas this thesis has shown that reading is relatively stable from childhood into adulthood.

Most individual reading measures included within the composites offer a generalized view of reading ability, without allowing us to distinguish whether they are ultimately measuring decoding or comprehension skills. The exceptions are the Southgate Test at age 7, which measures decoding, and the Comprehension Tests at ages 11 and 16, however at each of these ages general reading measures are also included in the composite. As such, these reading measures should be considered as general measures of functional reading ability, rather than specifically measuring decoding or comprehension. The Simple View of Reading indicates that at least some level of decoding is required for reading (Gough & Tunmer, 1986), implying that a level of decoding may be assumed in all composite variables, however the content of the individual items indicate that the majority of variables are also measuring a level of comprehension, for example, at age 16 it is unlikely that an individual can read well enough to cope if they cannot comprehend what they are reading. All composites show strong correlations with each other, supporting existing evidence that decoding and comprehension are highly correlated (Garcia & Cain, 2014).
This chapter, then, has added to the current body of literature through preparation and presentation of genetic and reading based data which will be of value to future studies. The next chapter of this thesis will go on to utilize these findings, by using reading composite scores generated through a similar process as the phenotypic outcome for a genome-wide association study, using the genetic data imputed in this chapter.
5.3. Additional References

*Excluding those cited in published work*


Chapter 6: The Genetic Architecture of Reading in the National Child Development Study in Relation to Reading-Related and Health- and Lifestyle-Related Genotypes

The National Child Development Study (NCDS) is a longitudinal British birth cohort study with accompanying genetic data (see Chapters 3 and 5 for full details). The NCDS contains a variety of reading measures, including objective reading tests, teacher ratings and self-report measures. These reading measures have been shown to produce valid composites for reading at ages seven, 11, and 16, and an overall reading composite comprising of reading-related measures from ages seven, 11, 16, 23, and 33 (see Chapter 5 for full details). The aims of this research project are as follows: (1) understand whether the genetic architecture of continuous reading ability remains constant across the different composites (2) assess how the genetic architecture of reading in the NCDS correlates with that of reading related measures from other cohorts and (3) assess whether genetic correlations measured with NCDS reading ability composites replicate genetic correlations identified with reading ability from psychometric tests. This project uses molecular methods to do this; first, by utilising Genome Wide Association Studies (GWAS) on the NCDS reading composites, then, by using LD Score Regression (Bulik-Sullivan, Finucane, et al., 2015; Bulik-Sullivan, Loh, et al., 2015) to calculate genetic correlations with the other variables. This chapter will end with a discussion of the results.

6.1. Background

Selecting Reading Measures

Many large scale behaviour genetics reading studies measure the phenotype of continuous reading ability or reading-related skills through formal standardised testing (e.g.,
Gialluisi et al., 2014; Luciano et al., 2013; Eising et al., 2022; Price et al., 2020). While standardised reading tests have the benefit of being easily quantifiable, they are not always available or do not always provide the required range of data. The National Child Development Study contains several measures of reading at different age points, and includes many styles of measuring reading, such as formal tests, teacher report, and self-report. The Southgate Group Reading Test used at age seven in the NCDS was designed to catch poor readers, rather than to show the full extent of variation in reading ability, and so a ceiling effect is present (Shepherd, 2012). As a result, the only standardised word-reading test in the NCDS cohort has limited usefulness when wishing to carry out research on quantitative reading ability. This thesis has sought to show that combining the various reading measures available in the NCDS can provide valid and informative measures of reading variation, which may be used to carry out further reading research in a large cohort sample.

In order to assess whether the reading composites developed in this project are measuring a similar genetic construct to standardised reading measures, they are compared with genetic results for several specific reading and language related phenotypes. These are word reading, spelling, phoneme awareness, nonword reading, nonword repetition, and self-reported dyslexia diagnosis (full details of these constructs are available in the Methods section). In previous research, these have all been found to genetically correlate strongly with each other (Eising et al., 2022; Doust et al., 2022), and so strong correlation of these measures with the NCDS reading composites would be indicative that the NCDS composites are indeed reliably measuring reading, despite the use of several subjective measures in the production of the composites.
Molecular Genetic Methods

In the past decade, Genome Wide Association Studies (GWAS) have been used to investigate the genetics of various psychological and cognitive traits, including reading ability and reading-related traits (e.g., Luciano et al., 2013; Price et al., 2020; Truong et al., 2019; Eising et al., 2022). GWAS uses common Single Nucleotide Polymorphisms (SNPs) from across the whole genome, in order to identify areas of the genome which show an association with phenotypic outcomes (Uffelmann et al., 2021). In this way, GWAS analysis can be used to identify loci which may have an influence on reading ability. Previous studies have identified only a limited number of SNPs showing association or suggestive association with reading ability (Luciano et al., 2013; Price et al., 2020; Eising et al., 2022). GWAS require large sample sizes to carry sufficient power to detect associations (Chen et al., 2021), and GWAS of reading ability have historically suffered from small sample sizes, causing difficulty in finding significant genetic results (Eising et al., 2022).

The NCDS reading data exhibits a relatively low sample size when considering GWAS on a continuous trait, however, this does not mean that molecular methods cannot be utilised for reading research in this dataset. Other recent research has demonstrated that it is possible to use GWAS summary statistics from modest sample sizes to successfully conduct downstream analysis on quantitative reading traits (Price et al., 2020; Price et al., 2023). Such analysis may include calculation of Polygenic Risk Scores (PRS) or genetic correlations.

Genetic correlations can be used to investigate the extent to which there are similarities in the genetics that underpin polygenic traits (Bulik-Sullivan, Finucane, et al., 2015). Genetic correlations function in the same way as a Pearson’s correlation, ranging in value from -1 to 1, with values further from zero indicating a stronger association (Bulik-Sullivan, Finucane, et al.,...
The presence of significant genetic correlations can indicate several possibilities. Cases of high genetic correlation can indicate the presence of biological pleiotropy, in which a genetic variant influences the presentation of multiple phenotypes through biological processes (van Rheenen et al., 2019; Gratten et al., 2016). Alternatively, a genetic correlation may indicate that the correlated variants influence one of the phenotypes, which in turn then influences the second phenotype (van Rheenen et al., 2019; Gratten et al., 2016). Presence of genetic correlations indicates that at least part of any phenotypic correlation is due to genetic influence (Lee et al., 2018). It should be noted that there are cases in which genetic correlation can also be an indicator of linkage disequilibrium, in which two separate genes are located close together on the genome and are therefore statistically associated with both phenotypes (Lande, 1984; Chebib & Guillaume, 2021).

This study utilises Linkage Disequilibrium Score Regression (LDSC) based genetic correlation analysis, which is based on the SNP data available in the GWAS summary statistics (Bulik-Sullivan, Loh et al., 2015; Bulik-Sullivan, Finucane, et al., 2015). SNP based genetic correlations (rG) use information from the whole genome, and so can be used to assess how SNPs with relatively low heritability, including SNPs below significance level, covary together for different traits (Bulik-Sullivan, Finucane, et al., 2015). LD Score regression is a time and computationally effective technique for estimating genetic correlations, and only requires GWAS summary statistics, rather than individual level genome data, making it a valuable tool for conducting genetic correlation analysis on traits for which access to individual level data is not available (Ni et al., 2018).

*Genetic Correlations Literature*
Little is known about how the SNP-based genetic architecture of reading functions with age, with most studies in this area having been conducted on twins. Twin based research broadly indicates, with a few exceptions, that the genetic underpinnings of reading ability are relatively stable throughout childhood (Betjemann et al., 2007; Wadsworth et al., 2001; Logan et al., 2013; Samuelsson et al., 2008) (for a full discussion of this topic, refer to Chapter 5). There is now a very small amount of evidence suggesting that SNP-based genetic influence on reading ability is stable with age. In a meta-analysis of quantitative reading traits, Eising and colleagues (2022) correlated genetic architecture of word reading in cohorts made up of individuals aged twelve and above with those aged under twelve, and found high genetic correlations for word-reading ($r_G = 0.89, SE = 0.16$) and non-word reading ($r_G = 0.88, SE = 0.24$) using LDSC. In addition to this, another study found consistent genetic correlations between reading related variables measured regularly between the ages of seven and 13 (Shapland et al., 2021). Further research in this area is required in order to examine stability of SNPs at a more granular level, and extending beyond late childhood.

Existing behaviour genetics research using molecular methods on quantitative reading has focused primarily on validated reading related variables, and it has been argued that this approach may reduce confounding and increase ease of replication (Bates et al., 2004). Reading and language related measures used in existing GWAS research include word reading, non-word reading, phoneme awareness, spelling, non-word repetition, rapid automated naming, rapid automated stimulus, and rapid alternating stimulus (Luciano et al., 2013; Eising et al., 2022; Truong et al., 2019), along with the first principal component of a series of standardised reading and language measures (Gialluisi et al., 2014). Phenotypic research often indicates moderate to strong associations between these variables (eg. Shapiro et al., 2013; Caravolas et al., 2001;
Cunningham et al., 2020) or with dyslexia (eg. Wolf et al., 2002; Coleman et al., 2009), which will largely capture those performing at the lower end of a normal distribution of reading ability (Hulme & Snowling, 2016). Several molecular genetics studies have now indicated that reading-related skills show a shared, if not identical, genetic basis. Eising and colleagues (2022) found significant pairwise genetic correlations, as measured using LDSC, between all tested variables after correction for multiple testing, except for phoneme awareness and non-word repetition, which did not survive the correction. In addition to this, Doust and colleagues (2022) found that self-reported dyslexia diagnosis shows significant negative genetic correlations with all five GenLang measures. Additionally, another study with a much smaller sample size (N = 6453), using GREML analysis rather than LDSC, has reported positive genetic correlations between word reading, non-word reading, phonological awareness, listening comprehension and non-word repetition (Shapland et al., 2021). This is unsurprising, given that the same cohort was used as part of the GenLang meta-analysis, which found similar results (Eising et al., 2022). The same study showed genetic stability of a series of specific reading-related measures between ages 7 and 13 through genetic correlations (Shapland et al., 2021), and was able to use further analysis to show that reading fluency could be considered the “genetic core” of the range of reading-related variables tested (Shapland et al., 2021).

Along with genetic similarities between reading and language related variables, molecular genetics based research has also indicated evidence of pleiotropy between reading ability and a series of other traits relating to cognition, education, and psychiatric traits. An early genomic research piece, which used variance components analysis methods, found a high genetic correlation between reading and mathematics ability in a sample of approximately 3000 twin pairs (Davis et al., 2014). Evidence of pleiotropy extends to other cognitive traits, including
general cognitive ability. Molecular genetic analysis using the bivariate GCTA method has uncovered a high genetic correlation \( (r_G = .89, SE = 0.26) \) between general cognitive ability and reading ability (Trzaskowski et al., 2013). The high standard error in this instance may be in-part due to the relatively small sample size used in the study, as standard error calculations using this method are very sensitive to sample size (Visscher et al., 2014). It is possible that the exceptionally high genetic correlation reported here between reading and cognitive ability is due to the measures used; the cognitive ability measure was a composite of a mixture of verbal and non-verbal tests, and concern has been expressed that verbal IQ and reading ability cannot be properly separated as low verbal IQ can be a consequence of low reading ability (Stanovich, 1991; Snowling et al., 2020).

The majority of molecular genetics research investigating the genetic architecture of reading ability in relation to other traits has been conducted using the polygenic score method (PGS), in which genetic data can be used to calculate an individual’s propensity to a certain trait based on their SNP information (Choi et al., 2020). PGS research has repeatedly indicated evidence of pleiotropy for reading and general cognitive ability. Studies have shown that an independently generated PGS for intelligence can be used to predict reading ability, and conversely, a PGS for reading and spelling can predict verbal-numerical cognitive scores (Price et al, 2020; Luciano et al., 2017). A small amount of existing research has been conducted on the genetics of reading in the NCDS (Davies et al., 2015).

Alongside cognitive abilities, there is also evidence to show that the genetics of reading are associated with educational attainment. For example, Belsky and colleagues (2016) have reported that a PGS for educational attainment is associated with earlier development of reading skills, and a faster trajectory of improvement. PGS for educational attainment has been
associated with a higher level of reading ability in several other studies (Price et al., 2020; Gialluisi et al., 2019; Gialluisi et al., 2021; Selzam et al., 2017). Conversely, PGS for reading skills has been associated with going to college or university (Luciano et al., 2017).

Evidence has also indicated that genetics associated with reading ability may also be associated with a series of psychiatric traits. Polygenic scores for several psychiatric traits have been associated with lower reading ability, including ADHD (Price et al., 2020; Gialluisi et al., 2019; Gialluisi et al., 2021), bipolar disorder (Price et al., 2020; Gialluisi et al., 2019; Gialluisi et al., 2021), and schizophrenia (Gialluisi et al., 2021). The association between the genetics of reading ability and schizophrenia has not always been replicated (e.g., Price et al., 2020; Gialluisi et al., 2019), meaning genetic similarity between schizophrenia and reading ability is currently unclear. Other psychiatric traits, such as depression and autism spectrum disorder, have not shown genetic similarity with reading through the use of polygenic traits (Luciano et al., 2017; Price et al., 2020; Gialluisi et al., 2019; Gialluisi et al., 2021). An association has been reported between a PGS for a reading composite and poorer self-reported health (Luciano et al., 2017).

A small amount of existing research has used the LDSC method to estimate molecular genetic correlations between reading ability and other traits, although this work is in its infancy. An early study in this area with a modest sample size reported significant genetic correlations between a reading composite measure and childhood intelligence (rG = 0.40, SE = 0.17), and educational attainment (rG = 0.70, SE = 0.14) (Luciano et al., 2017). Two large, recent molecular genetics studies of reading support these findings (Doust et al., 2022; Eising et al., 2022). The study by Doust and colleagues (2022), which uses the phenotype of self-reported dyslexia diagnosis in adults (cases = 51,800, controls = 1,087,070), reported genetic correlations
between dyslexia and several health and lifestyle outcomes, calculated using the LDSC method. Genetic propensity for dyslexia was significantly associated with higher genetic propensity for poor occupational conditions (e.g., shift work, manual work), smoking and alcohol consumption, pain, life dissatisfaction scores, and ADHD. Lower propensity for dyslexia was associated with genetic propensity for better educational outcomes. Similar results were found in the Eising et al., study (2022) which assessed quantitative reading ability in up to 34,000 individuals, depending on reading measure. Negative, significant genetic correlations were found between reading ability and physical health outcomes, depression, and risky lifestyle choices (e.g., alcohol intake, driving too fast, cigarette smoking), and measures of emotional wellbeing (e.g., mood swings, fed up feelings). Reading ability showed positive genetic correlations with educational measures and physical measurements, such as height.

Most importantly for this study, one previous study has investigated genetic correlations between reading ability and a small number of other traits within the NCDS (Davies et al., 2015). This study selected one reading measure each from age seven, 11, and 16 to represent reading; the chosen variables were the Southgate Reading Test at age seven, and reading comprehension tests at ages 11 and 16. GREML was used to calculate genetic correlations with other variables from within the NCDS. Reading at age seven was not significantly genetically correlated with number of O-Levels (rG = 0.20, P = 0.30), a measure of educational attainment, however reading at age 11 was (rG = 0.61, P=0.02) (Davies et al., 2015). Composites of teacher reported ability showed a significant genetic correlation with number of O-Levels at both ages, and the age seven rating included the teacher rating of reading ability in the child, along with ability in general other areas. Reading at age 16 was significantly associated with number of O-Levels (rG = 0.53, P = 0.05), however teacher-rated ability in English, which is included in the reading composite
reported for this study, was not (rG = 0.45, P = 0.11). This study used the 1000 Genomes imputation of the genetic data.

It is hypothesised that the NCDS reading measures will be strongly genetically correlated with each other across age, in light of their high levels of phenotypic correlation as demonstrated in Chapter 5, and in support of previous literature referred to in this section. Additionally, it is hypothesised that the NCDS reading measures will correlate moderately to strongly with the set of additional reading-related measures tested from independent samples, demonstrating the same pattern of relationships as previous research using molecular genetics methods (e.g., Eising et al., 2022; Doust et al., 2022. It is also hypothesised that the NCDS reading measures will replicate genetic correlations with other variables as identified in large scale reading GWAS studies, including those reported by Eising and colleagues (2022) and Doust and colleagues (2022). This work will particularly build on earlier work on the genetics of reading in the NCDS reported by Davies and colleagues (2015), using valid reading composites to explore genetic correlations in a wider range of traits, using data from other GWAS studies. Support of the stated hypotheses will establish the NCDS reading measure composites as reliable and valid measures of reading skill.

6.2. Materials and Methods

Data and Measures

The National Child Development Study. The primary dataset used in this study is the National Child Development Study, which is described in detail in Chapters 3 and 5. As mentioned previously, at age 44 a biomedical survey was undertaken for a subsample of study participants. This included collection of genetic data on several genotyping arrays. The genetic
data was imputed to the Haplotype Reference Consortium (HRC) r1.1 reference panel according to the procedure described in Bridges and colleagues (2023) for use in this study. The HRC panel was chosen as it is an up-to-date reference panel which provides better imputation quality than alternatives (Haplotype Reference Consortium, 2016). Seven arrays were imputed in total, however, only five were used in this study due to high sample overlap between Affymetrix 500k and other arrays and low imputation quality and sample size of the Illumina 15k array (see Chapter 5 for full details). Affymetrix v6, Illumina 1.2m, Infinium v1.1, Infinium v3 and Illumina HumanHap Quad-660 arrays were used for this analysis.

Post-imputation, the full imputed datasets were checked for quality control using the ‘ic’ script (Rayner, 2016). The data were manually checked using BCFtools v1.16 to ensure that variants on the sex chromosomes were properly encoded (Danecek et al., 2021). The data were then filtered before downstream analysis. BCFtools v1.16 was used to create a new version of the genetic dataset containing only variants with an $R^2$ imputation quality score of > 0.8, to ensure only the highest quality variants were used (Davies et al., 2015), and a minor allele frequency (MAF) of > 0.05, to ensure that only common SNPs were included in analysis (Goswami et al., 2021).

The phenotypic measure in this study was reading ability, as measured at ages seven, 11, 16, 23, and 33. The reading measures differed at each age. Multiple reading measures exist at ages seven, eleven, and 16. For those ages, reading composite values were created using a similar method to that described in Bridges and colleagues, 2023. Modifications to that process are described below. An overall reading composite was also included, including data from all age points.
As demonstrated in Chapter 5, Bridges and colleagues (2023) found that several reading measures in the NCDS show a high level of internal consistency, as measured by Chronbach’s Alpha, and so can be used to create composite measures for reading. Composites were created for ages seven, 11, 16, and overall, using the reading variables determined appropriate at each age as described by Bridges and colleagues (2023). Z-scores were generated for each of the variables, which were multiplied by the principal component loadings found in Bridges and colleagues (2023) to weight the measures in the composite. A mean score was then taken for each individual, and only individuals who had answered all questions at the relevant age point were included (see Results for further information). This was to ensure that genetic correlations would be clearly measuring the difference between the composites, and would not be subject to bias due to certain questions not being answered.

**External Summary Statistics.** Along with the summary statistics generated from the NCDS GWAS, this study also included GWAS summary statistics generated on a range of reading-related measures. These measures were included in order to assess the genetic similarity between the NCDS reading composite developed here and other, more precise measures of reading-related skills. All but one of the external summary statistics were provided from meta-analyses by Eising and colleagues (2022). This project included a large scale GWAS meta-analysis of several reading and language related traits, measured quantitatively. The reading related skills analysed in these meta-analyses are defined as follows, with all wording quoted from Eising and colleagues, 2022, pg.2, Table 1:

- “Word reading: Number of correct words read aloud from a list in a time-restricted or unrestricted fashion
- Nonword reading: Number of nonwords read aloud correctly from a list in a time-restricted or unrestricted fashion
- Spelling: Number of words correctly spelled orally or in writing after being dictated as single words or in a sentence
- Phoneme awareness: Number of words correctly altered in phoneme deletion/elision and spoonerism tasks
- Nonword repetition: number of nonwords or phonemes repeated aloud correctly

An additional set of summary statistics was also included in this analysis, which investigated the phenotype of self-reported dyslexia diagnosis in a sample of individuals who had signed up for commercial genetic testing with the company 23andme (Doust et al., 2022).

Additional details regarding the summary statistics are available in Table 1. All external GWAS analyses had been adjusted for genomic control, which is intended to reduce artificial genomic inflation caused by sample bias (Devlin & Roeder, 1999; Bacanu & Roeder, 2000).

**Table 1:** Sample information about external summary statistics for a set of reading and language related variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Age Range</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Reading</td>
<td>5 – 26 years</td>
<td>33959</td>
</tr>
<tr>
<td>Spelling</td>
<td>5 – 26 years</td>
<td>18541</td>
</tr>
<tr>
<td>Phoneme awareness</td>
<td>5 – 18 years</td>
<td>13633</td>
</tr>
<tr>
<td>Nonword Reading</td>
<td>5 – 26 years</td>
<td>17984</td>
</tr>
<tr>
<td>Nonword Repetition</td>
<td>5 – 26 years</td>
<td>14046</td>
</tr>
<tr>
<td>Dyslexia</td>
<td>18+ years</td>
<td>1138870</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Cases = 51,800)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Controls = 1,087,070)</td>
</tr>
</tbody>
</table>
Analysis

GWAS Analysis. Due to the differential quality of the five arrays, we opted to analyse the five arrays separately and then meta-analyse the results, rather than combining the genetic data and analysing the combined dataset. GWAS was carried out on the filtered, imputed data files using reading composites and a series of additional covariates. Sex was added as a covariate, as recorded at birth and cross-checked with genetic data to ensure that all sex data was consistent (e.g., Eising et al., 2022). The first five ancestry principal components were also included as covariates, in order to correct for any population stratification (Price et al., 2006). Principal Components Analysis (PCA) was conducted on the full sample for each array to generate the ancestry components, before removing duplicates across arrays. The full sample available for each array was used to avoid very large discrepancies in sample size which could influence the results (McVean, 2009). PCA was carried out in Plink v1.90b4, on pruned markers, with a window size of 1000, a window shift of 5 and an r^2 threshold of 0.2. 10 PCs were generated.

Each of the GWAS was conducted in RVTESTS, a command-line GWAS software that’s main purpose is to carry out rare-variant analysis, but can also carry out analysis of common variants (Zhan et al., 2016). Reading at ages 7, 11, and 16, and an overall reading composite was analysed alongside genetic data to produce a score for meta-analysis in the form of a regression coefficient, beta. The imputation dosage values were used as input (Zheng et al., 2011), and the calculated output for all analyses was the meta-analysis score (Zhan et al., 2016). In cases of sample overlap between the arrays at each age point, the highest quality array was used for each duplicated individual, and duplicates were excluded (see Chapter 5 for full breakdown).
**GWAS Meta-Analysis.** Upon completion of the GWAS for individual arrays, the data were re-formatted in R v3.3.2 and meta-analysis was conducted at ages seven, 11, 16, and for the overall composite. Meta-analysis was conducted using METAL, a command-line software program designed for GWAS meta-analysis. METAL offers two core methods for carrying out GWAS meta-analysis; one based on z-scores, and the other based on standard errors (Begum et al., 2012). This study opted for the standard errors based approach given the same variable was being meta-analysed; additionally, it produces output most compatible with the genetic correlation analysis. After meta-analysis, R v3.3.2 and Unix were used format the output of the summary statistics in order to conduct genetic correlation analysis.

**Genetic Correlation Analysis.** Genetic correlations were carried out between the NCDS reading composites and (a) each other; (b) results of GWAS of a series of reading and language related traits from other studies; (c) a series of socioeconomic, occupational, and physical and mental health phenotypes. The Genoma Virtual Lab (Genoma VL) ([https://vl.genoma.io/](https://vl.genoma.io/)) was used to carry out this analysis (Cuellar-Partida et al., 2019, preprint). This is an online server which uses LDSC regression to assess genetic correlations between traits (Bulik & Sullivan, 2015a; Bulik & Sullivan, 2015b).

The Genoma VL stores publicly available GWAS summary statistics for over 1000 traits, primarily from UK BioBank summary statistics (Cuellar-Partida et al., 2019, preprint; Sudlow et al., 2015), and the functionality of the server allows one to conduct genetic correlation analysis with each trait through uploading their own data. Summary statistics from the NCDS GWAS were uploaded to the server, and LDSC was carried out between all of these traits in order to assess whether the reading measures used are genetically stable with age. Additionally, the GWAS summary statistics from the reading related measures described above were also
uploaded to the server. LDSC was conducted between each of the NCDS reading ability outcomes and each of these additional traits, in order to check the genetic validity of the NCDS reading measures and to assess whether results were able to replicate those found by Eising and colleagues (2022) and Doust and colleagues (2022).

The NCDS age seven reading composite and the NCDS overall composite were then analysed against each set of available summary statistics available on the Genoma VL server. The age seven composite was chosen as it has been recommended that including factors which may contribute before disease onset can reduce error caused by reverse causation (Bulik-Sullivan, Finucane, et al., 2015). The overall NCDS reading composite was chosen for comparison, as this covers a range of reading measures spanning from childhood to mid-life, and the author was interested to determine whether the addition of further information to the composite would result in a different outcome.

6.3. Results

**Reading Composites: Descriptive Information**

The generated reading composites for ages seven and 16 showed a highly skewed distribution (Figure 1). Means and standard deviations for these composites can be found in Table 2. High numbers of individuals scored highly on the Southgate Reading Test, which likely contributed significantly to this effect at age seven. A similar effect was likely in place at age 16, which may be due to inclusion of the binary ‘Can Cope’ variable, for which a large proportion could cope well enough to read. The reading measure for age eleven appears to be close to normally distributed (Figure 1). The overall reading composite is somewhat normally distributed,
however has a long tail to the left (Figure 1). This may be due to the relatively small number of individuals who said that they had difficulties reading at ages 23 and 33 (Table 3).

**Figure 1:** Kernel density curves showing distribution of four reading composite scores in the NCDS.
Table 2: Descriptive statistics relating to four reading composites in the NCDS.

<table>
<thead>
<tr>
<th>Reading Composite</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 7</td>
<td>0.00</td>
<td>0.80</td>
<td>5730</td>
</tr>
<tr>
<td>Age 11</td>
<td>0.00</td>
<td>0.70</td>
<td>5546</td>
</tr>
<tr>
<td>Age 16</td>
<td>0.01</td>
<td>0.70</td>
<td>4824</td>
</tr>
<tr>
<td>Overall</td>
<td>0.05</td>
<td>0.51</td>
<td>3101</td>
</tr>
</tbody>
</table>

Note: Means and standard deviations differ as standardization occurred before weighting by factor loadings, which were calculated for the full NCDS sample. See Chapter 5 for more details.

Table 3: Number of individuals who reported reading difficulties since leaving school in the NCDS.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Yes</th>
<th>No</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading difficulties: age 23</td>
<td>169</td>
<td>5408</td>
<td>857</td>
</tr>
<tr>
<td>Reading difficulties: age 33</td>
<td>204</td>
<td>5520</td>
<td>710</td>
</tr>
</tbody>
</table>

Genetic Analysis Results

Post Imputation Quality Control. Post-imputation quality control indicated that imputation quality was not of the highest standard, as a large number of variants had an INFO score of zero (see Chapter 5, Supplementary Information). The number of variants available for analysis on each array after quality control and removal of variants with a low MAF can be seen in Table 4.
Table 4: The number of arrays available for analysis on each array after removing low quality variants and low MAF variants.

<table>
<thead>
<tr>
<th>Array</th>
<th>Number of Variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affymetrix v6</td>
<td>117511993</td>
</tr>
<tr>
<td>Illumina 1.2million</td>
<td>87264358</td>
</tr>
<tr>
<td>Infinium v1.1</td>
<td>57392078</td>
</tr>
<tr>
<td>Infinium v3</td>
<td>107694225</td>
</tr>
<tr>
<td>Illumina 660-Quad</td>
<td>31955411</td>
</tr>
</tbody>
</table>

For each age point, GWAS was carried out on each array individually for use in meta-analysis. Because of this, the outcome of those GWAS is not the main focus of this study, and so will not be reported on in detail here. However, the quality of GWAS analysis should be noted. The number of individual samples analysed for each individual differed substantially between arrays, and missing data in the sample meant that only a very small number of individuals were available on certain arrays at certain age points (Table 5). Additionally, it should be noted that, at all age points, the Infinium v1.1 array analysis produced a number of errors due to calculation of invalid standard errors. This is likely due to the small number of individuals analysed on the array.
Table 5: The number of individuals analysed for each GWAS analysis and meta-analysis by array and age.

<table>
<thead>
<tr>
<th>Array Type</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 16</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affymetrix v6</td>
<td>161</td>
<td>155</td>
<td>148</td>
<td>101</td>
</tr>
<tr>
<td>Illumina 1.2m</td>
<td>2502</td>
<td>2403</td>
<td>2084</td>
<td>1345</td>
</tr>
<tr>
<td>Infinium v1.1</td>
<td>53</td>
<td>51</td>
<td>43</td>
<td>25</td>
</tr>
<tr>
<td>Infinium v3</td>
<td>2258</td>
<td>2203</td>
<td>1904</td>
<td>1230</td>
</tr>
<tr>
<td>Illumina 660-Quad</td>
<td>738</td>
<td>716</td>
<td>630</td>
<td>387</td>
</tr>
<tr>
<td>Meta-Analysed</td>
<td>5712</td>
<td>5528</td>
<td>4809</td>
<td>3088</td>
</tr>
</tbody>
</table>

At each age point, the GWAS meta-analysis was then conducted. Each completed meta-analysis output was uploaded to FUMA, an online server which can produce graphics of GWAS results, for visual inspection to ensure that no anomalies were present (Watanabe et al, 2017). The QQ plots for each age group largely indicate that P-values are as expected under the null hypothesis, indicating the results are not affected by population stratification (Ehret, 2010) (Figure 2). As expected due to the small sample size, no SNPs reached genome wide significance (Figure 3).

**Genetic Correlations between NCDS Composites.** The summary statistics for each of the reading composites in the NCDS were compared against each other in order to assess the genetic stability of the reading construct over time in the cohort. Results from age seven showed very high genetic correlations with results from ages 11 and 16, indicating that the composites are measuring a genetically stable reading construct (Table 6). Additionally, the composites from ages seven, 11, and 16 showed significant high positive genetic correlations with the overall reading composite, indicating that the additional adulthood measures used in this composite have a similar genetic basis. All correlations were statistically significant. This indicates that, while
the reading measures differ at each age of the NCDS, a reliable construct of reading can be taken at age seven, 11, or 16, along with an overall reading composite encompassing ages seven, 11, 16, 23, and 33. Heritability estimates were in line with previous results for GWAS of quantitative reading ability, calculated using the LDSC method (Eising et al., 2022).

**Figure 2**: A series of QQ plots produced in relation to GWAS for reading ability at four age points.
**Figure 3:** Manhattan plots showing the outcome of a series of GWAS on reading ability.
Table 6: Genetic correlations between NCDS reading composites at different ages.

<table>
<thead>
<tr>
<th>Variable A</th>
<th>Variable B</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 16</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Age 7</td>
<td>Age 11</td>
<td>Age 16</td>
<td>Overall</td>
</tr>
<tr>
<td>Age 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$r_G$</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$P$ Value</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$r_G$ SE</td>
<td>$&lt; 0.00$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$-h^2$ (Var. B)</td>
<td>0.32</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$h^2$ SE</td>
<td>0.11</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$r_{GE}$</td>
<td>0.84</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$P$ Value</td>
<td>$&lt; 0.00$</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$r_G$ SE</td>
<td>0.10</td>
<td>$&lt; 0.00$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$h^2$ (Var. B)</td>
<td>0.35</td>
<td>0.35</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$h^2$ SE</td>
<td>0.11</td>
<td>0.11</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$r_{GE}$</td>
<td>0.74</td>
<td>1.02</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$P$ Value</td>
<td>$&lt; 0.00$</td>
<td>$&lt; 0.00$</td>
<td>0.00</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$r_G$ SE</td>
<td>0.18</td>
<td>0.11</td>
<td>$&lt; 0.00$</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$h^2$ (Var. B)</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$h^2$ SE</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$r_{GE}$</td>
<td>0.70</td>
<td>0.84</td>
<td>0.80</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$P$ Value</td>
<td>$&lt; 0.00$</td>
<td>$&lt; 0.00$</td>
<td>$&lt; 0.00$</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>$r_G$ SE</td>
<td>0.14</td>
<td>0.11</td>
<td>0.13</td>
<td>$&lt; 0.00$</td>
</tr>
<tr>
<td></td>
<td>$h^2$ (Var. B)</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>$h^2$ SE</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
</tbody>
</table>
Genetic Correlations with Additional Reading and Language Measures. Table 7 indicates a statistically significant genetic correlation between the reading composites at ages seven, 11, 16, and overall with each of the additional reading related measures that were included in the analysis. All correlations were moderate to strong, however the genetic correlation between the NCDS reading measures and self-reported dyslexia was consistently the weakest, with rG ranging from -0.41 to – 0.58 (p ≤ 0.00 in all instances).

Genetic Correlations with Other Variables. Genetic correlations were produced using all available traits on the Genoma VL server, resulting in 1436 correlations for inspection at age seven and another set for the overall composite. The purpose of this analysis was an attempt to replicate and expand on previous findings of significant genetic correlations between continuous reading ability and and dyslexia, and socioeconomic, educational, occupational, and physical and mental health and wellbeing (Eising et al, 2022; Doust et al., 2022). This section begins with an overview of estimated genetic correlations, before examining interesting findings in more detail.
Table 7: Genetic correlations between reading composites in the NCDS and reading and language measures from other cohort studies.

<table>
<thead>
<tr>
<th>Variable A</th>
<th>Variable B</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 16</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoneme Awareness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rG</td>
<td>0.87</td>
<td>0.69</td>
<td>1.08</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>P Value</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.21</td>
<td>0.19</td>
<td>0.31</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>h² (Var. B)</td>
<td>0.21</td>
<td>0.21</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>h² SE</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Non-Word Repetition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rG</td>
<td>0.63</td>
<td>0.67</td>
<td>0.76</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>P Value</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.22</td>
<td>0.25</td>
<td>0.33</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>h² (Var. B)</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>h² SE</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Spelling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rG</td>
<td>0.49</td>
<td>0.62</td>
<td>0.93</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>P Value</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.13</td>
<td>0.15</td>
<td>0.26</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>h² (Var. B)</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>h² SE</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Non-Word Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rG</td>
<td>0.69</td>
<td>0.75</td>
<td>0.98</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>P Value</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.17</td>
<td>0.17</td>
<td>0.28</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>h² (Var. B)</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>h² SE</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Word Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rG</td>
<td>0.84</td>
<td>0.85</td>
<td>1.07</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>P Value</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.19</td>
<td>0.16</td>
<td>0.27</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>h² (Var. B)</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>h² SE</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Dyslexia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rG</td>
<td>-0.58</td>
<td>-0.42</td>
<td>-0.53</td>
<td>-0.41</td>
</tr>
<tr>
<td></td>
<td>P Value</td>
<td>&lt; 0.00</td>
<td>&lt; 0.00</td>
<td>0.00</td>
<td>&lt; 0.00</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.13</td>
<td>0.10</td>
<td>0.17</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>h² (Var. B)</td>
<td>0.03^a</td>
<td>0.03^a</td>
<td>0.03^a</td>
<td>0.03^a</td>
</tr>
<tr>
<td></td>
<td>h² SE</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note^a: The genoma server was not able to properly estimate the heritability of this trait due to it’s binary nature. This estimate is much lower than the estimate provided by Doust et al (2022) using LDSC in a different programme.
**Genetic Correlations between Reading Ability and Other Traits.** Genetic correlation analysis was split into seven categories for ease of interpretation; cognitive, education, mental health and wellbeing, occupation, health, pain, and SES. Results are reported with respect to both the usual alpha value of 0.05, and additionally, with a Bonferroni-corrected value for 165 tests at 0.0003. Note that this correction is likely to be overly conservative due to the substantial overlap that is expected between the traits, and so values approaching significance at the Bonferroni level will be discussed. Significant results at the 5% level (2 significant figures) can be found in Figures 4 and 5, while the full set of selected variables can be found in the appendix. Analysis was continued using these two measures as they each have different properties of interest; the age seven result has the largest sample size, and the overall result, while the same size is small, showed high genetic correlations with all other reading measures and shows the highest heritability of all traits, indicating that less SES is being captured in this measure than in others. Results indicated that, at age seven, reading ability is negatively correlated with measures of occupational conditions, with negative correlations being identified with workplaces that are often noisy (rG = -0.49, P = 0.00053), dusty (rG = -0.52, P = 0.00072), cold (rG = -0.42, P = 0.01244), or contain potentially hazardous chemicals (Table S1). While several of these tests approached significance at the Bonferroni level, none reached the threshold. Conversely, a positive correlation was identified between reading ability and being in employment or doing voluntary work, although again, the Bonferroni-corrected threshold was not reached. Occupational variables which did not show a genetic correlation included measures which could be considered as age-related, such as being retired (rG = 0.15, P = 0.43) or looking after the home or family (0.14, P = 0.49) (Table S1). The majority of occupational variables showed no
genetic correlation with the overall composite, with only three reaching the 0.05 threshold (noisy workplace, asbestos, and shift work).

Mental health and wellbeing measures showed inconsistent genetic correlations with reading ability. For age seven reading, four measures of mental wellbeing reach significance at the 5% level, and in all cases greater reading skill was associated with greater wellbeing. None of these measures were clinical diagnoses, instead focusing on the general feelings of the participant, such as mood swings (rG = -0.30, P = 0.00020), which was the only measure in this category to surpass the Bonferroni threshold of 0.00030. Not all wellbeing measures were significantly genetically correlated with reading ability, with several others, including risk taking (rG = -0.02, P = 0.78700) and history of self-harm (rG = 0.18, P = 0.28590) showing no significant association (Table S2). Several psychiatric conditions were also tested, but none showed a significant association with reading ability, including Major Depressive Disorder (MDD) (rG = -0.10, P = 0.23470), schizophrenia (rG = -0.01, P = 0.85010), obsessive compulsive disorder (OCD) (rG = -0.12, P = 0.66540), and anxiety disorders (rG = -0.04, P = 0.76630) (Table S2). The neurodevelopmental trait of ADHD showed a negative genetic correlation with reading ability (rG = -0.50, P = 0.00), although this did not survive Bonferroni correction. These results did not replicate for the overall reading composite, with the only two variables approaching or surpassing the 5% threshold, and none surviving Bonferroni correction.
Figure 4: Significant genetic correlations (P ≤ 0.05) between reading ability composite at age seven and a series of cognitive, educational, mental health and wellbeing, and occupational measures.
Figure 5: Significant genetic correlations (P ≤ 0.05) between reading ability composite at age seven and a series of health, pain, and SES measures.

Genetic correlations were as expected for both educational attainment and cognitive variables (Figure 4, full details in supplement). At age seven, reading ability showed positive significant genetic correlations with all cognitive measures included, including verbal-numerical reasoning (named fluid intelligence in UKB) (rG = 0.61, P = 0.00000) and cognitive
performance ($r_G = 0.69, P = 0.00000$). Almost all of these associations survived Bonferroni correction, and results were very similar for the overall reading composite, with only two results failing to replicate at the Bonferroni level. Cognitive performance here was drawn from a large meta-analysis of several cohorts, with a variety of cognitive tests used in childhood and adulthood (Rietveld et al., 2014). The majority of individual segments of the ‘fluid intelligence’ test also showed a strong genetic correlation, identifiable in tables and figures by beginning with the ‘FI’ prefix. At the Bonferroni level, reading ability showed positive genetic correlations with obtaining higher qualifications and overall educational attainment, and negative genetic correlation with achieving lower level or no qualifications at both age seven and overall, supporting earlier work in this dataset (Davies et al., 2015). Similarly, genetic correlations between reading ability and measures of SES were as expected. At age seven, several tested variables were significant at the Bonferroni level, with positive correlations reported for variables associated with high SES, such as outright home ownership ($r_G = 0.39, P = 0.00028$), and negative correlations reported for variables associated with lower SES, such as a deprivation index ($r_G = -0.35, P = 0.00020$) and being unable to work due to ill health ($r_G = -0.59, P = 0.00001$) (Table S3). Only one of these associations was replicated at the Bonferroni level for the overall reading composite, however most were replicated at the 5% level.

Multiple measures for pain and general health were significantly genetically correlated with reading ability. A range of health related variables indicated that, in general, greater reading ability at age seven is genetically correlated with poorer health outcomes, including dental issues ($r_G = -0.37, P = 0.00002$), and hypertension ($r_G = -0.18, P = 0.01000$), although only hypertension, dental problems, and exercise in the last four weeks survived the Bonferroni correction. Several other variables demonstrated low P values, nonetheless. Hayfever showed a
positive genetic correlation with reading ability ($r_G = 0.30$, $P = 0.00620$), and BMI showed a negative genetic correlation with reading ability ($r_G = -0.17$, $P = 0.00863$) (Table S4). Dental problems was the only general health measure to replicate for overall reading ability at the Bonferroni level.

Several pain variables were shown to have significant negative genetic correlations with reading ability, replicating results found by Doust and colleagues and published in 2022. All pain variables that showed a significant genetic correlation were negative, indicating that genetics for lower reading ability are associated with likelihood of experiencing pain. Given the strong replication of a previous finding which was extensive across pain type, the next section explores this genetic association in more detail, by interrogating further pain related variables for genetic correlations with reading.

**Genetic Correlations between Reading Ability and Specific Pain Complaints.** Table S5 details the genetic correlations calculated for the NCDS age seven and overall reading composites with variables relating to specific pain complaints, i.e., complaint of pain at a particular location in the body without reference to a condition or disorder. The age seven reading composite significantly correlates with three of the measured pain types at the Bonferroni level, while the overall composite did not show any significant correlations at this level, although many more variables were implicated at the 5% level for the age seven composite and three reached significance at the 5% level for the overall composite. Significant associations across all pain types were negative in direction, further indicating an inverse relationship between reading ability and pain, in other words, greater propensity for high reading ability is associated with lower propensity for experiencing the measured pain types. Multiple significant
correlations were identified in all categories: back, shoulder, and neck pain; knee, hip, and leg pain; stomach and abdominal pain; and pain above the neck.

**Genetic Correlations Between Reading Ability and Non-Specific Pain.** This section reports the genetic correlations between continuous reading ability and non-specific pain, defined here as pain experienced in an unidentified location in the body without explicit connection to a specific physiological disorder (Table 8). These results are particularly striking, as all correlations between both composites and all measured pain types are statistically significant at the 5% level for both ages, with two significant at the Bonferroni level at age seven and multiple chronic pain approaching significance at the Bonferroni level for the overall composite. All correlations measuring the presence of pain are negative, again indicating that increased propensity for good reading ability is associated with decreased propensity for experiencing pain. Conversely, the variable *Pain type(s) experienced in last month: None of the above* measures the absence of pain, and is positively correlated with reading ability, indicating propensity to read well is associated with reduced propensity to experience non-specific pain.
Table 8: Genetic correlations between reading ability composites in the NCDS and non-specific pain complaints

<table>
<thead>
<tr>
<th>Non-Specific Pain</th>
<th>Age 7 Composite</th>
<th>Overall Composite</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (pain GWAS)</td>
<td>rG</td>
</tr>
<tr>
<td>Tense, sore, or aching muscles during worst period of anxiety</td>
<td>33301</td>
<td>-0.65</td>
</tr>
<tr>
<td>Pain type(s) experienced in last month: Pain all over the body</td>
<td>360391</td>
<td>-0.43</td>
</tr>
<tr>
<td>Pain type(s) experienced in last month: None of the above</td>
<td>60391</td>
<td>0.37</td>
</tr>
<tr>
<td>Multiple chronic pain</td>
<td>380000</td>
<td>-0.41</td>
</tr>
</tbody>
</table>

Genetic Correlations between Reading Ability and Medication Use. Table S6 shows the results of genetic correlation analysis between reading ability and taking various medications. Medications for investigation were selected due to their status as painkillers, their use as treatment for pain-related disorders, or their use as supplements for common joint complaints. Investigating the results by medication type reveals some interesting insights. There are three non-steroidal anti-inflammatories (NSAIDS) that show a negative genetic correlation with the age seven reading composite, ranging from -0.22 to -0.56 (P ≤ 0.05 in all instances), and only one opioid of four shows this same effect (rG = -0.68, P = 0.00690). For the overall composite, only one correlation was identified among the NSAIDS and one among the opioids (P ≤ 0.05). The common medication paracetamol shows a weak to moderate significant genetic correlation with reading at age seven (rG = -0.33, P = 0.00150), and the anti-convulsive gabapentin, which is also used to treat neuropathic pain (Kukkar et al., 2013), shows a similar result (rG = -0.62, P = 0.01816). Glucosamine and chondroitin, both supplements used with intent to protect the joints
(McAlindon et al., 2000), show moderate to strong positive genetic correlations with reading at age seven (P ≤ 0.05). None of the associations were significant at the Bonferroni level.

**Genetic Correlations Between Reading Ability and Pain Related Conditions.** Table S7 shows the genetic correlations between reading at age seven and overall reading ability and various health conditions for which a key symptom is pain. The majority of associations in this section are not statistically significant. The section regarding spinal disorders shows the greatest number of significant correlations; three for reading at age seven, and two for overall reading ability (P ≤ 0.05). As with the majority of results reported in this piece of work, each significant correlation in this section is negative in direction, indicating that increased propensity to good reading ability goes with decreased propensity for experiencing a pain-related health condition. A negative correlation was found between reading ability and nerve disorders at age seven (rG = -0.33, P = 0.00738) and overall (rG = -0.34, P = 0.01115).

**Genetic Correlations between Reading Ability and Angina.** Almost all genetic correlations between chest pain or angina and reading ability at age seven were statistically significant in a negative direction, although none reached the Bonferroni threshold (Table 9). Most of these associations were replicated for the overall composite, and this set of results shows the most consistent level of agreement between the age seven results and the overall results.
Table 9: Genetic correlations between reading composites in the NCDS and chest pain and angina.

<table>
<thead>
<tr>
<th>Chest pain and angina</th>
<th>Age 7 Composite</th>
<th>Overall Composite</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>rG</td>
</tr>
<tr>
<td>Chest pain or discomfort</td>
<td>357507</td>
<td>-0.22</td>
</tr>
<tr>
<td>Chest pain or discomfort walking normally</td>
<td>55737</td>
<td>-0.37</td>
</tr>
<tr>
<td>Chest pain or discomfort when walking uphill or hurrying</td>
<td>44867</td>
<td>0.09</td>
</tr>
<tr>
<td>Vascular/heart problems diagnosed by doctor: Angina</td>
<td>360420</td>
<td>-0.26</td>
</tr>
<tr>
<td>Unstable angina pectoris</td>
<td>361194</td>
<td>-0.29</td>
</tr>
</tbody>
</table>

Genetic Correlations between Reading Ability and Joint Disorders. Table S8 shows the genetic correlations between reading ability and joint disorders. Only one variable, Other specific joint derangements/joint disorders, showed a statistically significant negative correlation with either reading composite (rG = -0.56, P = 0.030120). This indicates a lack of evidence to conclude that reading ability and joint disorders share a common genetic basis, which will be explored further in the discussion.
6.4 Discussion

In this project, GWAS was carried out on four composites of reading measures from a large UK Cohort study. Meta-analysis was carried out on data collected across different arrays, in order to ensure accuracy. The resulting summary statistics were compared with each other using LDSC to generate genetic correlations, which indicated high genetic similarly of reading at the three separate age points and in the overall composite. Further genetic correlation analysis identified moderate to strong correlations between the quantitative reading composites and reading and language traits measured in other better-powered samples, including phoneme awareness, spelling, and self-reported diagnosed dyslexia. Additionally, significant genetic correlations were identified for measures of SES, educational attainment, cognitive ability, physical and mental health outcomes, and pain. Significant negative genetic correlations between reading and specific pain complaints were extensive and interrogated further. This study did not find convincing evidence of genetic pleiotropy between reading ability and most pain-related disorders, however some suggestive associations were identified with painkiller use: those with high reading ability in childhood have greater genetic propensity for taking NSAIDS.

Genetic Correlations Across age

Reading measured across ages showed strong positive genetic correlations. This indicates that the genetics of reading ability are stable across development in this cohort, even though the reading measures used differed between ages. These correlations support a body of existing work indicating that reading is genetically stable over time (e.g., Eising et al., 2022; Betjemann et al., 2007; Wadsworth et al., 2001). However, it is interesting to note that the age seven reading measure correlates more weakly with each subsequent measure. This may be due to the reading measures used in the generation of the composite, as age seven measures all strongly focus on
the reading ability of the child, including psychometric measures (Southgate reading test), as well as a teacher rating of the reading grade the child had attained (Bridges et al., 2022). As age increases, the measures included in the composite become more subjective, including teacher rating of ability at 11 and 16, and self-report ratings of ability at ages 23 and 33. This may indicate that the composite is capturing some error not caused by reading ability at those ages, or alternatively, it may be that additional genetic factors come into play with age. Existing research to date has indicated that additional factors may become active between first and seventh grade (Erbeli et al., 2017). It is also possible that the weakening correlations may be due to the variables included to form the composites. The age seven reading measure has a stronger focus on word reading, due to the inclusion of the Southgate test, while ages 11 and 16 included comprehension tests; this is supported by the fact that decoding and comprehension share some genetic basis, however there are also independent effects which act on comprehension (Keenan et al., 2006). Finally, another explanation may be due to developmental differences early on in schooling; some children may struggle with reading early on in school, and these difficulties may resolve as they progress through their education (Torppa et al., 2015; Aro & Wimmer, 2003). Nevertheless, the genetic correlations between all reading composites are strong, with all being greater than 0.70. This is significant in the context of the broader field for two reasons: (1) this is the first study that the author is aware of to use molecular methods to estimate the genetic similarity of reading across development over such a significant length of time, and the findings support previous existing research, and (2) this provides strong evidence which validates the use of the NCDS reading composites for future reading research. The combined strength of the phenotypic (Bridges et al., 2022) and genetic correlations across ages indicates that, while measurement of the variables differs, all composites are, in fact, measuring reading ability.
The results reported here also indicate that the reading measures are genetically correlated with other reading measures from independent samples, including dyslexia and quantitative decoding-related skills. This strengthens the argument that the NCDS reading composites are truly measuring reading skill, although it should be noted that all NCDS composite measures are likely to be encompassing some level of comprehension in addition to word-reading skill. Genetic correlations between other cognitive traits and both reading at age seven and overall reading are similar, with strong positive correlations found for cognitive performance ($r_G = 0.69$), verbal numerical reasoning ($r_G = 0.61$), and educational attainment ($r_G = 0.67$). This is expected, given past research which has indicated that reading ability and cognitive ability share some genetic bases (eg. Harlaar et al., 2005; Eising et al., 2022).

The majority of previous reading GWAS have focused on either case control designs or have used standardised testing to determine values for the reading phenotype. It has been previously argued that this may aid in replication of results (Bates et al., 2004), however, it has more recently been suggested that reliance on standardised reading tests may be detrimental to the field of behaviour genetics reading research, as it limits the sample sizes that may be used in genetic analysis (Eising et al., 2022). This study has provided evidence to show that teacher report and self-reported measures can be valuable measures to use for future research on the genetics of reading, particularly if used as part of a composite measure as is the case in this study. It is possible that this may extend to other cohorts outside of the NCDS, and this should be an area of future research to ensure that GWAS sample sizes for reading are as large as possible, to give them enough power to detect causal variants (Visscher et al., 2017).

*Genetic Correlations with Other Reading Measures*
The genetic correlations reported with other measures in this study are predominately consistent with the small amount of molecular genetic literature already in existence, and also concur with the phenotypic literature. It has been repeatedly shown that reading ability and SES are phenotypically associated (e.g., Herbers et al., 2012; Larson et al., 2015; Peterson & Pennington, 2015), including in the NCDS (Russell et al., 2018). Existing research has indicated genetic pleiotropy between reading ability and SES (Eising et al., 2022; Doust et al., 2022). Similarly, phenotypic evidence shows strong associations between reading ability and educational attainment (e.g. McGee et al., 2002; Levlin et al., 2022), including in the NCDS (Davies et al., 2015), and there is early molecular genetic literature indicating pleiotropy (Belsky et al., 2016). It has been suggested that pleiotropy between reading ability and cognitive traits, such as intelligence, could be due to generalist genes (Trzaskowski et al., 2013). The Generalist Genes hypothesis proposes that a collection of common SNPs show significant combined influence on many human traits (Kovas & Plomin, 2006). Evidence has shown that a PRS for educational attainment can predict adult SES outcomes, and the same study found that 46% of that association was mediated by a combination of cognitive abilities and non-cognitive skills (Belsky et al., 2016). While it may be intuitive to assume that the relationship between SES and reading skills acts through general cognitive ability, this suggests that cognitive ability may not be the only variable through which this association acts, and other factors such as personality and motivation may also be relevant in how the quad of reading, cognition, SES, and educational attainment interact (Belsky et al., 2016). Further research could explore the mechanism for this association further.

There is less research regarding an association between occupational conditions and reading ability, however phenotypic research again shows a strong association: those with
reading difficulties are more likely to work in low status roles, including manual labour (Smart et al., 2017). A small amount of existing molecular genetic research in this area has suggested that manual work is genetically associated with both quantitative reading ability (Eising et al., 2022) and dyslexia (Doust et al., 2022). This result is not surprising, given that occupation is often included in measures of SES, and reading ability and SES show a moderate genetic correlation, however the relationships and potential mediating factors are currently poorly understood, and this is an area where further multivariate genetic research could help to clarify these associations.

This paper also found indications that reading ability is genetically associated with poor health outcomes. It is possible that this association is due to common biological mechanisms, however, the range of conditions which were found to be significant must be considered. Heart attack, diabetes, physical activity levels, and BMI were all found to be associated with reading ability, however, all of these measures are also associated with lower SES (O’Rand & Hamil-Luker, 2005; Everson et al., 2002; Giles-Corti & Donovan, 2002; Wang et al., 2007). Because of this, it should be considered that the genetic correlations presented here are indicative of vertical pleiotropy, in which genetic variants are predictive of reading ability, and also predictive of SES, which in turn has a causal relationship with these health outcomes. This is particularly relevant given that reading ability in childhood has shown a strong association with adulthood SES in the NCDS itself (Russell et al., 2018). A positive genetic correlation was found between reading ability and birth weight, and higher birth weight is associated with higher parental SES (Thomson et al., 2021).

The mental health and wellbeing outcomes assessed in this study corroborate previous evidence regarding genetic associations with reading ability, in the cases of those variables for which a genetic correlation was found, and also in those cases where one was not. For example,
previous research has not found a consistent association between reading ability and depression, or schizophrenia, both of which failed to reach significance in this study. It is interesting, therefore, that associations were found between reading ability and sub-clinical reports of low mood, which replicated previous results (Eising et al., 2022). Those with poorer reading ability are more likely to experience adverse life circumstances, such as unemployment (Aro et al., 2018), therefore it is possible that they are more likely to experience temporary feelings of low mood or mood swings.

Perhaps the most interesting findings in this study are the genetic associations reported between reading ability and pain. These findings replicate findings reported by Doust and colleagues (2022), but very little other research can be found which documents either a genetic or a phenotypic relationship between the two. The results reported here indicate that associations between reading ability and pain are predominately for multisite pain, specific pain locations, and angina, and do not hold for health conditions for which pain is a key characteristic, other than angina. It should not be ignored that pain also shows existing phenotypic associations with SES (Booher, 2019). In some cases, it may be that the association reported here is yet another example of vertical pleiotropy. Some work on this topic is already underway for the association between pain an ADHD, for example, an association between ADHD is hypothesised to be caused by neuroinflammation (Kerekes et al., 2021), and the common ADHD medication Methylphenidate may influence how people with ADHD experience pain (Bozkurt & Balta, 2023).

One interesting, and unexpected, finding from this research was the suggestive associations reported between the genetics of reading ability and painkiller use. Several of these associations hold only with the age seven reading measure, and not the overall composite. The
higher heritability of the overall composite suggests that environmental factors, such as socioeconomic status, hold less influence on reading ability in adulthood, which would be in line with findings from other cognitive traits such as IQ (Bouchard, 2013). It has also been shown that heritability of reading ability increases with age using twin methodology (Wadsworth et al., 2001), and for reading comprehension, heritability increases and shared environment decreases (Soden et al., 2015). This is supported by the fact that most of the genetic correlations between reading ability and SES measures at age seven survived Bonferroni correction, while the majority of those for the overall composite did not, although almost all were still significant at the 5% level (Table S3). It may be that for those associations that were identified for the age seven composite but not the overall composite, SES is acting as a confounding factor. It has been shown that those of low SES are more likely to be prescribed prescription painkillers in a Scottish sample (Wakefield et al., 2015), and more likely to take them in a German sample (Sarganas et al., 2015). However, the converse has been true for over-the-counter pain medications, which are more likely to be taken by those of higher SES (Sarganas et al., 2015). It is also possible that the differential results could be caused by differential use of painkillers due to other factors, such as ease of obtaining particular medicines, with many of the medications tested in this study requiring a consultation to obtain, in the UK. Therefore, it may be that those who are taking prescription pain medications are taking them because their pain levels are exceptionally bad. In addition to these possible confounding factors, it should be noted that heritability of painkiller use was low in many instances, further indicating that more research is needed in order to properly understand the relationship between reading, pain, painkiller use, and their genetic and environmental associations.

Limitations
While this study has generated some interesting results, it is important that we consider potential limitations and how they may have influenced the findings. A key limitation of this study is the available sample size, particularly for the overall reading composite measure. The LDSC documentation (Bulik, n.d.) indicates that a minimum sample size of 5000 is required in order to ensure accurate results for generation of genetic correlations. For the overall reading composite, this number was not achieved. This was due to high levels of missing data across the cohort study caused by many participants failing to answer all questions at each timepoint. The sample size could have been increased by opting to include all individuals who had answered any of the reading questions waves of data collection. Because of the low sample size, the genetic correlation between the overall composite results and the other variables should be interpreted cautiously, as results may not be accurate.

Future research could repeat the overall reading skill GWAS using all individuals who had answered any question, if further genetic correlations were to be explored, if to do so would be appropriate for the research question. Alternatively, future researchers may use an alternative method for estimating genetic correlations, such as GREML, which is more robust to small sample sizes (Ni et al., 2018). GREML was not used for this analysis, as it requires full genotype data for all individuals analysed, whereas LD Score Regression only requires summary statistics in order to calculate genetic correlations (Ni et al., 2018).

A further limitation is the fact that LD documentation (r https://github.com/bulik/ldsc) indicates that, ideally, variants should be filtered before inclusion in the genetic correlation analysis before use. The authors recommend only using variants which have an INFO score > 0.9, to ensure quality of results. This was not possible in this analysis, as the GWAS results were meta-analysed before LD score regression. However, prior to GWAS analysis, the data were
filtered so that only variants with an INFO score > 0.7 were included (Bridges et al., 2023),
which may go some way to ensuring high imputation quality in the meta-analysis.

It must be emphasized the composites used in this analysis are, above all, measures of
functional reading ability; in other words, measures of whether the reading skills of the
individual are adequate for daily life. As a result, the composite scores do not offer any
differentiation between different elements of reading skills, such as decoding or comprehension.
As mentioned above, the genetic basis for these underlying skills differs somewhat (Keenan et
al., 2006), studies using these composites are not able to confer information about which
particular elements of reading ability, if any, share genetic effects with other traits. Similarly, due
to the fact that this study utilized pair-wise genetic correlations, it was not possible to distinguish
how much of the genetic relationships identified may be underpinned by a shared genetic basis
with cognitive ability. Use of PGS methodology would be valuable for exploring this further, as
this method provides the capability to control for the genetic propensity to cognitive ability while
exploring the relationship between the genetics of reading and other outcomes (Luciano et al.,
2017). In addition, the genomic structural equation modelling method may be used to explore
shared genetic covariance between reading, cognitive ability, and other outcomes (Grotzinger et
al., 2019).

To conclude, this study has reported a series of genetic correlations between reading
ability and cognition, educational attainment, SES, physical health, mental health, and pain.
Phenotypic literature indicates that all of these variables are interconnected with each other,
however, the genetic underpinnings of many of these associations are unclear. This study has
gone some way to opening up new avenues of exploration for genetic associations between
reading ability and health outcomes, particularly pain, by replicating some early findings in this
area and carrying out a more thorough investigation of the genetic association between reading ability and pain. At this stage, however, the true meaning of the genetic correlations, and any causality involved, is unclear. Further research must make use of multivariate genetic methods, such as multi-polygenic score analysis and genomic structural equation modelling, to assist understanding of the genetic relationships between reading ability and physical and mental health outcomes, and how SES may play a role in each of these associations. This may have significant clinical applications, whereby effective prevention programmes may be implemented early, such as more regular screenings for illnesses in those with reading difficulties, so that overall health may be maintained for longer. This is particularly relevant in light of predictions that chronic pain prevalence will increase against the backdrop of an ageing population (Fayaz et al., 2016).
6.5. References


https://github.com/bulik/ldsc


polygenicity in genome-wide association studies. Nature Genetics, 47(3), 291–295. https://doi.org/10.1038/ng.3211


42 genome-wide significant loci associated with dyslexia. *Nature Genetics, 54*(11), 1621–1629. https://doi.org/10.1038/s41588-022-01192-y


Chapter 7: Childhood Reading Ability and Pain in Childhood through to Midlife

The preceding chapter found evidence of a genetic correlation between quantitative reading ability and pain, a finding which replicates previous findings which indicated a genetic association between quantitative reading and pain (Eising et al., 2022) and dyslexia and pain (Doust et al., 2022). As indicated in the submitted work, it is possible that this genetic association is confounded by SES, as SES is associated with both reading ability (Russell et al., 2016) and pain (Poleshuck & Green, 2008). The submitted work presented in this chapter represents an attempt to elucidate whether there is a phenotypic relationship between childhood reading ability and pain in childhood and adulthood in the NCDS. In addition, this work presents a causal mediation analysis which attempts to determine whether reading and pain associations are mediated by SES, which would indicate some evidence of vertical pleiotropy, which occurs when the same genetic variant influences more than one trait through downstream causal pathways, rather than through the same variant directly affecting different biological mechanisms (Verbanck et al., 2018).

The following piece of work has been submitted to the Journal of Pain (US Version) as found below (other than minor clarifications and typographical corrections). Revisions were made in response to reviewer and editor feedback, and the version that follows below is the revised version, on which a decision is yet to be made.
7.1. Submitted Work
Childhood Reading Ability and Pain in Childhood through to Midlife

Elinor C. Bridges¹, Carole Torsney²³, Timothy C. Bates¹, Michelle Luciano¹

¹School of Philosophy, Psychology and Language Sciences, University of Edinburgh, 7 George Square, Edinburgh, EH8 9AD, United Kingdom

²Centre for Discovery Brain Sciences, University of Edinburgh, Edinburgh, UK

³Simons Initiative for the Developing Brain, University of Edinburgh, Edinburgh, UK

Corresponding author: Dr Michelle Luciano, School of Philosophy, Psychology and Language Science, University of Edinburgh, 7 George Square, Edinburgh, EH8 9AD

Email:

Disclosures: This work was funded by a PhD studentship from the Economic and Social Research Council with the Scottish Graduate School of Social Sciences. We have no conflict of interest to disclose.

Running Title: Association Between Reading Ability and Pain
Abstract

Dyslexia and pain have recently been shown to correlate on a genetic level, but there has been little exploration of this association on the phenotypic level despite reports of increased pain in Attention-Deficit Hyperactivity Disorder (ADHD) which commonly co-occurs with dyslexia. In this study we test for an association between reading ability, which is the primary feature of dyslexia, and pain both in childhood and adulthood. Logistic regression modelling was used to test associations between reading ability in childhood and pain from childhood to midlife in a large UK birth cohort; the 1958 National Child Development Study. Associations were found between poor childhood reading ability and increased headache and abdominal pain in childhood, and between poor childhood reading ability and headache, eye pain, back pain, and rheumatism in adulthood. Mediation analyses indicated that socio-economic status (SES: defined by employment), fully mediated the association between poor reading ability in childhood and back pain at age 42. By contrast, the association between reading ability and eye pain acted independently of SES. Different mechanisms were thus indicated for association of reading with different pain types, including manual labour and a potential shared biological pathway.

Perspective: This study found a relationship between poor reading ability in childhood and pain in adulthood. Those with reading difficulties should be monitored for pain symptoms. Future research may uncover shared biological mechanisms, increasing our understanding of pain and potential treatments.

Keywords: reading, headache, cohort study, mediation, socio-economic status
Introduction

Low reading skill remains a prevalent global issue, with 21% of United States population showing poor English literacy skills (Mamedova & Pawlowski, 2019). Pain is also a common problem worldwide. A meta-analysis of UK studies has estimated that chronic pain prevalence in UK adults, defined as, “…pain in one or more body locations, lasting for a period of 3 months or longer,” is between 35.0% - 51.3% (Fayaz et al., 2016, pg.2). Chronic pain can be treatment-resistant (Borsook et al. 2018), and those in the most socioeconomically deprived areas more likely to suffer chronic pain and to be prescribed opioids (Smith et al., 2019; Torrance et al., 2018).

Reading difficulties may be defined either through a diagnosis of dyslexia, or as the bottom portion of a normal distribution of reading skill (Snowling & Hulme, 2020). Genome-wide association research has indicated a genetic association between reading ability and pain through significant genetic correlations (Doust et al., 2022; Eising et al., 2022). Genetic correlations between traits can indicate that some of the same genetic variants act in the same direction to contribute to the expression of both traits (Bulik-Sullivan et al., 2015), which can be indicative of pleiotropy and therefore a shared biological mechanism (Chebib et al., 2021). Alternatively, genetic correlations can reflect mediated pleiotropy, in which one trait causes another but the genetic variants which act directly on the causal trait are detected in the secondary trait (Solovieff et al., 2013). The relationship between the causal and secondary traits may or may not be genetic in nature (Solovieff et al., 2013). Doust and colleagues (2022) reported significant genetic correlations between self-reported dyslexia and several pain types. Eising and colleagues (2022) found significant genetic correlations between quantitatively measured reading traits and various pain measures in the direction of poorer reading skills and
greater reporting of pain. In addition, reading difficulties frequently co-occur with Attention Deficit Hyperactivity Disorder (ADHD) symptoms (McGee et al., 2002; Gooch et al., 2013; Wilcutt et al., 2000), which may share a common genetic basis (Daucourt et al., 2019; Doust et al., 2022). ADHD has been repeatedly associated with pain and the pain syndrome fibromyalgia (Stickley et al., 2016; Stray et al., 2013; Pan et al., 2021; Revero et al., 2011; van Rensburg et al., 2017; Yilnaz et al., 2018).

This paper investigates whether the odds of experiencing pain is higher for those who show lower reading ability in childhood. A small amount of existing research indicates no association between poor childhood reading skills and headache or migraine (Andrasik et al., 1988; Waldie et al., 2002), and we expect to replicate these early findings for other pain types in our study of the National Child Development Study (Centre for Longitudinal Studies, n.d.). However, there are no prior studies on childhood reading skill and pain in adulthood, and we aim to address this gap in knowledge in our study. In light of the reported genetic associations, we hypothesize an association between childhood reading score and adulthood pain. We expect that this association is not predominately genetic in nature, and is instead mediated by adulthood socioeconomic status (SES), specifically by occupation; this is consistent with the observed moderate genetic correlation between dyslexia and manual work (Doust et al., 2022). A mechanism for this association has not, to our knowledge, been identified. Building on this, we suggest reading skill will influence adulthood SES (Smart et al., 2017), as poor childhood readers may be more likely to work in jobs without high literacy demands, notably manual labour (Baker et al., 1997), causing pain (Prego-Dominquez et al., 2021). Other concepts associated with low adulthood SES, such as poor housing quality or differential healthcare quality, may also be causal for pain (Urwin et al., 1998; Cookson et al., 2016). A lack of
mediation via SES would strengthen the alternative explanation of shared biological mechanisms causing the association between reading ability and pain.

Understanding the association between reading and pain, and the mechanism(s) responsible, could lead to better outcomes for patients with pain through early identification and intervention where possible. If poor reading is associated with pain, children and adults with reading difficulty can be monitored for pain and intervention, for example, preventative physiotherapy could be delivered early to reduce later backpain. Early intervention has been shown to reduce opioid prescription in the case of low back pain (Martin et al., 2020), and so early intervention among poor readers may reduce opioid use in this group, who are more likely to be of low SES and therefore at higher risk of opioid prescription. This may reduce the associated risk of addiction and other negative health outcomes in a potentially vulnerable population (Chou et al., 2015). In addition, an association between poor reading ability in childhood and pain in adulthood provides strong further motivation to identify and support children with reading difficulties, as failure to do so may lead to poor health outcomes much later in life.

Materials and Methods

Study Design

The study presented here utilises logistic and linear regression analyses in a large, British birth cohort to test for contemporaneous associations between reading ability and pain in childhood, and longitudinal associations between reading in childhood and pain in adulthood. In all models, pain was the outcome variable, predicted by reading ability and a series of other
covariates. All data manipulation and analysis was carried out in RStudio v1.1.463. All models were tested for model fit against the null model using the `lmtest()` package (Zeileis et al., 2002). Following from this, causal mediation analysis is used to determine whether associations between childhood reading ability and adulthood pain are mediated by adulthood SES.

**Dataset**

Data for this study were taken from the National Child Development Study (NCDS), also known as the 1958 British Birth Cohort. The NCDS began as a perinatal mortality survey and attempted to collect data on all babies born, and their mothers, in a single week during 1958 (Power & Elliot, 2006). As a result, the sample for the first wave of the cohort was very large (n = 17416, Power & Elliot, 2006). Follow up surveys have occurred on the same sample throughout childhood and adulthood, and data collection is ongoing today (Centre for Longitudinal Studies, n.d.). Despite some attrition, sample sizes remain large; for example, at age 42, n = 11419 (Atherton et al., 2008). A description of data collection methods at each wave has been made available by the Centre for Longitudinal Studies (n.d.), and information regarding follow up is available from Power and Elliott (2006). Data is available for download to researchers from the UK Data service, as cited below. Due to the nature of secondary data use, no direct ethical approval was required for this study. Informed consent was obtained from participants (or parents of participants, where appropriate) at the time of data collection (University of London. Institute of Education. Centre for Longitudinal Studies, 2014a)

The NCDS contains a rich variety of variables, covering aspects from health measures at birth, cognitive abilities, employment status, physical and mental health, and more. As a result,
this is an excellent dataset for investigating whether associations hold over time. For this study, NCDS data was used from ages seven, 11, 16 (University of London. Institute of Education. Centre for Longitudinal Studies, 2014b), age 23 (University of London, Institute of Education, Centre for Longitudinal Studies, 2008-a), age 33 (University of London, Institute of Education, Centre for Longitudinal Studies, 2008-b), age 42 (University of London, Institute of Education, Centre for Longitudinal Studies, 2008-c) and age 50 (University of London, Institute of Education, Centre for Longitudinal Studies, 2012). Further details about each of the variables used from these datasets can be found below.

**Measures**

Below is a summary of the measures used in this study. This information is available in table format in supplementary Table S1 and S2.

**Reading Composite.** Reading skill was represented by a composite score, which was created using a similar procedure to that found in Bridges and colleagues (2023). For childhood analyses up to the age of sixteen, a composite of three reading measures from age seven reading was used. The three measures are the Southgate Group Reading Test, a 30 item word-reading test; a teacher rating of the child’s reading ability compared to other children their age; and the grade a child had achieved on a progressive reading scheme (Bridges et al., 2023). For adulthood analyses, a composite comprising of reading measures from ages seven, eleven, and sixteen was used. Two additional measures were taken from age 11; a 35 item reading comprehension test and a teacher rating of the child’s use of books compared to other children their age. Three additional measures were also included from age 16; a 35-item reading comprehension test, a teacher rating of the participant’s ability in the subject of English, and whether the teacher believed the participant could read well enough to cope.
Polychoric correlations of reading variables from all three ages were generated. A Principal Components Analysis (PCA) was carried out on all variables, based on previous research on reading in this cohort (Bridges et al., 2023), and the loadings for the first Principal Component were used to weight the variables for the composite.

For each of the two reading composites, centred z-scores were calculated for each of the three variables for each respondent, and these values were then multiplied by their respective PC loading. The mean of each row was then used as the composite score. In cases where some variables were missing, a mean was generated with only the data that was available. For adulthood analyses, inclusion of reading variables from ages seven, 11 and 16 years was intended to capture the true score of the participant across their childhood, as it has been found that reading ability has not yet stabilised for a substantial minority of children by age seven (Catts et al., 2012). Childhood reading was used rather than adulthood reading for two reasons: 1) it was hypothesised that the accumulation of negative outcomes due to poor reading ability may be influential in the frequency of pain experience, and 2) reading measures in adulthood were self-report only, rather than a mix of self-report, objective tests, and reports from others. Participants who suffered from significant vision problems at age seven, eleven, or sixteen were not included in any analyses.

**Pain Measures.** In this dataset, pain measures were not consistent between data waves. At age 7, the mother of the cohort member was asked if the child suffered from periodic abdominal pain, (1=yes, 0=no), and if the child complained of frequent headaches or migraine (1=yes, 0=no). At age 11, the parent reported recurrent abdominal pain and any headaches or migraines in the past year. At age 16, pain measures consisted of teacher report of whether the
participant often complains of aches or pains, binary coded as *certainly applies*, and parental report of headache or migraine in the past month as a binary outcome.

Pain measures became more consistent in adulthood. At age 23 there were five pain outcomes. Often experiencing headache, eye pain, back pain, and rheumatism were measured as part of the Malaise Inventory, a scale which has been commonly used to assess emotional difficulties (Hirst et al., 1983). All pain types were self-reported by the respondent. Variations on the Malaise Inventory were also administered at ages 33, 42, and 50, although the pain measures were not included at age 50. At ages 23 and 33, migraine was also measured by asking the respondent if they had experienced migraine in the preceding 12 months. At age 50, pain was measured through report of back pain, whether they had seen a doctor in the past year for migraine, and a pain score, which is taken from the Bodily Pain scale of the SF-36 Health Survey (Ware et al., 1983). It is derived from two variables; pain intensity and whether pain experienced interferes with daily life (Ware et al., 1983, Section 3.7). A higher score ordinarily indicates less pain.

**Sex.** Sex was used as a covariate in all analyses, due to its potential confounding influence. Females tend to score more highly on reading measures (Chiu et al., 2006) and there are sex differences in pain prevalence and severity (Bartley et al., 2013; Mogil et al., 2012). Sex as reported at birth was used for all analyses.

**Irritability.** Irritability is a predictor of pain (Keefe et al., 1986), and is also a key measure of neuroticism (Costa et al., 1992), which is associated with self-reported pain (Ramírez-Maestre et al., 2004). Other facets of neuroticism include depression and anxiety (Uljaszek et al., 2009). The NCDS does not contain personality testing for cohort members until age 50, and so inclusion of irritability as a covariate may capture some elements of neuroticism.
not covered by depression and anxiety. Irritability at age seven was defined by parental report, and coded as a binary variable, with frequent irritability coded as 1 and a response of sometimes or never irritable coded as 0. This question was repeated and used again at age 11. At age 16, irritability was measured by teacher report of whether child was irritable, touchy, or flies of the handle. At ages 23, 33, and 42, and 50, irritability was represented by self-report on whether the respondent was easily upset or irritated, as part of the Malaise Inventory.

Disability. Binary disability status was controlled for to ensure that pain caused by disability was not confounding the results, while also considering that some instances of seemingly idiopathic regular or chronic pain could be considered a disability, for example, in the case of fibromyalgia (Giorgi et al., 2022). Disability was defined by parental report at ages seven, 11, and 16, and self-report in adulthood. Disability was not included at age 42, due to the small number of respondents reporting as disabled (n=21).

Restlessness. Given the frequently reported association of ADHD with both reading difficulties and pain, we wished to control for ADHD symptoms in our analyses. A direct measure of ADHD was not available in childhood in the NCDS, and so the Bristol Social Adjustment Guide (BSAG) score for restlessness was used as a proxy, given that restlessness is a commonly reported manifestation of ADHD (Lewandowski et al., 2007), which provides a score from zero (least restless) to four (most restless). At age 7, the BSAG administered at age seven was used. For all remaining waves, restlessness was measured by the age 11 BSAG score, as ADHD has usually presented by age 11 (Kessler et al., 2005).

Coordination. Developmental Coordination Disorder (DCD) is another developmental disorder that shows high co-occurrence with both dyslexia and ADHD (Lino & Chieffo., 2022). Early research suggests that DCD may be associated with a higher-than-average incidence of
joint pain and a potential link to hypermobility syndromes (Kirby et al., 2006). The NCDS does not contain a childhood measure for DCD, and so a teacher rating of whether the child has poor coordination was used. This question was asked at age seven, and this measure was used in the age seven analysis, and again at age 11. The age 11 response was used in all other waves, as DCD has usually presented by age 11 (Smits-Engelsman et al., 2015).

**Depression.** Higher levels of depressive symptoms have frequently been associated with both low reading ability (Willcutt et al., 2000; Eloranta et al., 2019) and different pain types (Lerman et al., 2015; Demyttenaere et al., 2007; Arnow et al., 2006). Evidence suggests that depression and pain share some common biological basis; genetic correlations have been reported between depression and several pain types, including headache, back pain and abdominal pain (Wang et al., 2019). As a result, depression was included as a covariate in our models to prevent any confounding effect. Measures for depression are inconsistent across study waves. A clinical diagnosis of depression was not measured at age seven, and so the BSAG score for depression was used. The BSAG score for depression ranges from zero to fourteen, representing least to most depressed. The BSAG score for depression at age 11 was used for the age 11 analysis. At age 16, parents were asked if the participant had ever seen a specialist for an emotional or behavioural problem, and given free-text space to provide any diagnoses, however in cases where multiple reasons were listed full information was not coded for in the available dataset. The original free-text responses were not available so we used teacher report of whether the child certainly appears *miserable, unhappy, tearful or distressed* often, to improve consistency with other childhood mental health measures and to avoid any systemic bias that may be presented by the coding process. Self-reported depression measures were used in all adulthood analyses; see Supplementary Table S2 for more details.
**Anxiety.** Similar to depression, higher levels of anxiety have been associated with low reading skills, although the direction of causation is unclear (Hendren et al., 2018). Anxious symptoms are also associated with presence of chronic pain (Lerman et al., 2015; McWilliams et al., 2003), and there is a hypothesised shared pathway for the two (Zhuo et al., 2016). Anxiety was therefore included as a covariate in our models. Measures of anxiety are inconsistent across the NCDS. At age 7, the BSAG score for anxiety, a scale from zero to twelve representing least to most anxious, was used. The BSAG Anxiety scale was not available for age eleven, and so in its place, anxiety was measured by parental report of whether the child worries about many things (yes, frequently = 1). At age 16, teacher report was used, based on whether the child certainly appears often worried or worries about many things. Self-reported anxiety measures were used in all adulthood analyses; see Supplementary Table S2 for more details.

**Childhood SES.** Low socioeconomic status has been highly associated with low reading ability in many studies (Fernald et al., 2012; Fung & Chung, 2019; Molfese et al., 2003; Russell et al., 2016; Williams & Silva., 1985). Associations have been found between low SES and likelihood of experiencing pain, for example, back pain (Ikeda et al., 2019) and chronic pain of various types (Prego-Domínguez et al., 2020). For these reasons, SES of the child’s family at the measurement age was included as a covariate at ages seven, 11, and 16. In line with previous research in the NCDS (Ritchie & Bates, 2014), SES was represented by a composite comprising of three measures: socioeconomic status of the father dependent on profession, measured on a five-point scale; whether the residence of the child was owned by the family; and how many people were living in the household per number of rooms. Centered Z-scores were generated for all three measures, and the mean of these values was taken for the composite. Cronbach’s alpha showed an acceptable level of internal consistency between the three variables at each age point.
(Supplementary Table S3) (Taber, 2017). Where data was missing, the composite was generated for the data that was available in order to maximise the sample size.

**Analysis**

**Childhood.** A series of logistic regressions were run to test for associations between reading skill at age seven and experience of different types of pain at ages seven, eleven and sixteen. Covariates were included in these models as described above. Individuals for whom a disabling visual defect was reported at ages seven, 11, or 16 were excluded from the point at which the defect was reported (Supplementary Table S4).

**Adulthood.** Similar analyses were repeated at ages 23, 33, 42 and 50. In each of these analyses, the overall childhood reading composite score, made up of 8 reading measures from ages seven, 11, and 16, was used to predict various types of pain in a series of linear and logistic regressions. Childhood reading score was chosen as we hypothesised that any relationship found between childhood reading score and pain would be mediated by adulthood SES by profession. It has been shown previously that self-reported dyslexia is genetically correlated with SES measures, including work in manual professions (Doust et al., 2022), while in turn, low SES is associated with chronic pain (Prego-Domínguez et al., 2020). All adulthood analyses were repeated with inclusion of childhood SES as a covariate.

**Mediation Analyses.** We hypothesised that an association between childhood reading score and pain in adulthood would be either fully or partially mediated by the adulthood SES of the participant, which we have defined by profession ranking. SES was measured on a six-point scale, operationalised continuously in our model, and each level was representative of a profession type, listed from lowest to highest, as follows: *Unskilled, Partly Skilled, Skilled*. 
Manual, Skilled Non-Manual, Managerial-technical, and Professional. A seventh category, Other, was coded as missing. We tested this hypothesis by carrying out causal mediation analysis, using the R package mediation() (Tingley et al., 2014). Mediation analysis was conducted to test the mediating effect of SES by profession on the relationship between reading and pain for all four pain types at age 42: headache, back pain, eye pain and rheumatism. Age 42 was chosen as all relationships were statistically significant, and we believed enough time had passed to reasonably test our mediation hypothesis. SES was measured on a six-point scale, and each level was representative of a profession, with lower values representing manual jobs, moving up the scale to skilled and professional roles.

Model based inference was used for the mediation analysis, with 1000 simulations and bootstrapping to produce confidence intervals. The mediate() package allows mediation analysis to be undertaken for models with a binary outcome variable by selecting two possible independent variables for comparison, known as the treatment value (Tingley et al., 2014). The treatment value for the mediation was -0.69, which is the 10\textsuperscript{th} percentile of the reading composite score, as this is commonly used as an arbitrary cut-off for the presence of reading difficulties (e.g., Psyridou et al., 2020, Gran Ekstrand et al., 2021, Raatikainen et al., 2021). The control value was 0.18, which is the value of the 50\textsuperscript{th} percentile, or the median value.

Results

Descriptive Statistics

Sample sizes for all models were large, although there were variations across waves (Table 1). Number of individuals reporting the various pain types in each wave can also be seen in Table 1. The NCDS is a nationally representative cohort study, meaning that sex split in the sample reflects that of the UK population (Power & Elliot, 2006). Measures of ethnicity included across
the NCDS are inconsistent, and future cohort studies should accurately report on the ethnicity of the participants from birth. Cronbach’s alpha scores were generated for all sets of variables used for the generation of composite measures, and all values exceeded the minimum acceptable threshold of 0.6 (Taber, 2018).

Table 1. Number of individuals experiencing pain used for each analysis.

<table>
<thead>
<tr>
<th>Pain Type</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 16</th>
<th>Age 23</th>
<th>Age 33</th>
<th>Age 42</th>
<th>Age 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Headache</td>
<td>1101</td>
<td>1432</td>
<td>207</td>
<td>689</td>
<td>684</td>
<td>781</td>
<td>-</td>
</tr>
<tr>
<td>Total N</td>
<td>13241</td>
<td>9135</td>
<td>4524</td>
<td>5545</td>
<td>5002</td>
<td>4984</td>
<td></td>
</tr>
<tr>
<td>N Abdominal pain</td>
<td>1675</td>
<td>993</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total N</td>
<td>13243</td>
<td>9108</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N Aches and pains</td>
<td>-</td>
<td>-</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total N</td>
<td></td>
<td></td>
<td>4748</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N Eye pain</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>251</td>
<td>187</td>
<td>139</td>
<td>-</td>
</tr>
<tr>
<td>Total N</td>
<td></td>
<td></td>
<td></td>
<td>5549</td>
<td>5004</td>
<td>4984</td>
<td></td>
</tr>
<tr>
<td>N Back pain</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1050</td>
<td>1199</td>
<td>1509</td>
<td>524</td>
</tr>
<tr>
<td>Total N</td>
<td></td>
<td></td>
<td></td>
<td>5549</td>
<td>5001</td>
<td>4985</td>
<td>4344</td>
</tr>
<tr>
<td>N Migraine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>789</td>
<td>627</td>
<td>-</td>
<td>125</td>
</tr>
<tr>
<td>Total N</td>
<td></td>
<td></td>
<td></td>
<td>5553</td>
<td>4962</td>
<td></td>
<td>4348</td>
</tr>
<tr>
<td>N Rheumatism</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>212</td>
<td>179</td>
<td>235</td>
<td>-</td>
</tr>
<tr>
<td>Total N</td>
<td></td>
<td></td>
<td></td>
<td>5553</td>
<td>5001</td>
<td>4983</td>
<td></td>
</tr>
<tr>
<td>Total N Pain Score</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3940</td>
</tr>
</tbody>
</table>

Mean 77.60
Standard Deviation 23.04
Reading Ability and Pain in Childhood

All models fit significantly better than the null model. A significant association was uncovered between reading and headache (Figure 1A) and abdominal pain (Figure 1B) at age seven (p < 0.01) (Table 2). The association was highly significant, but the effect size was small. The model predicts that boys in the 75th percentile for reading with no depression, anxiety, irritability, restlessness, disability or signs of poor coordination would have a 6.20% chance of experiencing headaches, compared to a 7.60% chance for those in the 25th percentile. The increase was similar in magnitude for abdominal pain.
Table 2. Results of a logistic regression estimating the odds of regularly experiencing pain at age 7 as predicted by a composite reading score for age 7 and covariates. (***/= P≤0.001, ** = P ≤ 0.01, * = P ≤ 0.05).

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Headache</th>
<th></th>
<th></th>
<th>Abdominal Pain</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio</td>
<td>P</td>
<td>(95% CI)</td>
<td>Odds Ratio</td>
<td>P</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>Reading</td>
<td>0.82***</td>
<td>&lt;0.01</td>
<td>(0.75-0.90)</td>
<td>0.88***</td>
<td>&lt;0.01</td>
<td>(0.81-0.95)</td>
</tr>
<tr>
<td>Irritability</td>
<td>2.33***</td>
<td>&lt;0.01</td>
<td>(1.99-2.73)</td>
<td>1.48***</td>
<td>&lt;0.01</td>
<td>(1.27-1.71)</td>
</tr>
<tr>
<td>Disability</td>
<td>1.69***</td>
<td>0.00</td>
<td>(1.99-2.73)</td>
<td>1.67***</td>
<td>&lt;0.01</td>
<td>(1.32-2.12)</td>
</tr>
<tr>
<td>Restlessness</td>
<td>0.95</td>
<td>0.36</td>
<td>(1.99-2.73)</td>
<td>0.90*</td>
<td>0.03</td>
<td>(0.82-0.99)</td>
</tr>
<tr>
<td>Coordination</td>
<td>1.23</td>
<td>0.27</td>
<td>(0.85-1.77)</td>
<td>0.86</td>
<td>0.37</td>
<td>(0.60-1.23)</td>
</tr>
<tr>
<td>Depression</td>
<td>1.01</td>
<td>0.48</td>
<td>(0.97-1.06)</td>
<td>1.01</td>
<td>0.76</td>
<td>(0.97-1.04)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>1.01</td>
<td>0.60</td>
<td>(0.97-1.05)</td>
<td>1.01</td>
<td>0.66</td>
<td>(0.97-1.04)</td>
</tr>
<tr>
<td>Childhood SES</td>
<td>0.92</td>
<td>0.08</td>
<td>(0.85-1.01)</td>
<td>1.05</td>
<td>0.21</td>
<td>(0.97-1.13)</td>
</tr>
<tr>
<td>Female</td>
<td>1.03</td>
<td>0.67</td>
<td>(0.91-1.17)</td>
<td>1.14*</td>
<td>0.02</td>
<td>(1.02-1.26)</td>
</tr>
</tbody>
</table>

Similar results were found at age eleven (p < 0.01 for headache, p = 0.02 for abdominal pain). (Table 3); an increase in reading composite score was associated with decreased odds of experiencing pain. The effect size was slightly larger than for age seven; for example, girls in the
25th percentile for reading showed a 14.19% chance of experiencing headache, compared to 12.56% for those in the 75th percentile.

**Table 3.** Results of a logistic regression estimating the odds of regularly experiencing pain at age 11 as predicted by a composite reading score for age 7 and covariates. (** = P ≤ 0.001, ** = P ≤ 0.01, * = P ≤ 0.05).

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Headache</th>
<th></th>
<th></th>
<th>Abdominal Pain</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio (95% CI)</td>
<td>P</td>
<td>Odds Ratio (95% CI)</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>0.84*** (0.77-0.91)</td>
<td>&lt;0.01</td>
<td>0.89* (0.81-0.98)</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritability</td>
<td>1.33*** (1.14-1.56)</td>
<td>0.00</td>
<td>1.40*** (1.17-1.68)</td>
<td>&lt;0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability</td>
<td>1.03 (0.76-1.39)</td>
<td>0.86</td>
<td>1.14 (0.80-1.62)</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restlessness</td>
<td>1.04 (0.94-1.15)</td>
<td>0.43</td>
<td>0.93 (0.81-1.06)</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coordination</td>
<td>1.11 (0.77-1.59)</td>
<td>0.58</td>
<td>0.49* (0.27-0.87)</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>1.06*** (0.77-1.59)</td>
<td>0.00</td>
<td>1.06** (1.02-1.11)</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>2.15*** (1.86-2.48)</td>
<td>&lt;0.01</td>
<td>2.19*** (1.86-2.59)</td>
<td>&lt;0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Childhood SES</td>
<td>0.85*** (0.78-0.91)</td>
<td>&lt;0.01</td>
<td>0.88** (0.80-0.97)</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.28*** (1.14-1.44)</td>
<td>&lt;0.01</td>
<td>2.50*** (2.17-2.89)</td>
<td>&lt;0.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Figure 1.** Probability of experiencing different types of pain at age 7 as predicted by logistic regression of childhood composite reading score, with confidence intervals.

No significant association was found neither between reading score and headaches nor reading score and aches and pains at age sixteen although the effect size estimates were comparable to those found at age 7 and 11 years (Table 4). Female sex and anxiety were significant predictors of both pain types.
Table 4. Results of a logistic regression estimating the odds of regularly experiencing pain at age 16 as predicted by a composite reading score for age 7 and covariates. (** = P ≤ 0.01, * = P ≤ 0.05).

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Headache</th>
<th>Aches and Pains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>(95% CI)</td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>0.95</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>(0.78-1.17)</td>
<td></td>
</tr>
<tr>
<td>Irritability</td>
<td>0.78</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>(0.39-1.59)</td>
<td></td>
</tr>
<tr>
<td>Disability</td>
<td>2.06***</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(1.29-3.29)</td>
<td></td>
</tr>
<tr>
<td>Restlessness</td>
<td>0.97</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>(0.73-1.28)</td>
<td></td>
</tr>
<tr>
<td>Coordination</td>
<td>0.67</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>(0.20-2.18)</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>2.00</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.84-4.76)</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>1.80*</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(1.01-3.18)</td>
<td></td>
</tr>
<tr>
<td>Childhood SES</td>
<td>0.81*</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.67-0.98)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.90***</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(1.33-2.39)</td>
<td></td>
</tr>
</tbody>
</table>
**Reading Ability and Pain in Adulthood**

Associations in adulthood were also inconsistent, but showed some patterns. A significant association was found between reading and eye pain and reading and headache for all ages at which those variables were available, with those scoring more highly on the reading composite less likely to experience both pain types (Tables 5-7). An association was found between increased reading score and decreased back pain at all ages for which back pain was measured (Tables 5-8). Prediction of migraine and rheumatism was inconsistent. Reading predicted migraine at age 23 and 33 only (Table 5), while an association for rheumatism was only found at ages 33 and 42 (Tables 5, 6).
**Table 5.** Results of a logistic regression estimating the odds of regularly experiencing pain at age 23 as predicted by a composite reading score from ages 7 to 16 and covariates. (**** = P ≤ 0.01, *** = P ≤ 0.001, * = P ≤ 0.05).

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Headache Dropout Rate</th>
<th>Eye Pain Dropout Rate</th>
<th>Back Pain Dropout Rate</th>
<th>Migraine Dropout Rate</th>
<th>Rheumatism Dropout Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
</tr>
<tr>
<td>0.69***</td>
<td>&lt;0.01</td>
<td>0.73***</td>
<td>0.00</td>
<td>0.65***</td>
<td>0.87*</td>
</tr>
<tr>
<td>(0.60-0.79)</td>
<td>(0.60-0.90)</td>
<td>(0.58-0.73)</td>
<td>(0.76-1.00)</td>
<td>(0.78-1.28)</td>
<td></td>
</tr>
<tr>
<td><strong>Irritability</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
</tr>
<tr>
<td>1.41***</td>
<td>&lt;0.01</td>
<td>1.46**</td>
<td>0.01</td>
<td>1.36***</td>
<td>1.22*</td>
</tr>
<tr>
<td>(1.17-1.70)</td>
<td>(1.10-1.95)</td>
<td>(1.16-1.61)</td>
<td>(1.02-1.46)</td>
<td>(0.80-1.54)</td>
<td></td>
</tr>
<tr>
<td><strong>Disability</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
</tr>
<tr>
<td>2.11***</td>
<td>&lt;0.01</td>
<td>1.75*</td>
<td>0.03</td>
<td>2.05***</td>
<td>1.75***</td>
</tr>
<tr>
<td>(1.45-3.06)</td>
<td>(1.05-2.94)</td>
<td>(1.50-2.81)</td>
<td>(1.22-2.51)</td>
<td>(1.50-4.28)</td>
<td></td>
</tr>
<tr>
<td><strong>Restlessness</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
</tr>
<tr>
<td>1.18*</td>
<td>0.03</td>
<td>1.09</td>
<td>0.44</td>
<td>1.01</td>
<td>1.15*</td>
</tr>
<tr>
<td>(1.01-1.37)</td>
<td>(0.88-1.35)</td>
<td>(0.89-1.15)</td>
<td>(0.10-1.33)</td>
<td>(0.66-1.20)</td>
<td></td>
</tr>
<tr>
<td><strong>Coordination</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
</tr>
<tr>
<td>1.12</td>
<td>0.71</td>
<td>1.10</td>
<td>0.83</td>
<td>0.71</td>
<td>1.50*</td>
</tr>
<tr>
<td>(0.61-2.07)</td>
<td>(0.47-2.57)</td>
<td>(0.41-1.24)</td>
<td>(0.89-2.55)</td>
<td>(0.19-2.35)</td>
<td></td>
</tr>
<tr>
<td><strong>Depression</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
</tr>
<tr>
<td>1.37</td>
<td>0.18</td>
<td>1.70</td>
<td>0.10</td>
<td>1.09</td>
<td>1.63*</td>
</tr>
<tr>
<td>(0.87-2.18)</td>
<td>(0.90-3.19)</td>
<td>(0.69-1.70)</td>
<td>(1.05-2.53)</td>
<td>(0.48-2.59)</td>
<td></td>
</tr>
<tr>
<td><strong>Anxiety</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
</tr>
<tr>
<td>2.26***</td>
<td>&lt;0.01</td>
<td>3.07***</td>
<td>&lt;0.01</td>
<td>1.58***</td>
<td>1.71***</td>
</tr>
<tr>
<td>(1.88-2.72)</td>
<td>(2.28-4.13)</td>
<td>(1.36-1.83)</td>
<td>(1.45-2.03)</td>
<td>(1.14-2.09)</td>
<td></td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
<td><strong>P</strong></td>
<td><strong>Odds Ratio</strong> (95% CI)</td>
</tr>
<tr>
<td>3.36**</td>
<td>0.00</td>
<td>0.79</td>
<td>0.10</td>
<td>1.65***</td>
<td>2.69***</td>
</tr>
<tr>
<td>(2.96-4.43)</td>
<td>(0.60-1.04)</td>
<td>(1.42-1.91)</td>
<td>(2.26-3.21)</td>
<td>(1.42-2.65)</td>
<td></td>
</tr>
</tbody>
</table>
**Table 6.** Results of a logistic regression estimating the odds of regularly experiencing pain at age 33 as predicted by a composite reading score from ages 7 to 16 and covariates. (** = P ≤ 0.01, * = P ≤ 0.05).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>0.68*** &lt;0.01</td>
<td>0.52*** &lt;0.01</td>
<td>0.73*** 0.00</td>
<td>0.81** 0.01</td>
<td>0.58*** &lt;0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.59-0.79)</td>
<td>(0.41-0.66)</td>
<td>(0.65-0.83)</td>
<td>(0.70-0.94)</td>
<td>(0.45-0.74)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritability</td>
<td>2.31*** &lt;0.01</td>
<td>2.42*** &lt;0.01</td>
<td>1.86*** &lt;0.01</td>
<td>1.31** 0.01</td>
<td>1.54* 0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.90-2.81)</td>
<td>(1.73-3.38)</td>
<td>(1.57-2.21)</td>
<td>(1.06-1.62)</td>
<td>(1.07-2.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability</td>
<td>1.81*** 0.00</td>
<td>2.26*** 0.00</td>
<td>1.71*** &lt;0.01</td>
<td>1.62** &lt;0.01</td>
<td>3.35*** &lt;0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.46-2.23)</td>
<td>(1.62-3.16)</td>
<td>(1.44-2.03)</td>
<td>(1.31-2.01)</td>
<td>(2.43-4.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restlessness</td>
<td>0.97 0.76</td>
<td>1.06 0.64</td>
<td>0.93 0.28</td>
<td>0.93 0.41</td>
<td>0.89 0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.83-1.15)</td>
<td>(0.83-1.35)</td>
<td>(0.82-1.06)</td>
<td>(0.78-1.11)</td>
<td>(0.66-1.20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coordination</td>
<td>0.87 0.686</td>
<td>0.97 0.96</td>
<td>0.74 0.26</td>
<td>1.04 0.91</td>
<td>0.20 0.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.46-1.68)</td>
<td>(0.37-2.55)</td>
<td>(0.44-1.25)</td>
<td>(0.55-1.96)</td>
<td>(0.03-1.51)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>1.56 0.186</td>
<td>2.03 0.11</td>
<td>1.47 0.21</td>
<td>1.33 0.43</td>
<td>1.79 0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.81-2.99)</td>
<td>(0.85-4.87)</td>
<td>(0.80-2.69)</td>
<td>(0.65-2.70)</td>
<td>(0.71-4.53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>1.17 0.729</td>
<td>1.37 0.59</td>
<td>1.15 0.74</td>
<td>0.36 0.10</td>
<td>1.14 0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.48-2.84)</td>
<td>(0.44-4.33)</td>
<td>(0.51-2.61)</td>
<td>(0.11-1.21)</td>
<td>(0.32-4.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2.63*** &lt;0.01</td>
<td>1.01 0.97</td>
<td>1.27*** 0.00</td>
<td>2.68*** &lt;0.01</td>
<td>1.77*** &lt;0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.19-3.16)</td>
<td>(0.74-1.37)</td>
<td>(1.10-1.45)</td>
<td>(2.23-3.21)</td>
<td>(1.28-2.44)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 7.** Results of a logistic regression estimating the odds of regularly experiencing pain at age 42 as predicted by a composite reading score from ages 7 to 16 and covariates. (** = P ≤ 0.01, * = P ≤ 0.05).
<table>
<thead>
<tr>
<th>Covariate</th>
<th>Headache Odds Ratio (95% CI)</th>
<th>Headache P</th>
<th>Eye Pain Odds Ratio (95% CI)</th>
<th>Eye Pain P</th>
<th>Back Pain Odds Ratio (95% CI)</th>
<th>Back Pain P</th>
<th>Rheumatism Odds Ratio (95% CI)</th>
<th>Rheumatism P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>0.73*** 0.63-0.83</td>
<td>&lt;0.01</td>
<td>0.49*** 0.38-0.64</td>
<td>&lt;0.01</td>
<td>0.81*** 0.73-0.91</td>
<td>&lt;0.01</td>
<td>0.67*** 0.54-0.83</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Irritability</td>
<td>0.94 0.81-1.10</td>
<td>0.47 0.76-1.52</td>
<td>1.07 0.69-1.52</td>
<td>0.69 0.88-1.12</td>
<td>0.99 0.92-1.12</td>
<td>0.92 0.70-1.19</td>
<td>0.91 0.70-1.19</td>
<td>0.51 1.19</td>
</tr>
<tr>
<td>Restlessness</td>
<td>1.04 0.89-1.22</td>
<td>0.61 0.76-1.37</td>
<td>1.02 0.90-1.37</td>
<td>0.90 0.92-1.17</td>
<td>1.04 0.54-1.54</td>
<td>0.54 0.99-1.54</td>
<td>1.24 1.17-1.54</td>
<td>0.06 1.54</td>
</tr>
<tr>
<td>Coordination</td>
<td>1.75* 1.06-2.88</td>
<td>0.03 1.37-2.88</td>
<td>2.98** 1.40-6.35</td>
<td>0.00 0.36-0.95</td>
<td>0.58* 0.95-1.52</td>
<td>0.03 0.47-2.60</td>
<td>1.10 0.95-2.60</td>
<td>0.82 2.60</td>
</tr>
<tr>
<td>Depression</td>
<td>2.18*** 1.40-3.40</td>
<td>&lt;0.01 0.71-3.88</td>
<td>8.51*** 1.33-6.66</td>
<td>&lt;0.01 2.95</td>
<td>1.98*** 3.88-1.33</td>
<td>&lt;0.01 2.28-6.66</td>
<td>3.76*** 2.12-6.66</td>
<td>&lt;0.01 6.66</td>
</tr>
<tr>
<td>Anxiety</td>
<td>2.00* 1.03-3.91</td>
<td>0.04 0.71-3.88</td>
<td>1.66 0.54-3.88</td>
<td>0.24 1.86</td>
<td>1.00 1.04-1.86</td>
<td>1.00 0.48-2.83</td>
<td>1.16 1.86-2.83</td>
<td>0.75 2.83</td>
</tr>
<tr>
<td>Female</td>
<td>2.83*** 2.39-3.35</td>
<td>&lt;0.01 0.56-1.13</td>
<td>0.79 0.94-1.13</td>
<td>0.20 1.21</td>
<td>0.20 1.07-1.21</td>
<td>0.31 1.06-1.83</td>
<td>1.40* 1.06-1.83</td>
<td>0.02 1.83</td>
</tr>
</tbody>
</table>

Table 8. Results of a logistic regression estimating the odds of regularly experiencing pain at age 50 as predicted by a composite reading score from ages 7 to 16 and covariates. (** = P ≤ 0.01, * = P ≤ 0.05.)
<table>
<thead>
<tr>
<th>Covariate</th>
<th>Odds Ratio</th>
<th>P</th>
<th>Odds Ratio</th>
<th>P</th>
<th>Coef. (^a)</th>
<th>Std Err</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(95% CI)</td>
<td></td>
<td>(95% CI)</td>
<td></td>
<td>(95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>0.72***</td>
<td>&lt;0.01</td>
<td>0.75</td>
<td>0.09</td>
<td>-6.00***</td>
<td>0.64</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>(0.61, 0.85)</td>
<td></td>
<td>(0.54, 1.05)</td>
<td></td>
<td>(-7.26, 4.74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritability</td>
<td>1.82***</td>
<td>&lt;0.01</td>
<td>1.33</td>
<td>0.15</td>
<td>7.41***</td>
<td>0.83</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>(1.48, 2.22)</td>
<td></td>
<td>(0.90, 1.98)</td>
<td></td>
<td>(5.78, 9.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability</td>
<td>1.10</td>
<td>0.40</td>
<td>1.73*</td>
<td>0.01</td>
<td>17.39***</td>
<td>0.83</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>(0.88, 1.36)</td>
<td></td>
<td>(1.18, 2.54)</td>
<td></td>
<td>(15.76, 19.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restlessness</td>
<td>0.96</td>
<td>0.66</td>
<td>1.34</td>
<td>0.06</td>
<td>-0.85</td>
<td>0.73</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>(0.80, 1.15)</td>
<td></td>
<td>(0.99, 1.82)</td>
<td></td>
<td>(-2.28, 0.59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coordination</td>
<td>0.67</td>
<td>0.30</td>
<td>0.49</td>
<td>0.48</td>
<td>-1.17</td>
<td>2.70</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>(0.30, 1.46)</td>
<td></td>
<td>(0.07, 3.588)</td>
<td></td>
<td>(-6.45, 4.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>1.46</td>
<td>0.11</td>
<td>1.12</td>
<td>0.79</td>
<td>17.39***</td>
<td>2.27</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>(0.92, 2.34)</td>
<td></td>
<td>(0.47, 2.68)</td>
<td></td>
<td>(12.94, 21.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>1.55</td>
<td>0.16</td>
<td>2.32</td>
<td>0.11</td>
<td>6.35*</td>
<td>3.24</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(0.84, 2.93)</td>
<td></td>
<td>(0.82, 6.55)</td>
<td></td>
<td>(-0.01, 12.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.93</td>
<td>0.45</td>
<td>3.88***</td>
<td>&lt;0.00</td>
<td>2.71***</td>
<td>0.70</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>(0.77, 1.12)</td>
<td></td>
<td>(2.49, 6.05)</td>
<td></td>
<td>(1.34, 4.08)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: \(^a\)This analysis was a linear rather than a logistic regression (note that lower pain scores represent increased pain in this analysis).

### Mediation Analyses
Mediation analysis was carried out for four pain types at age 42, headache, back pain, eye pain, and rheumatism (Figure 2), given they were each significantly predicted by reading score. The comparison of mediation effects for multiple pain types can help determine whether this is a general or specific association. Results for the models which included childhood SES as a covariate are available in Supplementary Table S9.
Figure 2. Probability of experiencing different types of pain at age 42 as predicted by logistic regression of childhood composite reading score, with confidence intervals. (color to be used on this figure)

**Back Pain.** Addition of SES to the regression model for back pain reduced the predictive power of reading so that it was no longer significant (OR = 0.94, P = 0.33), suggesting complete mediation (Baron & Kenny, 1986). This was confirmed by the causal mediation analysis, which provided a significant estimate for the indirect mechanism of reading on back pain (P < 0.01) (Table 9).
Table 9. Results of four causal mediation analyses testing whether occupation, tested by adulthood SES is a mediator of the relationship between childhood reading ability and four pain types at age 42.

<table>
<thead>
<tr>
<th></th>
<th>Backache</th>
<th>Rheumatism</th>
<th>Eye Pain</th>
<th>Headache</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACME</strong></td>
<td>0.01***</td>
<td>-0.16</td>
<td>-0.00</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>&lt;0.01</td>
<td>0.06</td>
<td>0.96</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>ADE</strong></td>
<td>0.01</td>
<td>0.03</td>
<td>0.02**</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>0.32</td>
<td>0.18</td>
<td>0.00</td>
<td>0.14</td>
</tr>
<tr>
<td><strong>Total Effect</strong></td>
<td>0.02*</td>
<td>-0.13</td>
<td>0.02***</td>
<td>0.02*</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>0.03</td>
<td>0.09</td>
<td>&lt;0.00</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Proportion Mediated</strong></td>
<td>0.52*</td>
<td>-c</td>
<td>-c</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>0.03</td>
<td>-c</td>
<td>-c</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Note: 

- **Average Causal Mediation Effect**, the effect which acts through the mediator variable
- **Average Direct Effect**
- **Values are not interpretable due to opposing effect directions**

**Rheumatism.** Addition of SES to the model reduced the effect of reading so that it became non-significant (OR = 0.82, P = 0.19), suggesting complete mediation. However, direct and indirect effects were in opposite directions, and neither was significant (Table 9), suggesting that the effects have cancelled each other out and this conceptualization is not a useful representation of the constructs involved.

**Eye Pain.** Addition of SES to the logistic regression did not alter the significance of reading ability as a predictor of eye pain, suggesting partial or no mediation by SES (OR = 0.45, P < 0.00). Only the direct effect was significant in the causal mediation analysis, providing
evidence that the association between reading ability and eye pain operates independently of SES (Table 9).

**Headache.** The predictive power of reading ability on headache was reduced below significance when SES was added to the model (OR = 0.86, P < 0.01). Causal mediation analysis indicated that neither the direct nor indirect effect was statistically significant, however both variables acted in the same direction to produce a total significant effect (P = 0.02, Table 9). Taken together, these results indicate that, in combination, both variables work to influence headache.

**Post-Hoc Analysis**

Following the discovery of a mediation effect of SES between reading ability and back pain, we conducted further analysis in an attempt to determine whether the effect of SES was due to the manual work involved in many lower SES jobs, or other SES factors such as access to healthcare. To do this, we conducted an additional mediation model with an added covariate: physical demands of the participant’s job. The participant was asked to rate how physically demanding their work is, choosing from *a lot, moderate amount*, and *very little*. This was added as a categorical covariate, and mediation analysis repeated. The predictive power of SES reduced when physical demand was added to the model, but remained significant (OR = 0.93, P = 0.02). The effect of a lot of physical demands from work, compared to very little, was highly significant (OR = 1.66, P < 0.01), while the effect of moderate physical demands was not significant (OR = 1.06, P = 0.51).
Discussion

We tested for associations between reading ability and different types of pain in childhood and adulthood using a large cohort sample. Associations were found at every age except for age 16, after controlling for relevant covariates. The most consistent associations between reading and pain across waves were for headache, eye pain, and back pain. At age 42, we found evidence of total mediation by SES between reading ability and back pain. We found evidence of confounding effects of SES for the associations between reading and rheumatism and reading and headache. Reading ability and eye pain have an association that acts independently of occupation.

Associations were inconsistent across age groups, which may be due to the differences between the models at each age group. Although all drawn from the same sample, the individuals analysed differed between models due to missing data being present at different time points and for different pain types, meaning models are not directly comparable. Missing data was excluded from the study through listwise deletion, as although there is some indication that better readers are more likely to remain in the NCDS, evidence indicates that this effect is small and the sample is generally representative (Atherton et al., 2008). Question wording used to measure covariates differed between waves, which may have influenced the consistency of results. The covariates in our study were limited by the data available in the NCDS. For example, there was no direct measure of ADHD or DCD in childhood. While the BSAG scales are somewhat useful, their meaning must be interpreted with “caution” (Shepherd, 2013). At several age waves, there is no data on diagnosis of depression or anxiety, and earlier waves often rely on outdated terms which may not align with modern diagnoses. Further research in other datasets is required in order to
assess whether these results are replicable when other variables are used to represent the covariates stated here.

The significant results at age seven and eleven were unexpected. The result for headache was inconsistent with previous research, which did not find an association between reading difficulties and migraines in children (Andrasik et al., 1988), or a significant genetic correlation between self-reported dyslexia and headache (Doust et al, 2022). This finding may have clinical ramifications; children complaining of pain should be screened for reading difficulties, and *vice versa*, in order to ensure that early opportunities for intervention and treatment are not missed. For example, a child who struggles with their reading may also be suffering from headaches, and conversely, children who struggle with headaches should be assessed for reading difficulties.

We propose two hypotheses for the childhood association between reading and pain: 1) Children who struggle with reading may report headache or abdominal pain in an effort to avoid going to school (Heyne et al., 2001; Gonzálvez et al., 2021), or 2) childhood reading difficulties and childhood pain experience could share an underlying biological mechanism. There is little existing evidence to support (2), however, future exploration is warranted. Transcranial magnetic stimulation may be effective for both improving headaches (Yang & Lee, 2020) and dyslexia (Rios et al., 2018), suggesting some biological commonality. Future research may be useful to determine direction of causation in childhood; bidirectional effects may present, whereby children who are experiencing frequent pain find it harder to concentrate on their reading. It has been shown in adults that chronic pain is associated with cognitive impairment (Moriarty et al., 2011).

All four pain types showed an association at age 42, however, mediation analysis showed different results across pain types. The association between back pain and reading ability was
fully mediated by adulthood SES, confirming our hypothesis that the association is indirect. Physical demands of the job significantly predicted backpain, yet the mediation effect of SES remained significant, suggesting that there are elements of SES aside from physical labor that contribute to the association, for example healthcare or living conditions. In order to mitigate some of the risk conferred by occupational status, it may be that additional occupational health support could be provided to poor readers who are at risk of back pain (Rantonen et al., 2018). It is likely that the previous genetic correlations identified between self-reported dyslexia and back pain (Doust et al., 2022), and quantitative reading ability and back pain for more than three months (Eising et al., 2022), arise partly because of a causal relationship from reading to occupation to back pain. One genetically sensitive twin study found an association between educational attainment and low back pain, however this association was no longer significant when genetics were accounted for (Zadro et al., 2017). Further studies are needed to establish whether genetic pleiotropy explains any of the genetic correlation between reading ability and back pain, and what this might tell us about each phenotype.

The associations between rheumatism and reading and headache and reading appear to be due to a confounding effect of adulthood SES. This is interesting given that headache was not genetically correlated with dyslexia in the study by Doust and colleagues (2022), but it was genetically correlated with quantitative reading ability in the study by Eising and colleagues (2022). The study by Eising et al (2020) may be detecting confounding by SES given that reading ability includes poor reading due to poorer literacy environments related to SES. This confounding would be reduced in the genetic study of dyslexia because dyslexia diagnosis in this sample of older participants would have relied on a reading-IQ score discrepancy criterion that shows bias against children of lower SES (Siegel et al., 1998). As for rheumatism, the NCDS
does not provide a definition of rheumatism, and several rheumatic disorders exist, all of which are characterised by pain (Sangha, 2000). It may be that no consistent direct effect could be found due to the individuals with rheumatism representing a heterogenous group.

No mediation effect of SES was found between reading ability and eye pain. Eye pain is unlikely to be made worse through physical labour, unlike back pain which is aggravated by manual work (Lötters et al., 2003). It may be that those who read better read more frequently, and so notice vision problems sooner, avoiding long-term eye-strain. Primary care providers should ensure that poor readers have their eyesight tested regularly, to avoid ongoing issues that go unnoticed. Alternatively, a common biological mechanism may be the underlying cause, but we are not aware of any neurobiological research that supports this hypothesis.

In all cases for which an association between reading ability and pain is present, we must consider the possible role of health literacy. Health literacy is defined as the ability to read and understand content in a healthcare setting, demonstrated validated measures of health literacy that are commonly used (DeWalt et al., 2004). Adequate reading skills are required in order to develop health literacy (Speros, 2005). Poor health literacy has been associated with worse health outcomes in several domains, including increased intensity of chronic pain (Köppen et al., 2018). It may therefore be that participants with poor childhood reading ability are also exhibiting to low health literacy, which is contributing to their difficulties with pain.

Throughout the lifespan, several covariates showed a significant predictive effect on the odds of experiencing pain. For example, being female was shown to increase odds of experiencing headache at every wave between age 11 and 42. This is in line with previous research indicating that global headache prevalence is higher in women than men (Stovner et al., 2022). Results regarding children and young people have been mixed, although a meta-analysis
has found a tendency towards headaches being more common in female children and young people (Albers et al., 2015). Migraine was also found to be more likely in women at almost every age, supporting existing research showing that migraine is more common in women (Buse et al., 2013; Peterlin et al., 2011). Results regarding children and young people have been mixed, although a meta-analysis has found a tendency towards headaches being more common in female children and young people (Albers et al., 2015). Migraine was also found to be more likely in women at every tested age, supporting existing research showing that migraine is more common in women (Buse et al., 2013; Peterlin et al., 2011). Future research relating to the reading-pain association should include separate sex-based analyses, given the strong predictive role of sex.

Either anxious symptoms, depressive symptoms, or both were shown to be significant predictors of odds of experiencing headache at ages 11, 16, 23, and 42. It should be noted that, in these cases, it is plausible that there is a bidirectional effect present in light of previous work. For example, in a low-income sample it was found that anxiety was more prevalent in study participants with headaches, and conversely, participants in the same study who reported experiencing two or more symptoms of anxiety were more 3.24 times more likely to experience headaches (Lucchetti et al., 2013). Similarly, Breslau and colleagues (2000) found evidence of a bidirectional effect at play in the association between depression and migraine, however, this was not replicated for non-migraine headache; headache was a predictor of depression, however, the inverse was not the case. This may explain why anxiety predicted headache at four waves, while depression only predicted headache at two. Given the role of depression and anxiety, it may be prudent to pay particularly close attention to poor readers with mental health difficulties and monitor them for pain symptoms, which may be alleviated with early intervention.
Conclusion

In this study, we have found several associations between reading ability and pain in childhood and adulthood. Some of our findings support a potential neurobiological commonality between reading ability and pain, providing phenotypic evidence to support genetic correlations between reading ability and adulthood pain (Doust et al., 2022; Eising et al., 2022). Future research should use a validated reading measure to further investigate associations between reading ability and pain, and should focus on determining whether reading and pain associations are a result of personal or psychological mediators, shared biological pathways or pleiotropy. Healthcare providers should monitor patients with literacy difficulties for pain symptoms, as reading is a predictor of pain through adulthood to age 50, and early intervention may reduce later prescription of opioids (Martin et al., 2020). Additional preventative support should be made available to poor readers, such as physiotherapy (Martin et al., 2020) and regular eye tests. In addition, earlier intervention may also be valuable. This may be done through ensuring all children are given the necessary support and resources to develop their reading skills, thus potentially improving childhood reading skills and preventing pain caused by poor reading ability. The incorporation of modern molecular genetic methods will be a valuable tool for furthering research in this area.
References


Attributions in Language and Literature. *Sustainability, 13*(13), 7512.
https://doi.org/10.3390/su13137512


McGee, R., Prior, M., Williams, S., Smart, D., & Sanson, A. (2002). The long-term significance of teacher-rated hyperactivity and reading ability in childhood: findings from two
https://doi.org/10.1111/1469-7610.00228


https://doi.org/10.1038/nrn3360

https://doi.org/10.1177/00222194030360010701


https://doi.org/10.1017/s0033291721004141

https://doi.org/10.1111/j.1526-4610.2011.01900.x


relation to social deprivation. Annals of the Rheumatic Diseases, 57(11), 649–655. https://doi.org/10.1136/ard.57.11.649


7.2. Discussion and Conclusion

The submitted work presented in this chapter has builds on the previous genetic work in this thesis by providing evidence for associations between childhood reading ability and childhood and adulthood pain for several pain types. Consistent associations were found between childhood reading ability and pain at age 42, and a causal mediation analysis was conducted to determine whether these associations were due to a mediation effect of SES by occupation. There was clear evidence for a mediation effect of SES between reading ability and back pain, and this association remained when physical demands of the participant’s job was added as a covariate, indicating that the hypothesis that the manual work element of low SES was not responsible for the mediation.

Previous work on longitudinal psychological and psychometric predictors of health has been conducted in the NCDS (Cheng et al., 2017; Carter et al., 2019; Dodgeon et al., 2020), however the research presented in this thesis adds new insights. All three papers cited here found longitudinal associations between childhood measures of cognitive traits and health in adulthood. One study found an association between cognitive ability at age 11 and back pain at age 55, however the association was no longer significant when controlling for adulthood achievement and SES (Cheng et al., 2017). Carter and colleagues (2019) found that a combined measure of reading and maths skill at age 16 was associated with poorer self-reported health at age 50, as part of a wider structural equation model investigating a range of causal pathways.

Finally, Dodgeon and colleagues (2020) found as part of a path analysis modelling study that a latent factor of cognitive ability at 16 showed an association with physical health at age 50. This is particularly relevant to this study, as the latent factor at age 16 included the reading
comprehension test, and the overall cognitive latent factor at 16 was predicted by a cognitive factor at 11 containing the comprehension test, which was in turn predicted by cognitive ability at age seven, which included the Southgate test. The measure of physical health used was a subset of questions of a multi-item health questionnaire (Doll et al., 2000), which was broken down into four domains by the authors, one of which was pain. The work presented in this chapter builds on each of these three studies by demonstrating that there is a specific and replicable association between earlier reading ability and later pain, and in addition, that associations are also present between reading ability and pain when both are measured during childhood. The role of childhood cognitive ability in this relationship should be explored in greater detail, given that previous research has indicated that earlier cognitive ability predicts poorer health outcomes later in the NCDS.

The fact that the work is in line with other results studying similar, or more general, phenotypes in this dataset provides another layer of evidence, on top of genetic evidence and the phenotypic evidence presented in this chapter, that the association between reading ability and pain is worth further investigation in future research. It may be, for example, that children who present with reading difficulties ought to be regularly assessed for problems with pain, and this is particularly significant given that there is evidence which suggests children with potentially co-occurring developmental disabilities such as Autism Spectrum Disorder may find it more difficult to communicate the fact that they are in pain (Brimo et al., 2021; Emerson & Bursch, 2020).
7.3. Additional References

Excluding those cited in submitted work


genome-wide significant loci associated with dyslexia. *Nature Genetics*, 54(11), 1621–1629. https://doi.org/10.1038/s41588-022-01192-y


SECTION FOUR:

GENERAL DISCUSSION
Chapter 8: General Discussion and Conclusion

This thesis set out to explore genetic and phenotypic associations with quantitative reading ability in the NCDS, building on existing work and ensuring that the dataset was properly prepared so that further research may be carried out. In other words, this thesis demonstrates a behaviour genetics project through almost every stage, from preparation of the genomic and phenotypic data, through to genetic analysis, and phenotypic follow-up analysis to further explore genetic associations. There are several key outputs that can be drawn out of this thesis. In Chapter 5, I have imputed a moderately sized genetic dataset with an accompanying rich phenotype dataset to the highest quality imputation reference panel currently available for European populations, the Haplotype Reference Consortium (HRC) reference panel (McCarthy et al., 2016). In addition, Chapter 5 has shown that it is possible to produce reliable and valid functional reading measures by principal components analysis of a small number of objective tests combined with subjective measures of reading ability.

In Chapter 6, the genetic data was used to show that reading composites made up of these measures are heritable, and genetically correlate strongly with each other across ages and with gold standard reading variables measured in independent samples. Furthermore, this study indicated genetic correlations between reading skill in the NCDS and a series of cognitive, health, occupational, SES, attainment, and lifestyle variables from independent samples. This included strong genetic correlations between adulthood non-specific pain complaints and reading at age seven and overall reading ability from age 7 to 33. To investigate this finding further, Chapter 7 reported a phenotypic study of the longitudinal relationship between reading ability and pain in NCDS participants. This study confirmed a phenotypic relationship between childhood reading ability and pain in childhood and adulthood. Mediation analysis between
childhood reading and pain in adulthood showed that the mechanism of this association differs depending on pain type; for some pain types the association was mediated by SES, and in other cases, it was not. This suggests that there may be a shared biological pathway between reading ability and some types of pain, or alternatively, there is another environmental cause for the confounding association.

8.1. Implications

There are two key implications of this thesis; one lies in the way in which it has built on existing research in this area, and the other relates to the novel findings that have been demonstrated. To begin with, I will discuss the ways in which this thesis has built upon previous work in this area. As discussed in Chapter 6, there has already been one molecular genetic study of reading in the NCDS, however this piece of work has been able expand upon the existing work in several ways (Davies et al., 2015). To begin with, in the original piece of work, genetic analysis was conducted on a version of the genomic dataset that was imputed to the less recent 1000 Genomes reference panel (Davies et al., 2015; Auton et al., 2015). It has since been shown that the HRC reference panel, which I used in my work, is able to produce higher quality imputation (McCarthy et al., 2016), and has even been shown to aid in the identification of significant variants in GWAS analysis when analysis is repeated with a HRC imputed version of the data rather than the 1000 Genomes version (Iglesias et al., 2017). This higher quality imputation is not only relevant for the accuracy of results in my own work, but for other, future studies too, as the imputed data is now available for researchers to use upon successful application to the Centre for Longitudinal Studies at University College London. Heritability of reading estimates derived in this thesis are similar to those generated by Davies and colleagues (2015), despite using a composite of reading measures rather than the three reading tests, as
Davies and colleagues used. The composites generated in my own research are each likely to be measuring a combination of decoding and comprehension, whereas the measures used by Davies and colleagues (2015) were specifically measuring word-reading or comprehension. The results of my own work indicated genetic correlation of overall reading ability with educational attainment in independent samples, building on the existing work, which only tested for genetic correlation of educational attainment within the NCDS sample. This indicates that the finding is not due to sampling error, and is robust, particularly given that slightly different reading measures were used in my own work.

Aside from the quality of the existing data, my work has provided a proof-of-concept indicating that it is both possible and feasible to use a composite measure combined of objective and subjective reading measures to capture stable variation in reading ability in the NCDS, and that this can be reliably used in genetic analyses. This supports earlier work in this area, in which high genetic correlation was identified between an objective word reading measure and teacher rating of reading ability, which were based on curriculum criteria, in a sample of seven year old twins (Harlaar et al., 2005). This is significant as the NCDS contains a wide array of rich variables, meaning that genetics of reading research can now be carried out which can investigate a wide range of gene-environment correlations and interactions over the life course. The richness, size, and many years of data collection make the NCDS a rare resource, and this thesis has shown that despite very few psychometric reading measures, the quantitative and categorical reading data converge and can be used for further research to understand how the genetics of overall general reading skill interacts with other traits and environmental factors.

Chapter 6 has added to this existing research on the genetics of reading in the NCDS (Davies et al., 2015) by investigating genetic correlations with a wide range of reading measures
from independent samples. These correlations have broadly been consistent with others published in the literature, particularly in the domains of SES, cognitive abilities, and health outcomes (e.g., Doust et al., 2022; Eising et al., 2022). Chapter 6 of this thesis has also provided evidence of genetic stability of overall quantitative reading from age seven to adulthood. Existing research from twin and adoption studies, including ILTS, has shown that the genetics of decoding and comprehension ability appear to be stable up to late adolescence when reading data is collected at the same point for the whole sample, however this research has not shown whether the genetics of reading are stable into adulthood (Wadsworth et al., 2001; Wadsworth et al., 2016). Other work, utilizing CLDRC data, has sought to test the genetic stability of reading across two time points and included individuals in their twenties at time point two, however data was not collected at the same age point for each individual, with a wide age range of 12.6–26.6 years included in the analysis (Wadsworth et al., 2015). The work of this thesis has been able to expand on these previous findings in two ways: (1) it has provided molecular genetic evidence which supports existing evidence using other methods, and (2) it has extended work on the genetic stability of reading beyond adolescence and into adulthood, using a large sample with data collected at the same age point for all individuals at each data collection point.

In addition to replicating previous associations, my work has further explored the shared genetic architecture of reading and pain, and found evidence to suggest that there may be some shared genetic architecture between these traits. Chapter 7 reported consistent phenotypic associations between reading ability and pain in this cohort, and Chapter 6 indicated that a range of pain measures, particularly general pain measures referring to the whole body or no specific area, are genetically correlated with quantitative reading ability. It has been argued that using genetic data alongside imperfect phenotypic studies can be useful for determining causality and
reducing the number of confounders, and therefore genetic results which support phenotypic results can strengthen confidence that an association exists (Bulik-Sullivan et al., 2015a). Taken together, then, Chapters 6 and 7 provide strong evidence of a relationship between reading ability and pain, on both a genetic and phenotypic level. A genetic association between reading ability and pain could be confounded by SES, however the work detailed in Chapter 7 indicates that phenotypic associations between reading and pain the NCDS dataset are only fully mediated by SES for one of the four pain types tested (back pain), and so a possible shared biological mechanism should be considered to explain reading and pain associations.

Evidence of pleiotropy between reading and pain, and wider health measures, may be indicative that an extension of the ‘generalist genes’ hypothesis is required. The generalist gene hypothesis suggests that all human cognitive abilities may be influenced by a large and overlapping collection of SNPs with small effect sizes, explaining why many cognitive traits are strongly correlated with each other (Plomin et al., 2007; Haworth et al., 2009). Consistently high genetic correlations between cognitive and other traits, particularly ones in the neurological domain such as pain, suggest that the generalist genes hypothesis may be extended to cover other common variable traits outside of cognitive abilities. Alternatively, the reported genetic correlations between reading ability and pain offer some support for the omnigenic model. The omigenic model proposes that one explanation for the low effect sizes of variants seen across GWAS of many traits is that many causal variants may be present in regulatory regions of the genome, and therefore impact the expression of phenotypes via their impact on the expression of other genes, rather than acting on the phenotypic outcome directly (Boyle et al., 2017). Under this model, it is suggested that regulatory variants that are associated with a disorder in a particular tissue are likely to also have an effect for other disorders which are expressed in the
same tissue (Boyle et al., 2017). Further work is needed on the tissue expression of causal genetic variants for reading in order to determine whether this is a plausible explanation.

These results shed new light on the concept of ‘health literacy’, a full discussion of which is available in Chapter 1. Briefly, health literacy is a heterogenous concept with varying definitions, but usually refers to the ability to read and understand medical information (DeWalt et al., 2004). Poor health literacy has been repeatedly associated with poor health outcomes (DeWalt et al., 2004), however, the research presented in this thesis suggests that poor reading ability, regardless of context, is associated with the poor health outcome of pain. In light of this, it must be considered whether the multiple associations reported between poor health literacy and poor health outcomes would be replicated if it was simply general word reading ability and comprehension which were being tested; in other words, it is possible that poor health literacy is simply a mediator between poor reading ability and poor health outcomes. Further research in this area is required in order to test a possible mediation effect of health literacy between reading ability and health, in order to determine whether it is really general reading ability that is predictive of poor health outcomes. Should this be the case, future work is needed to test whether underlying component processes of reading show specific effects, as this thesis has only demonstrated a relationship between overall reading ability and pain, without distinguishing between comprehension and word reading. This would be informative in helping us to understand potential mechanisms through which reading ability acts on pain and other health outcomes; for example, whether there indeed a shared biological element, or whether poor general reading ability leads to difficulty reading health information, and therefore difficulty in understanding and acting upon health advice (Paasche-Orlow & Wolf, 2007). In addition to this, it should be noted that a similar study in the NCDS found an association between cognitive
ability at age 11 and back pain in adulthood, however this association was no longer significant when controlling for a series of adulthood SES variables, including educational attainment and occupation, and adulthood personality factors (Cheng et al., 2017). It is possible, then, that some of the original effect was partially mediated by SES, which would align with the results of this thesis.

8.2. Limitations of this Work

While the work reported in this thesis is robust, there are several limitations which must be considered, partly in order to understand the extent to which these results can be relied upon and partly to identify weaknesses which future research could improve upon. While the NCDS dataset has multiple strengths, including its large sample size, rich phenotyping, long-running nature, and collection of genetic data, both the phenotypic and genetic data also have several weaknesses. While the overall sample size of the NCDS is large, the genetic subsample is much smaller, and no significant SNPs for reading could be identified due to the small sample size (Uffelmann et al., 2021). In addition to this, coverage on some of the arrays used for collecting genetic data was low, likely contributing to the high number of variants of low imputation quality reported in Chapter 5. This reduced the number of variants that would eventually be used in the analysis, and GWAS are most effective when as many variants as possible are used, as this increases the power to detect associations (Das et al., 2018). Further, limited information was available about the genotyping procedures and initial data cleaning, meaning that this thesis was not able to report the historical quality control or individual exclusion criteria.

Aside from weaknesses in the genetic data, the NCDS phenotypic data also has limitations. The key limitation of this dataset is the measurements used; there is a lack of psychometric testing, and while this thesis has shown that this can be overcome to show the full
continuum of overall reading ability by using composite variables, it means that this thesis has only been able to analyse a general measure of functional reading ability rather than specific component processes, such as word-reading and comprehension. This is a limitation as it has been argued that using component processes of reading for genetic research may increase replicability of results across studies, as these phenotypes are reflective of the underlying elements which make up overall reading ability and are phenotypically distinct from each other (Bates et al., 2004), although this replicability is yet to be demonstrated in the case of GWAS studies, as the largest GWAS of quantitative reading phenotypes did not find support for previously identified SNPs at the level of genome-wide significance (Eising et al, 2022). It should also be noted that while there is genetic overlap between processes such as decoding and comprehension, there are also likely to be differences in their genetic underpinnings, as indicated by twin data (Keenan et al., 2006). As a result, the use of the NCDS composites in future reading research must done with caution, and with an understanding that they are only able to inform on general reading performance, rather than any one particular underlying component of it. Future research could compare the genetics of these composites to those of reading comprehension measures, or overall reading meaures, from other datasets, potentially providing additional insight into what the NCDS composites are ultimately measuring.

Aside from being unable to draw a clear distinction between decoding and comprehension, the work in this thesis has not addressed the possible role that cognitive ability may play in any phenotypic or genetic relationship between reading ability and pain, and so the results must be interpreted with an understanding that there is genetic overlap between the two. Existing research indicates that the two are genetically related (eg. Harlaar, Hayiou-Thomas et al., 2005; Eising et al., 2020), however disentangling them is complex, in part due to the
potentially causal nature of reading on cognitive ability (Harlaar, Hayiou-Thomas et al., 2005; Luciano et al., 2017; Ritchie et al., 2004). As a result, it is likely cognitive ability has a role to play in any relationship between reading ability and pain. Now that an association between reading and pain has been established, future research should endeavor to understand the role of cognitive ability in this relationship, from both the genetic and the phenotypic perspective. For example, future research should address whether cognitive ability acts as mediator of this relationship, and in addition, should address whether a proportion of the shared genetics between reading and pain are also shared by cognitive ability. Genomoic structural equation modelling could be valuable for this (Grotzinger et al., 2019).

Other phenotypic measurements from the NCDS have also provided challenges; measurements across waves are inconsistent for a wide range of variables, including reading, SES and health related variables, and data collection did not occur at regular intervals. Due to the age of the dataset, information about some variables from the earlier waves is scarce, meaning it is difficult to determine the true meaning of some of the data. An example of this is the teacher grading at age seven for the reading book level the participant has achieved; very little further information is available for this measure, limiting its informativeness; for example, there is no information regarding what specific criteria must be met for a child to progress to the next reading level, or what level a typically developing child would be expected to achieve at a given age. In addition to this, some variables which would have been useful for this study were not available. For example, ADHD diagnosis was not available in the data, likely due to the age of the dataset, and so a scale measure of restlessness rated between 0 and 4 was used in its place. While restlessness is a feature of ADHD (Lewandowski et al., 2007), the phenotype of restlessness may be more likely to capture hyperactive- rather than inattentive-subtypes of
ADHD, which are characterised by externalising and internalising symptoms, respectively (Bell, 2010). As a result, there may be some measurement error regarding this variable.

A limitation to this project overall is the lack of generalisability of the project, however, this is likely to be true of any equivalent project. The historic nature of the data makes it difficult to generalise results, particularly those which relate to environmental factors, to the current day. It is possible that associations that have been identified in this study may not hold in other cohorts due to cohort effects; it has been shown, for example, that the strength of the relationship between birth weight and cognitive ability in childhood is much higher in the NCDS than it is in the Millennium Cohort Study, which is a birth cohort study beginning in the year 2000 (Goisis et al., 2016). It may be that factors such as changes in schooling since the time of data collection could create cohort effects. Changes in reading instruction have occurred during this period, including the widespread introduction of systematic teaching of synthetic phonics in British schools (Wyse & Goswami, 2008), and the increasing access to and standardisation of schooling more generally (Scoppio, 2000). This presents a clear challenge, as longitudinal data is required to understand the outcomes of reading ability across the lifespan, meaning that a trade-off is required to gain information about longitudinal associations for the full life-span.

8.3. Future Directions

This thesis has shown that it is possible to carry out valid genetics of reading research in the NCDS, which opens up several possibilities for further research going forward. As my research has demonstrated that quantitative reading composites in the NCDS that cover the full variation of reading ability are highly genetically correlated with reading-related measures from independent samples, there is an opportunity to further increase the sample size of future meta-analyses of quantitative reading, such as GenLang, by adding the NCDS sample. All composites
tested in this thesis would be valid contributions, however it may be that the best measure to include would be the overall reading composite, which shows the highest heritability estimate \( h^2 = 0.50, \ SE = 0.17 \), indicating that this measure captures less SES than the other age points. The smaller sample size of the overall composite would not hinder analysis if the data were to be included in meta-analysis, although any of the other composite measures could be included if a larger sample size from this dataset was desired. It should be noted that the overall NCDS reading composite includes measures of both decoding and comprehension, and so would only be suitable in cases when lack of distinction between component processes is not a barrier to inclusion.

The rich range of variables in the NCDS makes this dataset a good choice to carry out research into gene-environment interactions. Gene-environment interactions occur when the environmental and genetic influences on a trait are not acting independently of each other; for example, where a genetic predisposition for a disorder exists which is then amplified by a poor environment, known as the diathesis-stress model (Rende & Plomin, 1992; Pennington et al., 2009), or conversely when environmental conditions are advantageous, underlying genetic differences are more likely to emerge, known as the bioecological model (Bronfenbrenner & Ceci, 1994; Pennington et al., 2009). Existing research on gene-environment interactions in quantitative reading ability has found support for the bioecological model, however this has not always been replicated; higher parental education has been associated with higher heritability of reading in some studies (Kremen et al., 2005; Rosenberg et al., 2011), but not always (Kirkpatrick et al., 2011). Almost all research in this area has been carried out in twin studies, and the NCDS dataset provides an opportunity to study gene-environment interactions in quantitative reading ability using molecular data. The NCDS genetic data may also be used in
gene-environment correlation studies, in order to determine whether genetic pre-disposition of reading influences reading behaviours, which then in turn influence the reading environment of children. It has been shown that exposure to books at home has a positive association with reading development, however this association appears when maternal reading and language skills are controlled for, indicating a genetic correlation is present which affects both reading environment and childhood reading skill (Puglisi et al., 2017). A subsample of participants in the NCDS provided data about their children and parenting, including reading, and so this could be formally test in the NCDS by investigating whether a polygenic score for reading is associated with reading to the child, and reading ability of the child. Similar research has been successful carried out for other traits (see e.g., Pasman et al., 2021)

Aside from use of the dataset, this thesis has clearly demonstrated that further research is needed to understand a potential relationship between reading ability in childhood and health variables in adulthood, in particular pain variables. The results of this thesis have shown that a relationship between the two traits exists, and this also appears to differ for different pain types. Future research should explore this association further, using a combination of phenotypic and multivariate molecular methods analyses to build on this work and to understand the mechanism through which associations between reading ability and different pain types are acting. Specific lines of enquiry that ought to be followed in future include using further multivariate genetic research, to better understand what evidence there is of pleiotropy between reading ability and different pain types, ideally using psychometric tests and clearly defined pain measures to reduce measurement error. Work is needed to more fully understand if there are specific elements of SES which contribute to the association between reading ability and certain pain types, and how that mechanism can potentially be interrupted in order to reduce risk where possible.
8.4. Conclusion

This thesis has walked through a project investigating the molecular genetics of quantitative reading ability and its genetic correlates in a population-based British birth cohort study, and has included phenotypic analysis to explore some of these associations. This thesis has built on earlier work by carefully updating a moderately sized genetic dataset to the HRC reference panel, and demonstrating that although the NCDS contains a limited number of psychometric reading measures, a combination of psychometric and subjective measures can be combined to create reliable composites for measuring longitudinal reading ability. Genome-wide association analysis of reading at multiple age points returned similar heritability estimates to other GWAS of reading, and indicated that the genetic contributions to reading ability are stable through childhood and into adulthood. Reading composites in the NCDS showed strong genetic correlations with component processes of reading, in independent samples, providing further evidence to indicate that reading composites in the NCDS are truly measuring reading ability. The reading measures also showed genetic correlations between several other types of trait measured in independent samples, including SES, attainment, cognitive ability, occupational measures, and health, including pain. Interestingly, a phenotypic relationship between reading ability and pain was shown in childhood, suggesting that the cause of this association may be rooted in early life, including neurodevelopment. In adulthood, SES, and specifically manual occupation, mediated the association between reading and back pain, suggesting that causal mechanisms between reading and pain differ as a function of both time point and pain type.
This thesis has demonstrated that the reading measures in the NCDS can be used to replicate previous findings in the genetics of reading ability or be included in future GWAS meta-analyses of quantitative reading ability, and the longitudinal nature of the dataset offers plenty of opportunities for novel research on gene-environment interplay, and the long-term outcomes of childhood reading ability. It is important that this opportunity is taken; the NCDS is a rare resource, and additional research can only strengthen our understanding of factors that support children’s literacy, regardless of genetic predisposition, so that they may take this skill forward and use it to improve their lives.
8.5. References


