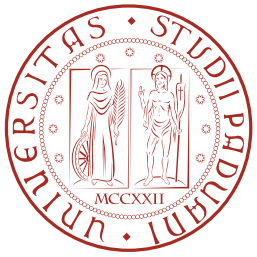

Università degli Studi di Padova

Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata

Dipartimento di Psicologia dello sviluppo e della Socializzazione

Laurea Magistrale in Psicologia Sociale, del Lavoro e delle Organizzazioni

Dyslexia Polygenic Score prediction of Specific Cognitive Abilities

Relatore:

Prof. Massimiliano Pastore

Laureanda:

Ambra Perugini

N. 2015584

Anno accademico 2021/2022

To my son

Contents

Riassunto

Questa tesi si pone l'obiettivo di fornire una panoramica sul disturbo dell'apprendimento meglio noto come Dislessia. L'interesse per questo argomento, tanto coinvolgente quanto complesso, è nato durante l'esperienza di tirocinio a Edimburgo: sotto la supervisione di Michelle Luciano, ho preso parte a un progetto di ricerca sul rischio poligenico di sviluppo della dislessia condotto su un campione di gemelli provenienti dal Brisbane Adolescence Twin Study (M. J. Wright & Martin, 2004; Wainwright, Wright, Geffen, Luciano, & Martin, 2005).

Questo lavoro mira ad esaminare su un campione di gemelli le capacità predittive del Punteggio di Rischio Poligenico (PRS) per la Dislessia (Doust et al., 2022) in relazione a diverse abilità cognitive, quali QI di Performance e Verbale misurate attraverso il Multidimensional Aptitude Battery (MAB), Successo Accademico e Creatività tramite il Queensland Core Skills Test (QCST) e infine Velocità di Elaborazione di Informazioni misurate attraverso il Choice Reaction Time (CRT), Inspection Time (IT) e Digit Symbol (Subtest del WAIS-R).

Dopo l'introduzione, la tesi si articola in quattro capitoli: il primo capitolo introduce il tema della dislessia toccando tematiche quali le problematiche di definizione e classificazione a cui è soggetta e i conseguenti problemi legati alla frequenza del disturbo nella popolazione, che varia dal 5% fino al 20% a seconda dei criteri dia-

gnostici utilizzati (Wagner et al., 2020). Infine, l'importanza di un'identificazione precoce è stata messa in luce (Ferrer et al., 2015; Lonigan, Purpura, Wilson, Walker, & Clancy-Menchetti, 2013).

Il secondo capitolo, che precede l'esposizione del progetto, affronta il tema dell'ereditarietà della Dislessia, con una rassegna della letteratura di riferimento, basata principalmente sugli studi sui gemelli. Sono stati analizzati diversi studi riguardo la relazione della dislessia con alcune capacità cognitive, prese in esame in questa tesi. È stato spiegato inoltre, in riferimento alla genetica di tale disturbo, cosa sono il Genome-Wide Association Study e le principali scoperte sui geni della dislessia, tenendo in considerazione il problema della replicabilità. Infine, sono stati introdotti i Punteggio di Rischio Poligenico (PRS) per poi analizzare gli ultimi studi che hanno generato PRS da grandi Genome-Wide Association Studies (GWAS) in relazione alla capacità di lettura (Gialluisi et al., 2019; Price et al., 2020; Gialluisi et al., 2021), con particolare interesse per lo studio di Doust et al.(2022) e il loro Punteggio di Rischio Poligenico che è quello utilizzato in questa ricerca.

Nel terzo capitolo vengono esposte le ipotesi della ricerca, la natura del campione, una descrizione dei Test Cognitivi e la metodologia. L'ipotesi di partenza prevede che il rischio poligenico avrebbe predetto basso successo accademico e bassi livelli di scrittura creativa, bassi punteggi in Verbal IQ ma non in Performance IQ e peggiori tempi di risposta in Velocità di Elaborazione di Informazioni per Inspection Time (IT), Choice Reaction Time (CRT) e Digit Symbol. La scelta dei test utilizzati è stata fatta attraverso la Parallel Factor Analysis e osservando la matrice di correlazione. Le distribuzioni dei Test Cognitivi presi in esame sono state illustrate con degli istogrammi e descritte in termini di Media, Deviazione Standard, Curtosi (Kurtosis) e Simmetria (Skewness). L'analisi dei dati è stata spiegata insieme alle statistiche prese in esame.

Il quarto capitolo si concentra sull'esposizione dei risultati. La ricerca è stata condotta con il software di analisi statistica R. I dati sono stati analizzati con un Modello Multivariato Multilivello con il pacchetto 'brms' (Bürkner, 2018). Nella fase iniziale è stato effettuato un confronto tra modelli per comparare il modello con il Punteggio di Rischio Poligenico con un modello nullo utilizzando i model weights. Quello che è emerso è una evidenza del modello target di poco superiore al modello nullo, evidenziando la necessità di interpretare con cautela i risultati dell'analisi. Quindi, i risultati sono stati divisi in sottosezioni per ciascun test cognitivo mostrando la stima per ciascuna equazione con il suo intervallo di credibilità e la varianza spiegata con gli R^2 . Sono inoltre inclusi dei grafici che rappresentano la capacità del modello di prevedere i dati simulati sulla base dei parametri e i grafici degli effetti condizionali.

I risultati hanno supportato le prime due ipotesi di partenza della nostra ricerca relative al successo accademico e al IQ, mentre la terza ipotesi sulla Velocità di Elaborazione di Informazioni è stata supportata solo per il sub-test Digit Symbol. È importante sottolineare come la varianza marginale, ricavata mettendo a confronto gli R^2 del modello target e il modello nullo, fosse sempre di piccole dimensioni e in alcuni casi nulla, mostrando quindi pochissima varianza spiegata dal Rischio Poligenico per la Dislessia.

Questo studio potrebbe costituire un punto di partenza per future ricerche, essendo il primo a utilizzare un Rischio Poligenico che spiega fino al 6% della varianza in test di abilità di lettura per indagare la relazione a Test Cognitivi Specifici. I risultati di questa ricerca possono indirizzare le ipotesi a priori di studi futuri permettendo così un approccio di tipo Bayesiano. Inoltre, lavorare sul rischio poligenico potrebbe aiutare a identificare con maggiore precisione e affidabilità i bambini a rischio di dislessia, prima che sviluppino le abilità di lettura, contribuendo a fornire

loro il supporto necessario e un intervento specifico sulla capacità di lettura il prima possibile, garantendo loro possibilità di miglioramento e di colmare il divario tra loro e i loro coetanei.

Introduction

The aim of this thesis is to give an overview on the Learning Difficulty well known as Dyslexia and to describe the study conducted on a twin sample from the Brisbane Adolescence Twin Study (M. J. Wright & Martin, 2004; Wainwright et al., 2005) using a Polygenic Risk Score for Dyslexia, which was first computed by Doust et al. (2022) in a Genome-Wide Association Study on 51,800 participants self-reporting a diagnosis of Dyslexia and 1,087,070 controls, to predict cognitive abilities and their correlates such as IQ, Academic Achievement and Information Processing Speed. The research project was developed during my internship at the University of Edinburgh where I learned in great detail about Dyslexia and related cognitive abilities and I was trained to handle complex data with the statistic software R under the supervision of Michelle Luciano, who guided me through the process and has been a great mentor.

In the first chapter, an introduction of the learning disability will be given, with an insight on the ongoing debate related to its classification, varying widely between different theoretical approaches. An overview of its frequency in the population will be given, which can be as low as 5% up to 20% (Wagner et al., 2020), again highlighting a problem in inconsistency between different definitions. Insight on the criticality of early identification will close this chapter, showing evidence of the

benefits of interventions as early as in kindergarten (Ferrer et al., 2015; Lonigan et al., 2013) allowing to bring the Dyslexic children closer to their peers in their reading abilities and even close the gap between them.

In the chapter preceding the exposition of the project, there will be an overview on what is known until now on the genetic inheritance of Developmental Dyslexia, with a review of related literature, which is mostly based on Twin and Family studies. A general explanation of the design of such studies will be given, to move into more detail in the relation of dyslexia with other cognitive abilities involved in this study. It will be explained what is a Genome-Wide Association Study (GWAS) and the main findings on the genes of Dyslexia, taking in consideration the problem of replicability. Lastly, Polygenic Risk Scores will be shortly introduced to then analyze the last studies extrapolating them from large GWAS in relation to reading ability (Gialluisi et al., 2019; Price et al., 2020; Gialluisi et al., 2021), with particular interest at the study from Doust et al.(2022) and their Polygenic Risk Score.

In the last section of this work the hypothesis of our study will be explained in great detail. The same will be done for the Cognitive Tests involved in the analysis as dependent variables, showing their distribution with histograms and the properties of the groups of participants. The data will be analysed with a Multivariate Multilevel Model with the package 'brms' (Bürkner, 2018) and a model comparison will be carried out in the initial phase to compare the model with the Dyslexia Polygenic Risk Score to a null model (2). Then, the results will be displayed in subsections for each Cognitive Test showing the Estimate for each equation with its Credibility Interval and the variance explained with R^2. Graphs representing the ability of the model to predict simulated data will be included, together with graphs of the conditional effects.

This work is of great value for the research community focusing on dyslexia because it uses for the first time such a powerful Polygenic Risk Score for Reading Ability to predict Cognitive Abilities and Academic Achievement. As it will be emphasised later on, it is of great social value to be able to accurately identify children at risk of developing Dyslexia in order to give them the support needed and a specific intervention on reading ability as soon as possible to give them a better chance to close the gap between them and their peers.

Chapter 1

Definition of Dyslexia

Reviewing the history of dyslexia since it was introduced to the academic world, one realizes that there are a multitude of studies trying to categorize the learning disability. Over the last decades more interest grew about dyslexia and more researchers tried to find the causes and effects of it. More than 50 definitions can be found in the literature in an attempt to identify the right one regarding dyslexia (Ott, 1997). Many people, including Government committees and Non Profit Organizations, have tried to accomplish such categorization with different degrees of success (Michail, 2010). The definitions vary and depend on the scientific backgrounds of the individual researchers and what they conceptualize as the underlying cause of dyslexia (Ott, 1997). Finding the definition that covers all the areas of interest is quite impossible but all those definitions have something in common, they all agree on a characteristic of dyslexia: the reading accuracy deficit; then other criteria are followed with less agreement: cognitive impairment, age discrepancy, IQ discrepancy and spelling fluency accuracy (Reiss & Brooks, 2004). Even the name used to define the disability brings researchers apart. Some emphasize the 'developmental' component by calling it "developmental dyslexia" to differentiate it from "acquired dyslexia" that is the impairment following a trauma or injury, others emphasize its

the specificity by calling it "specific reading disability"(Friend, Pennington, Smith, & Gilger, 2013).

Regardless of all the ongoing discussion on the definition and denomination, when it comes down to a clinical diagnosis, the DSM-V is one to be used (Association, 2013). According to the Diagnostic Manual dyslexia is under the category of Specific Learning Disability and can be found in the section of Neurodevelopmental Disorders. The main criteria are: difficulties in learning and in the use of academic skills and lack of response to intervention after 6 months and impairment of scholastic abilities compared to their peers. Importance is given to exclusionary criteria as well: the subject should not have a diagnosis for mental retardation, problems of uncorrected hearing and vision, other mental or neurological disorders, psychological or social adversity, language problems or inadequate education. But all those criteria are general to the Specific Learning Disability, the specification for dyslexia is given by problems in accurate or fluent word recognition, poor decoding and poor spelling abilities. The DSM-V gives the criteria for a diagnosis but not a clear definition specific to dyslexia.

A widely accepted definition is yet to be found but one of the most predominant is the one used by Lyon, Shaywitz, and Shaywitz(2003):

"Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowl-

edge."

This descriptive definition of Dyslexia on one hand has been very useful to inform people about the different characteristics and manifestations of Dyslexia (Michail, 2010), but it gained some critiques on the terminology used. For Snowling, Hulme, and Nation(2020) the disability is far from specific. It includes aspects that are useful and can provide guidance to practitioners in order to help them identify and assess dyslexia and can be useful for teachers as well and other professionals who are involved to a child/adult's education as it gives the characteristics that are associated with dyslexia. But as it is explained in the Dual Route Model by Coltheart, Rastle, Perry, Langdon, and Ziegler(2001), dyslexia can be divided into subtypes, the main ones being surface and phonological dyslexia, with studies supporting this theory (Castles & Coltheart, 1993; Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996; Stanovich, Siegel, & Gottardo, 1997). There is now compelling evidence for the heterogeneity within the population classified as having developmental dyslexia.

Since the causes are not so easily identifiable and it seems like each case is specific and different from the others, Castles(2006) suggests to move towards model-based diagnoses of dyslexia. The teachers can still use Lyon et al.(2003) definition as a general guide to look for signs of the learning disability and monitor the progress of the child/adult that displays them and they can ask for further support and request for the individual to be assessed for dyslexia if they persist through time and adequate support. On the other hand, scientists need much more precise criteria to conduct research (Gaddes, 2013). Therefore, a revision of it's definition has been suggested (Castles, 2006). Yet the definition by Lyon et al. (2003) among specialists and researchers is still the most widely used.

To draw a conclusion, what emerges is a difficulty in having a clear classification

of the disorder. When defining the criteria needed for the diagnosis of this learning disability they can vary greatly between the research criteria and the criteria used to qualify the child for special education (Friend et al., 2013).

Frequency of Dyslexia

The prevalence of dyslexia varies widely and this can be attributed to the discrepancies within the different definitions. Rates can be less than 5% up to 20% depending on the differences between diagnosis and problems with the classification (Wagner et al., 2020), leaving individuals with dyslexia, their families and professionals of the field with confusion. The indicators of this Learning Disability are distributed on a continuum throughout the population (Fletcher, Lyon, Fuchs, & Barnes, 2018), and a cut point has yet to be established to determine whether the condition is present. The position on the continuous of the cut-point by itself has an impact on the prevalence estimates. When the scoring criterion is 1.5 standard deviations or greater below the mean for the measure used to assess reading, the estimates of prevalence usually range from 3% to 7% (Fletcher et al., 2018; Peterson & Pennington, 2012; Snowling & Melby-Lervåg, 2016). If the cut off used is less stringent, then results tend to show higher prevalence.

Problems arise even when different operational definitions are used. As reported in Brown Waesche, Schatschneider, Maner, Ahmed, and Wagner(2011), depending on the definition used for the diagnosis, agreement on the same sample can be as low as 31%. When reassessing longitudinally, the stability for the definitions varied from 24% to 41% depending on the criteria used. Wagner et al.(2020) suggest two ways to improve the reliability of the diagnosis of dyslexia: firstly, the operational definition of Dyslexia should expand to include multiple indicators, then a Bayesian model that uses prevalence estimates to form the informative priors, such as comorbidity

of ADHD or having an affected parent, would also improve the diagnosis. The author also brought awareness on the problem of defining dyslexia as a spectrum. The governaments have policies for learnind disabilities and they need a clear cut-off to determine if an individual is eligible for special education but if the disability is seen as a continuum it is harder to determine which cases should be included for special support. It is now known that dyslexia can occur throughout the range of cognitive abilities (Wagner et al., 2020) and regardless of the discrepancy in IQ and eligibility for special education, individuals and their families should be made aware of their condition in order to acquire access to a more specific support during their education.

Early Identification

Early diagnostics of dyslexia can be a key step in commencing interventions early on aiming to reducing and alleviating the negative aspects of dyslexia which include both academic difficulties and socioemotional problems (Ozernov-Palchik et al., 2017). Early research conducted by Fawcett and Nicolson(1995), showed that the earlier the diagnosis the better the chances of remediation, with 82% of children diagnosed in grades 1 and 2 catching up with their chronological age group, compared with 46% of those diagnosed in grade 3, and lowering to only 10-15% for those as late as in grades 5 to 7 grade.

When adequate support from the teachers in developing reading skills is neglected, serious consequences arise. Children with dyslexia, as they progress in school, have a chance to improve in reading accuracy when given good instructions (Ferrer et al., 2015), what tend to remain a problem is the lack of fluency, that persists as a lifelong problem. In their overall academic achievement kids with reading disability tend to have poorer performance compared to their peers. A study

conducted by Ferrer et al.(2015) found that dyslexic readers show an achievement gap from first grade, persisting through their education. This gap between dyslexic readers and their peers without learning difficulties leads to serious consequences. In general, children who did not achieve fluent reading skills by third grade had lower graduation rates in high school (Hernandez, 2011), even lower when SES of the family was taken into account.

Another two studies (Campbell et al., 2014; Heckman, Pinto, & Savelyev, 2013) showed that when a proper early intervention is not carried out, those individuals will become adults with lower earnings because of lowered college attainment and higher levels of unemployment. On the contrary, those studies highlight as well how high quality early childhood programs focused on reading skills have shown to be effective in reducing crime, promoting education and raising Socioeconomic Status.

Ferrer et al.(2015) suggested that an early intervention for kids at risk of developing dyslexia should start as early as in kindergarten or even preschool. A study by Lonigan et al.(2013) showed the benefits of early interventions on 324 preschool children (mean age = 54.32 months, SD = 5.88) from low-income backgrounds. In this way, it may be possible to bring their trajectories closer or even close the gap between dyslexic children and their peers. This is also why so much effort and budget is used to try to identify the genetic links of dyslexia, so that children at risk could be identified more accurately early in preschool and intervention to strengthen their reading ability could start as early and accurately as possible.

In line with the studies cited above, the analysis of a study conducted by Richardson and Wydell(2003) suggests that dyslexia has deleterious consequences for the likelihood of academic progression, completion and achievement in students who are enrolled in programs of study at institutions of higher education, but it also shows

the importance of adequate support from the institutions that can make the completion rates of dyslexic students match those of students with no disability. When dyslexic students are given appropriate resources from the university, and with a good amount of commitment, they can reach high levels of success.

This evidence supports the importance of the research on the causes and genetic links of Dyslexia: finding better assessment tools so that children at risk can get the support they need as soon as possible and keep up with their peers. Vellutino, Scanlon, and Tanzman(1998) showed in their research that an early intervention on children who perform poorly in reading can allow the children to be brought up to an average level of performance if they are identified in an early stage of the reading development.

Chapter 2

Dyslexia and genetic links

It is important, when talking about dyslexia, to take into account the genetic component of the disability. Genetic inheritance is one of the major risk factors for reading disorder, together with home learning environment and family history (Luciano, 2017). The specific genes influencing reading ability are still largely unknown, but twin and family studies reveal the genetic and environmental structure underpinning reading ability, the way it relates to development, and how environmental factors modify it (Luciano, 2017). In this chapter we will look at past literature of the research designs used in behavioral genetics and the heritability of dyslexia, the heritability of the cognitive abilities involved in the study, the specific genes discovered to be linked to dyslexia and the use in research of Polygenic Risk Scores.

Quantitative genetics is often used to study behavior and the influence of the relative contribution of genetics and environment to individual differences using twin, adoption and family data (Rijsdijk & Sham, 2002; Hensler, Schatschneider, Taylor, & Wagner, 2010). Those last two study designs, involving adoption cases and families, have some limitations: the family study is not able to discriminate environmental from genetic factors and adoption data is often classified and hard

to get access to it, and even if it can be obtained selective placement and prenatal influences can be a bias. Therefore, twin studies are the most popular design in the behavioural genetics research field (Rijsdijk & Sham, 2002). The existence of two types of twin pairs, monozygotic (MZ) and dizygotic (DZ), provides a natural experiment for disentangling genetic from environmental factors and to compare a trait's resemblance. MZ twins are identical and share 100% of the genes, while DZ twins are fraternal twins, therefore they share approximately 50% of the genes like siblings. Twins grow up in the same family, therefore they go through some common experiences. In the case of MZ twins, all the trait differences are then attributable to the environmental differences. In DZ twins, the differences are both genetic and environmental.

As explained by Rijsdijk and Sham (2002), statistical models, such as structural equation models(SEM), can use this data to analyze the contributions of genes, shared environment and unique environment related to a specific trait. With SEM the relative importance of those latent factors can be inferred in relation to the observed traits.

The sources of genetic and environmental variation considered in behavioral genetics are as follows (Rijsdijk & Sham, 2002; Hensler et al., 2010):

- additive genetic influences (A), the sum of the effects of the individual alleles at all loci that influence the trait.

- non-additive genetic influences (D), including dominance (where some alleles or versions of a gene only exert their effect when present on both chromosomal copies in a cell) and epistasis (interaction effects between genes).

- common environmental variation (C), environmental influences shared by fam-

ily members, such as socio-economic status (SES), childhood diet, parenting style etc.

- unique environmental differences (E), all the events that differ throughout the members of the family and residual variance not otherwise explained, eg. different parental treatment, accidents, school environment etc.

Those four factors summed together represent the total phenotypic variance (P). Thanks to Twin data the different variance components can be estimated, because MZ and DZ twins have different degrees of correlation for the genetic components A and D but the same degrees of correlation for the environmental components C and E.

Whereas, when the correlation between MZ twins his higher than the correlation between DZ twins genetic influences can be implied (Andreola et al., 2021; Hensler et al., 2010), A, C and E effects are then expected. When the correlation of MZ twins is twice as much the DZ correlation, non additive genetic effects are expected. Moroever, when the DZ twin pair carrelation is higher than half of the correlation between MZ twins, an influence of the common environment, which is shared within twins, is implied. Unique environmental factors are considered to be specific to the person and therefore are not shared between twins. Since identical twins share not only genetic make-up but also part of the environment, a deviation from 1 in the correlation of the MZ twins provides an estimate directly for non shared enviromental influences. In the classical twin design two variables can not be estimated: D influences and C influences, they are confounded and need different models to be estimated (Andreola et al., 2021)".

To estimate the heritability (h^2) of a trait the Falconer's formula uses twin correlation to give an index of the relative contribution of genetic effects to the total

phenotipic variance (Falconer & Mackay, 1960).

$$h^2 = 2(r_{MZ} - r_{DZ})$$

In the formula r is the intraclass correlation coefficient. The relative contributions of the shared and non-shared environmental effects are: $c^2 = r_{MZ} - h^2$ (or $2r_{DZ} - r_{MZ}$) and $e^2 = 1 - h^2 + c^2$.

Rijsdijk and Sham (2002) highlight the importance of the assumptions involved in classical twin studies. Researcher must be aware of what the assumptions imply and the extent to which they can be considered realistic in relation to the trait of interest. Those are the four assumptions:

- MZ and DZ twin pairs share their environments to the same extent.
- Gene-environment correlations and interactions are minimal for the trait.
- Twins are no different from the general population in terms of the trait.
- Mating in the population occur at random (no assortment).

Only if those assumptions are met then the classical twin design can be a powerful tool to estimate separately genetic and environmental factors of a trait. The most debated one is the equal environment assumption, stating that the similarities in both twin types due to environmental influences are roughly the same for those reared in the same family. According to Rijsdijk and Sham (2002) if this assumption was not to be true and MZ twins experienced more similar treatment compared to DZ twins, then there would be an over estimation of the genetic effect in MZ twins and an underestimation of the shared environmental effect, resulting in increased

MZ correlations relative to DZ correlations.

Now that we have given an overview of twin design studies we can move forward to analyze the literature on Dyslexia heritability and the heritability of the related cognitive abilities used in our research.

Heritability of reading ability

Behavioral genetic studies in the past showed that dyslexia and reading-related skills are highly heritable (Light, Defries, & Olson, 1998; Byrne et al., 2006; Dale, Harlaar, & Plomin, 2005; Betjemann et al., 2008; Stevenson, Graham, Fredman, & Mcloughli, 1987; Andreola et al., 2021; Light & DeFries, 1995; Hawke, Wadsworth, & DeFries, 2006; Harlaar, Spinath, Dale, & Plomin, 2005; Stevenson, 1991).

Many twin studies demonstrated reading to be highly heritable throughout the developmental stages. Already back in 1998, Light et al.(1998) found a high heritability estimate of reading in both the reading-disabled and control twin samples, 0.70 and 0.42, respectively. More recently, an international longitudinal study on preschoolers showed that factors having an influence on pre-reading abilities such as Phonological Awareness (PA) have a h^2 of .61 (Byrne et al., 2006). Following up on the same children revealed that in kindergarten PA continued to show even higher heritability ($h^2 = 0.63$) and, assessing early reading ability, they found it to be highly heritable as well ($h^2 = 0.70$). Even in second grade heritability on the same kids was as high as 0.84 for word reading, 0.74 for phonological decoding and 0.67 for reading comprehension (Byrne et al., 2009). In the United Kingdom a much bigger sample of 5544 children was tested for reading skills at 7 years old, finding similar heritability estimates for word reading ($h^2 = 0.63 - 0.74$) and teacher-assessed literacy ($h^2 = 0.65 - 0.66$)(Dale et al., 2005).

In a study on an older US sample (mean age, 10.3), reading continues to have substantial genetic influence later on in school years on word reading ($h^2 = 0.76$) and reading comprehension ($h^2 = 0.67$) (Betjemann et al., 2008). At 5 years follow up the same measures showed again high heritability estimates ($h^2 = 0.68$) for word reading and $h^2 = 0.60$ for reading comprehension). Lower estimates were found for another sample in UK (mean age, 13) with heritability of 0.19 for word reading and 0.44 for reading comprehension (Stevenson et al., 1987).

A more recent study by Andreola et al. (2021) gives an overview of heritability for the main phenotypes involved in reading ability. The meta-analysis uses a multilevel approach to synthesize the results of 49 twin studies on behavioral genetics focusing on reading-related neurocognitive components; the sample size in total is of 38,000 individuals spanning 4.1-18.5 years of age. All the studies were based on the classic twin method that was explained in the previous paragraph.

The phenotypes included in the analysis are the following:

- General reading (reading speed and accuracy), divided into Letter-word knowledge, Phonological Decoding and Reading Comprehension.

- Spelling, defined as the ability to form words with the correct letters in the correct orders.

- Phonological Awareness (PA), the ability to recognize and manipulate linguistic sounds apart from their meanings.

- Rapid Automatic Naming (RAN), the ability to quickly name aloud objects, pictures, colors, or symbols (letters or digits).

- Language, such as receptive and expressive vocabulary, oral language and naming abilities.

As a first step, twin correlations were metanalyzed (MZ correlation and DZ correlation, respectively). Both the MZ correlations (r_{MZ}) and the DZ correlations (r_{DZ}) were transformed into the Fisher's Z scores ES_Z (ES_{ZMZ} and ES_{ZDZ}) and the sampling variance for each neurocognitive component was estimated. Next, the ES_{ZMZ} and ES_{ZDZ} were transformed back to MZ correlations (rMZ) and DZ correlations (r_{DZ}) for interpretation purposes. Additionally, the heritability was calculated by applying the Falconer's formula (Falconer & Mackay, 1960).

The results of this meta-analysis showed that for General Reading the overall MZ correlation was 0.79 ($ESZ_{MZ} = 1.08$, $S.E. = 0.09$, $t = 11.92$, $p < 0.001$, 95 % $CI = 0.90 - 1.26$) and the DZ correlation was 0.46 ($ES_{ZDZ} = 0.50$, $S.E. = 0.02$, $t = 29.23$, $p < 0.001$, 95 % $CI = 0.47$). The overall heritability, calculated with Faloner's formula (Falconer & Mackay, 1960) was 66%, with a shared environment effect of 13%, and the non-shared environment effect, which included measurement error, was of 21%. It was also revealed how 15% of the variation in general reading performance in middle school is independent from the performance in elementary school, indicating that the school grade level is a moderator. MZ and DZ correlations of this moderator showed a decrease of both heritability and shared environmental effects, and an increase in non-shared environmental effects, from elementary school to middle school. Genetic and environmental influences seem to be similar for sex and spoken language, those two factors didn't show moderating effects on the magnitude of MZ and DZ.

For Letter-word knowledge, the meta-analysis yielded an overall MZ correlation of 0.79 ($ES_{ZMZ} = 1.08$, $S.E. = 0.06$, $t = 17.98$, $p < 0.001$, 95% $CI = 0.96 -$

1.20) and an overall DZ correlation of 0.48 ($ES_{ZDZ} = 0.53$, $S.E. = 0.03$, $t = 17.73$, $p < 0.001$, 95% $CI = 0.47$). The heritability was estimated at 62%, with 17% shared environmental effect, and 21% of phenotypic variance accounted by non-shared environmental factor and measurement error. Once again sex did not moderate the magnitude of MZ and DZ correlations for this phenotype.

Phonological Decoding was found to have a correlation of 0.78 on MZ ($ES_{ZMZ} = 1.04$, $S.E. = 0.07$, $t = 13.99$, $p < 0.001$, 95% $CI = 0.88 - 1.19$) and of 0.44 for DZ ($ES_{ZDZ} = 0.47$, $S.E. = 0.03$, $t = 16.51$, $p < 0.001$, 95% $CI = 0.41$). The overall heritability was 68%, with shared environment effects of 10%, and non-shared environment effects of 22%.

The overall correlation for Reading Comprehension in MZ was 0.79 ($ES_{ZMZ} = 1.07$, $S.E. = 0.15$, $t = 6.96$, $p < 0.001$, 95% $CI = 0.77 - 1.38$) and for DZ was 0.45 (ESZDZ = 0.49, S.E. = 0.03, t = 19.89, p < 0.001, 95% CI = 0.44). Heritability of 68% was found, 11% of variance explained by shared environment effects, and 21% by non-shared environment effects and measurement error. For this Phenotype, moderators analysis showed a moderation of school grade levels on the magnitude of both MZ and DZ correlations with about 20% of the variation in reading comprehension performance during middle school being independent from reading comprehension performance in elementary school. A decrease of both heritability and shared environmental effects, and an increase in non-shared environmental effects was founds, from elementary school to middle school.

Regarding Phonological Awareness (PA), an overall MZ correlation of 0.75 was found ($E_{SZMZ} = 0.97$, $S.E. = 0.12$, $t = 8.04$, $p < 0.001$, 95% $CI = 0.73 - 1.22$) and a correlation of 0.49 for DZ ($ES_{ZDZ} = 0.54$, $S.E. = 0.03$, $t = 16.28$, $p < 0.001$, 95% $CI = 0.47$). The overall heritability for PA was 52%, the shared environmental effect

23%, and the non-shared environmental effect 25%. Moderators analysis showed that school grade levels moderated the magnitude of the correlations between grades, with about 30% of the variation in PA performance during elementary school independent from the performance in preschool/kindergarten. A decrease of both heritability and shared environmental effects, and an increase in non-shared environmental effects was found by Andreola et al. (2021) from preschool/kindergarten to elementary school.

For Rapid Automatic Naming the meta-analysis showed an overall MZ correlation of 0.61 ($ES_{ZMZ} = 0.71$, $S.E. = 0.06$, $t = 12.53$, $p < 0.001$, 95% $CI = 0.59$) and an overall DZ correlation of 0.38 ($ES_{ZDZ} = 0.40$, $S.E. = 0.08$, $t = 4.90$, $p < 0.001$, 95% $CI = 0.23$). The Falconer's formula (Falconer, 1960) was applied to meta-analytic MZ and DZ correlations to get the overall heritability, which was 46%, 15% for shared environment and 39% for non-shared environment.

For Spelling Abilities the overall MZ correlation was of 0.79 ($ES_{ZMZ} = 1.08$, $S.E. = 0.20$, $t = 5.56$, $p < 0.001$, 95% $CI = 0.69 - 1.48$) and of 0.39 for DZ ($ES_{ZDZ} = 0.42$, $S.E. = 0.03$, $t = 13.00$, $p < 0.001$, 95% $CI = 0.35$). The overall heritability was of 80%, with an effect of 0% for shared environment and of 20% for non-shared environment.

Last but not least, Language Skills overall MZ correlation was 0.81 ($ES_{ZMZ} = 1.13$, $S.E. = 0.14$, $t = 7.81$, $p < 0.001$, 95% $CI = 0.83 - 1.42$) and for DZ of 0.64 ($ES_{ZDZ} = 0.75$, $S.E. = 0.09$, $t = 8.59$, $p < 0.001$, 95% $CI = 0.57$), with overall heritability of 34%, an overall effect of 47% for shared environment and of 19% for non-shared environment.

The results of this meta-analysis indicate a similar pattern of causal architecture across most of the reading-related neurocognitive functions analyzed by these

studies, with moderate-to-substantial meta-heritability estimates, smaller environmental contributions, a significant effect of school grade levels, and no significant effects of moderators such as sex, and spoken language. This could in part be due to the fact that the genetic covariance among these neurocognitive components is high. Specifically, the phenotypic variance of general reading, letter-word knowledge, phonological decoding, reading comprehension, PA, RAN, and spelling, was primarily explained by additive genetic and non-shared environmental factors, while shared environment appeared to play a less important role.

These meta-analytic results agree with the well-established notion that the contribution of shared environmental influences is in general relatively small, and accounts for lower phenotypic variance compared to the non-shared environmental influences in reading skills. The results of the meta-analysis of the twin correlations result in a heritability above 50% for general reading, letter-word knowledge, phonological decoding, reading comprehension, PA, RAN, and spelling, suggesting that there is indeed a robust genetic effect on these reading-related skills. Furthermore, those findings support the notion that more research is now needed to specifically distill what unique environmental effects create individual differences in children growing up in the same family.

These findings were highly consistent with the previous studies that we took into consideration. Taken together, meta-analytic approach shows that genetic as well as non-shared environmental factors contribute to individual differences in reading-related neurocognitive components. Except for language for which shared environment seems to play a more important role, the causal architecture across most of the reading-related neurocognitive components can be represented, according to the author of the meta-analysis (Andreola et al., 2021), by the following equation $a^2 > e^2 > c^2$.

Moving to estimates of genetic influences for dyslexia we can see that they are high as well. In the same study from Dale et al. (2005) they analyzed the 7 years old sample falling in the lowest 13.4% of a reading composite score of word recognition and phonological decoding finding group heritability between 0.42 and 0.56; the number went up (between 0.50 and 0.67) for the same sample in another study (Harlaar et al., 2005) considering the lowest 10% on the same measures. Other two studies (Light & DeFries, 1995; Hawke et al., 2006) on an older group sample of dyslexic children with mean age of 12 (ranging from 8 to 20 years old) found a heritability for a reading composite that was comparable to the previous study in magnitude ($h^2 = 0.51 - 0.65$). But there was a study by Stevenson (1991) that reported much lower group heritability ($h^2 = 0.03$ for word recognition, $h^2 = 0.41$ for phonological decoding and $h^2 = 0.21$ for a reading composite).

The h^2 estimates of those different studies suggest that individual differences in reading are substantially influenced by genetic factors regardless of reading level (Light et al., 1998; Dale et al., 2005; Light & DeFries, 1995; Hawke et al., 2006). However, in the study by Light et al. (1995) the substantially higher heritability estimate for the reading ability within the reading-disabled sample compared with the control sample suggests that the etiology of reading disability may differ from that of individual differences in reading performance.

Hensler et al. (2010) pointed out some limitations of all those studies. The estimates of heritability are a sample estimate of the population parameters, therefore they are highly dependent on the sample and not always representative of the entire population. In those studies there is a prevalence of white families with an upper Socio-Economic Status (SES) and this could cause a bias in the estimation of environmental and heritability factors. This is why those results might not replicate in a different sample with different ancestry or SES.

Cognitive abilities and Dyslexia

Relation to academic achievement

In our research we will use as an estimator of Academic Acheivment (AC) the Queensland Core Skill Test. This test is known to have a strong genetic composition (Wainwright et al., 2005; Lazaroo et al., 2019). In a study conducted by Wainwright et al. (2005) genetic influences explained 72% of the variance in the total QCST score. In an article by (Lazaroo et al., 2019) the heritability was estimated for three subscales of the QCST, resulting in a heritability of .45 for Create and Present (CP), .70 for apply techniques and procedures (ATP) and .76 for comprehend and collect (CC). In the same study they also wanted to see if phonemic decoding (PD, a skill involved in dislexia) had an individual genetic effect that was different from the general genetic effect for cognitive abilities, Lazaroo et al. (2019) found that PD was moderately influenced by the general factor, but a larger amount of its genetic variance was explained by the phonemic decoding factor. Phonemic Decoding (PD) itself was found to be highly heritable, with a 0.52 and 0.68 for the reading disabled and control samples, respectively, in a Twin design study conducted by Light et al. (1998), suggesting that genetic influences significantly contribute to variation in the PD scores regardless of reading level.

A relationship that will be taken into account in our research is the one between reading ability and creativity. Past literature has found a correlation between the two as a compensatory enhancement for reading disability leading to increased creativity (Everatt, Steffert, & Smythe, 1999). What is found by Ritchie, Luciano, Hansell, Wright, and Bates (2013) is that the actual correlation in their sample is in the opposite direction, higher scores in reading ability were correlated with higher creativity scores. A number of studies supported the compensation hypothesis but the literature review carried out by Everatt et al. (1999) analyzing them showed

that the sample size was quite small and the concept of the reading disorder in itself varied a lot. Even though the review was in favor of a link between reading disorder and creativity, and Dyslexia adults showed a small advantage over non diagnosed adults in measures of self report and laboratory based tasks, the sample size was not wide enough to be considered reliable (Ritchie et al., 2013).

Validity in research is particularly important, in the creativity area, referring to laboratory tasks and self reports, this validity is only a modest predictive criterion (Dietrich, 2007; Dietrich & Kanso, 2010), indicating the need for new measures of creativity. Two such methods have gained particular prominence: Measuring real-world creative output, such as occupational type, and using comprehensive personality inventories. Ritchie et al. (2013) gives an overview of the studies in support of those two ways of measuring creativity identifying Openness and creative writing as good indicators. They also pointed out that the latter might have an associations resulting from active course-selection effects by students based on their skills in reading. Ritchie et al. (2013) then studied the relation of creativity and Dyslexia in a large sample of Twins from the Brisbane Adolescence Twin Study using Creativity was assessed using two versions of the Openness to Experience scale, the short-form 12-item NEO-FFI and the long-form 48-item NEO-PI-R (Wainwright, Wright, Luciano, Geffen, & Martin, 2008) and Create and Present (CP) subscale from the QCST to assess creative writing (Wainwright et al., 2005). The results of their research pointed out a positive association between reading ability and Openness and CP (from .17 to .28). Contrary to the compensation hypothesis, poor reading, spelling, and nonword repetition ability was associated with lower, rather than higher scores on all measures of creativity, even when controlling for IQ. Furthermore, the compensation model did not hold in the extremes; the poorest readers, beyond a threshold of one standard deviation below the mean, tended to have the lowest scores on the creativity measures, and the most creative individuals,

beyond one standard deviation above the mean, tended to have the highest reading ability. The main conclusion of Ritchie et al. (2013) is that reading disability may not involve innate compensatory enhancements, but instead may lead some individuals to learn compensatory skills.

In our study we will use our Polygenic Score for Dyslexia, collected on the same sample of Twins, as a predictor of creativity, calculated with Create and Present (CP) to see if a genetic predictor will have results similar to those of Ritchie et al. (2013) where a positive association of reading ability and creativity was found.

Dyslexia and processing speed

As reported by de Oliveira, da Silva, Dias, Seabra, and Macedo(2014) Processing speed is a component of the cognitive model of reading, together with word recognition and linguistic comprehension. This model can give some insight about the specific deficit present in dyslexic children. Dyslexic subjects have problems in processing visual and auditory information quickly, as well as difficulties in processing speed and rapid automatized naming (RAN) (Wolf & Denckla, 2005). In the double deficit hypothesis, Wolf and Bowers(1999) describe processing speed and phonological awareness as two important skills in the development of reading ability. Therefore, in dyslexic subjects there could be a deficit in one or both those skills. de Oliveira et al.(2014) studied a sample where the dyslexic group performed worse in word recognition and, as predicted, the dyslexic group required more time during the naming test, corroborating the literature on word recognition and processing speed deficits in dyslexia.

The tests that we will take into account are Inspection Time (IT) and Choice Reaction Time (CRT) and Digit Symbol; all are a measure of Processing Speed. In our study as a measure of IT a line discrimination task was administered (Luciano,

Smith, et al., 2001), for CRT the participants were presented with a visual form of dripping taps and the participant had to press the appropriate computer key as quick as possible to stop the tap from dropping (Luciano, Wright, et al., 2001). In the next Chapter we will use those two measures of Processing Speed to see if the Polygenic Score can predict lower speed. Now we will focus on the literature regarding their heritability and relation to other cognitive abilities.

Using a Twin Sample, Luciano, Smith, et al. (2001) analyzed the relation of IT and IQ. A common genetic factor mediated the relationship among standard deviation of the IT curve (SDIT) (36%), Performance IQ (PIQ) (30%), and Verbal IQ (VIQ) (18%). Twin correlations for IT demonstrated significant familial aggregation. MZ co-twins showed greater similarity (r=.34) than DZs (r=.21), but this correlation was short of the test-retest correlation indicating that stable features of the unique environment are also important. A significant correlation between IT and Full-scale IQ (FIQ) (- .35) was found in the study, which was higher for PIQ (- .33) than VIQ (- .26) when these were analysed separately. Bivariate model fitting indicated that the covariance between IT and FIQ was due to a common genetic factor, which explained 36% of the variance in IT and 32% of the variance in FIQ. In the trivariate analysis, this factor explained significantly more variance in PIQ (30%) than in VIQ (18%). There was no significant common environmental factor. The genetic correlations between IT and each of the IQ measures were higher than the phenotypic correlation (especially for PIQ); indicating that variation in genes that produce faster ITs are strongly related to the variation in genes that promote higher IQs. Another study by Luciano et al. (2004) presented the covariation between IT and CRT and Intelligence. The general genetic factor displayed factor loadings ranging between 0.35 and 0.66 for the IQ sub-tests, with IT and CRT loadings of -0.47 and -0.24, respectively. A unitary factor was not sufficient to describe the relationship between the variables, an independent genetic effect is likely to have an

influence on Processing Speed.

Relation to IQ and IQ sub-tests

Intelligence has a strong genetic component (Rack & Olson, 1993; Light et al., 1998; Luciano, Smith, et al., 2001; Luciano et al., 2004; Lazaroo et al., 2019; Wainwright et al., 2008). In a study by Lazaroo et al. (2019) the heritability of the two subscores of IQ were respectively 0.63 (0.58-0.68) for VIQ and 0.71 (0.64-0.77) for PIQ; the genetic correlation between non-word reading tests and VIQ was .54 [0.48-0.60] and .31 [0.26-0.36] for PIQ, as expected the score in the reading measure is more correlated with the VIQ (since it needs verbal fluency to complete the tasks) and lower for PIQ where the reading ability is less determinant of the result.

In the differential diagnosis of dyslexia it is important that the reading disability is not related to low IQ. Both IQ and Developmental Dyslexia have a strong genetic component. The aim in this research is to investigate if a Polygenic Score does not predict lower scores on IQ.

A study conducted by Rack and Olson (1993) investigated the relation of IQ and Developmental Dyslexia. Subjects with high full scale IQ had a more heritable word recognition deficit ($h_g^2 = .67$), compared to subjects with a low full scale IQ who had a heritability of .40. We have seen before that there is an ongoing debate on weather IQ should be taken into account for the diagnosis of dyslexia, Rack and Olson (1993) argues that reading problems may vary as a function of intelligence and therefore it is premature to abandon its use in testing for reading disability.

Genes of Dyslexia

Even though the hypothesis that Developmental Dyslexia is a heritable trait has been established over the last decades, molecular genetic studies have always struggled to identify genes with a large associations with reading ability/difficulties (Luciano, 2017), with a substantial problem in replicability. When studying the genes involved in specific traits, studies have taken two basic forms: whole-genome linkage and whole-genome association studies. In genetic linkage studies samples of family groups are used in order to map those regions of the genome which contain genetic variants that co-occur with a trait, in our case reading disability, it is called co-segregation. Those studies are usually well powered to detect the chromosomal regions in which a single genetic loci or cluster of genetic loci is able to explain at least 10% of the total variance. In a review by Scerri and Schulte-Körne (2010) of eight independent genome-wide studies for dyslexia involving linkage analysis all lacked the ability to find evidence of genes having an effect this large.

The second case of studies involved in discovering genes associated with a trait are the Genome-wide association studies (GWAS). GWAS test hundreds of thousands of genetic variants across many genomes with the goal to find the ones statistically associated with the trait of interest (Uffelmann et al., 2021). This methodology has generated in many different fields a myriad of robust associations for a range of traits and diseases, and regardless of the trait of interest the number of associated variants is expected to grow steadily as GWAS sample sizes increase. Our interest is on those studies which used GWAS to test for linear association between the genetic variants of thousands of individuals and quantitative reading scores or in those comparing variant frequencies between reading-impaired cases and controls in unrelated individuals.

GWAS seek to find genes with small variants called Single nucleotide polymor-

phisms (SNPs), which are the most common kind of genetic variation in the human population (Court, 2007). A SNP represents a difference in the transcription of a nucleotide, for example a SNP may replace cytosine (C) with thymine (T) in a certain block of DNA. SNPs are normal to occur throughout the entire DNA of a person, with an average of 1 every 1,000 nucleotides, resulting in about 4 to 5 million SNPs in total in an individual's genome. A variation, to be classified as a SNP has to be found at least in 1% of the population. So far scientists have found more that 600 million SNPs in different populations all around the world (Rocha et al., 2006). The most important function in research of SNPs is to act as biological markers to help scientists locate specific genes associated with complex disease, such as diabetes, cancer, heart disease, or in our case reading disability.

Until recent time GWAS have failed to find any large associations (Luciano et al., 2013), genetic variants uncovered by genome-wide association studies of reading ability/disability had not accounted for more than 0.4% of variance (Gialluisi et al., 2014; Luciano et al., 2013; Meaburn, Harlaar, Craig, Schalkwyk, & Plomin, 2008). The most robust candidate genes identified so far include *DYX1C1* (15q21), *DCDC2* and *KIAA0319* (6p22.3), *GCFC2* and *MRPL19* (2p12), and *ROBO1* (3p12.3-p12.3) (Gialluisi et al., 2021). But only in 2019 GWAS studies reached significant statistical levels. Truong et al.(2019), using Rapid Automatic Naming (RAN) and Rapid Alternating Stimulus (RAS) as predictors of reading disability, found genome-wide significant effects at rs1555839 ($p = 4.03 \times 10^{-8}$) which was also replicated in an independent sample of European ancestry and support for a novel trait locus at chromosome 10q23.31. Then again Gialluisi et al.(2019) observed a genome-wide significant effect on RAN for letters for variants on 18q12.2, within MIR924HG (micro-RNA 924 host gene; rs17663182 $p = 4.73 \times 10^{-9}$), and a suggestive association on 8q12.3 within NKAIN3 (encoding a cation transporter; rs16928927, $p = 2.25 \times 10^{-8}$). In the same study rs17663182 (18q12.2) also showed genome-wide significant

multivariate associations with RAN measures ($p = 1.15 \times 10^{-8}$).

Gialluisi et al.(2021) carried out another genome-wide association study (GWAS) on 2274 dyslexia cases and 6272 controls, partially overlapping with his previous work (Gialluisi et al., 2019), testing associations at the single variant, gene, and pathway level, and estimating heritability using single-nucleotide polymorphism (SNP) data. In this research, no SNP reached statistical genome-wide significance for Developmental Dyslexia. The most suggestive association was at rs6035856 (G/T, MAF = 0.45; p = 9.9 × 10-8), an intronic variant located within the gene LOC388780 (chr20p13); then the major allele G ($p = 3.2 \times 10^{-7}$) showed high quality imputation statistics across datasets for the risk of Developmental Dyslexia.

The most recent GWAS is the one carried out by Doust et al. (2022). This is the study that created the polygenic risk score that we will use in our research project. This study is the first one representing a 20 fold increase in sample-size from prior work, with a sample of 51,800 adults with a self-reporting diagnosis of Dyslexia and 1,807,070 controls. All participants had turned 18 (with mean age of 49.9 for the dyslexic group and 51.8 for the control). The PRS based on the 23ndMe GWAS was computed in four independent cohort explaining up to 6% of variance in reading related tests, 3.6% of variance explained in non-word-reading and 5.6% in word recognition tests. Overall, Dyslexia Polygenic Scores from this study were correlated with lower achievement on both reading and spelling tests and on a measure of phonological decoding which is typically impaired in dyslexic individuals.

Doust et al. (2022) identified 42 significant indipendent genome-wide loci ($p < 5 \times 10^{-8}$) and 64 loci with suggestive significance ($p < 1 \times 10{-}6$): 17 of those are genes linked to cognitive ability/educational attainment; other 25 are novel genes

and may be more specifically associated with dyslexia. 23 loci, of which 12 were novel, were validated in independent cohorts of Chinese and European ancestry. A similar genetic aetiology of dyslexia between sexes was confirmed in the study, and a genetic covariance between traits was found, which included ambidexterity but was not identified for neuroanatomical measures of language-related circuitry. This study reported no significance of the 75 previously reported dyslexia associations in the present genome-wide analysis.

Use of Polygenic Risk Scores in research

It is now important to give to the reader an overview of what are Polygenic Risk Scores (PRS) and their application in the research field, with a particular look at Learning Disabilities.

Since a multitude of human traits are heritable to varying degrees, PRS are extremely useful in research because they allow to study the influence of genes for those traits which are not caused by one gene or a small set of genes with a large effect, those traits are called Mendelian traits, but are likely caused by a weak to moderate contribution of many genes (Lambert, Abraham, & Inouye, 2019). PRS, sometimes called genomic risk scores, allow to predict the predisposition of an individual to develop a disease based on genomic information. In its simplest and most common form, PRS are sums of the effects of m SNPs, based on the estimated SNP effect sizes $\hat{\beta}$ (obtained from GWAS summary statistics).

Even though PRS are extrapolated from GWAS results, those two are based on different principles. In the case of GWAS, they are designed to detect SNPs associated with a disease with a low false positive rate. PRS have the advantage to allow for methods with a more flexibility in signal to noise trade-off. Therefore, it

is possible to obtain a powerful PRS by incorporating a larger number of SNPs by having a more noisy effect size estimate (Lambert et al., 2019). To decide how much noise should be allowed in a PRS there is no universal set of parameters; the choice should be made based on the genetic architecture of the disease, sample size and phenotype density. In research, usually more trials with different amount of noise are carried out to identify the PRS that predicts better the trait of interest. It is important to use an independent validation dataset or a cross-validation to obtain an unbiased estimation of the predictive performance, otherwise overfitting is very likely to happen. In the second step, the predictiveness of the PRS has to be tested on a different cohort from the one used for the GWAS or for tuning the PRS.

The accuracy of a PRS is highly dependent on the heritability of the trait of interest. When explaining the results of a PRS the effect size per standard deviation (SD) is usually reported, toghether with the proportion of variance explained (R^2) by the model (Lambert et al., 2019). Instead, when the trait has only binary outcome (is either present or it is not), the effect size is expressed as odd ratio (OR). Moreover, the model's performance is usually measured using two kinds of variance explained: Nagelkerke's or pseudo-R2; otherwise classification accuracy using area under the receiver-operating characteristic curve (AUC), the area under the precision-recall curve or Harrell's C-index can be used as well (Steyerberg et al., 2010). However, with the last indices caution must be exercised when interpreting prediction; metrics such as AUC or C-index without sufficient context can lead to severe bias; even small increases in these metrics can lead to several percent of the population being reclassified into different risk categories. In addition, Lambert et al.(2019) made clear that when metrics are compared across studies, factors such as ancestry as well as study design, for example the covariates included in the risk model, can affect these measures and even the SD of the PRS.

In the past PRSs have been drawn for different dyslexia related measured. A PRS for educational attainment was drawn out from a very large GWAS of Lee et al. (2018), accounting for up to 5.1% of the variance in reading measures in a sample from the EduYears study conducted by (Selzam et al., 2017) with the subjects being 14 years old. This result has been replicated in a more recent study by Price et al. (2020). Lower variance explained was found by Gialluisi et al. (2019) in the NeuroDys sample, explaining less than 2% of the total variance in reading abilities, including spelling, non-word reading, phonological awareness, and digit span. The PRSs have also shown association between other cognitive abilities and psychiatric disorders and dyslexia (Price et al., 2020; Gialluisi et al., 2019, 2021). In particular, the more recent study by Gialluisi et al. (2021) found that the PRS for bipolar disorder had a stronger association than the PRS for ADHD ($p_{Bipolar} = 1.33 \times 10^{-43}$; $p_{ADHD} = 7.66 \times 10^{-13}$) and explained a larger proportion (2.8% vs. 0.7%) of the variance in dyslexia risk. Also the PRS for Schizophrenia showed a substantial association with dyslexia ($p = 3.66 \times 10^{-22}$) and explained up to 1.4% of the variance. The GWAS by Doust et al. (2022) is the one to have found the most powerful PRS for Dyslexia so far being able to explain up to 6% of the variance in reading abilities.

Chapter 3

Aims of the project and hypothesis

This project aims to investigate how a polygenic risk score for dyslexia developed by Doust et al. (2022) predicts a variety of cognitive abilities. As we have seen in the previous paragraph, both dyslexia and cognitive abilities such as IQ, IT and CRT are highly heritable. Since those abilities have been studied in relation to reading ability as well, we want to see how our Polygenic Score for Dyslexia predicts them. More in detail we developed three main hypothesis:

- In the first hypothesis the aim is to see if the Polygenic Score would predict a lower academic achievement. The academic achievement is evaluated through the Queensland Core Skills Test (QCST). More specifically, the Polygenic Score might also be associated with creative writing. As the literature regarding dyslexia and creativity is addressed, some studies found that children with dyslexia had higher scores of creativity but this finding was not replicated in the study of Ritchie et al. (2013). This research wants to test whether Dyslexia Polygenic Scores predict creativity in adolescents, with the hypothesis being that it should predict lower, not higher, levels of creativity for the Create and Present task from the QCST.

- The Polygenic Score should not predict a lower nonverbal IQ, given that the diagnosis for dyslexia is based on reading ability discrepancy with IQ (i.e. IQ is unaffected). But the tool used to assess IQ (Multidimensional Aptitude Battery; 5 sub-tests) requires the participants to read the verbal sub-test items, therefore, for the second hypothesis expect to find that the IQ is negatively affected for those tasks dependent on reading. It is therefore expected that the high Polygenic Score for Dyslexia should predict lower cognitive performance on these tests.

- For the third hypothesis the interest is on information processing speed and how the Dyslexia Polygenic Score can predict higher scores on tests measuring the construct (Inspection Time, Choice Reaction Time, WAIS Digit Symbol) because lower scores symbolize better performance. The hypothesis is that information processing speed is affected in subjects with a high Polygenic Score. It is expected that higher Dyslexia Polygenic Score will associate with slower performance on Information Processing Speed tasks.

Description of the sample

The data available for this study was a database of 2469 subjects but due to the complex structure of the data and the great amount of tests collected, not all data was available for all subjects and many could not be used in the analysis. The final analysis is carried out on a sample of a maximum of 701 twins from the Brisbane Adolescent Twin Study (M. Wright et al., 2001; M. J. Wright & Martin, 2004), a study started in the '90's in Australia with the aim to collect as much data as possible on many different cognitive abilities, psychological traits and other health parameters.

	MZ	DZ	age	sex
VIQ	369	855	16.72 ($SD = 1.65$)	769 F 637 M
PIQ	369	854	16.71 ($SD = 1.65$)	768 F 636 M
DIGIT	364	825	16.72 ($SD = 1.67$)	751 F 620 M
CC	289	671	17.18 ($SD = 0.41$)	645 F 465 M
CP	289	671	17.18 ($SD = 0.41$)	645 F 465 M
IT	251	598	16.36 ($SD = 0.68$)	539 F 428 M
CRT	272	638	16.4 ($SD = 0.7$)	574 F 466 M

Table 1: Description of the sample for those participants who had Dyslexia Polygenic Score. Abbreviations: Verbal IQ (VIQ), Performance IQ (PIQ), Digit Symbol (DIGIT), Comprehend and Collect (CC), Create and Present (CP), Inspection Time (IT), Choice Reaction Time (CRT)

Since the great amount of data collected on many different batteries of tests, the samples are based on different subgroups of participants. The description of each group is reported in the table below (1) for the sub-tests of interest for our project. For VIQ the sample is of 369 MZ and 855 DZ twins, for PIQ there are 369 MZ and 854 DZ twins, data was missing on one DZ twin, for Digit Symbol 364 MZ and 825 DZ twins. For those subtests, all part of the MAB and WAIS-R, subjects were tested at a mean age of 16.7 ($SD = 1.65 - 1.67$). For the subtests of the QCST we had 289 MZ and 671 DZ twins tested at a mean age of 17.18 ($SD = 0.41$). For the Inspection Time (IT) and Choice Reaction Time (CRT) there were 251 MZ and 598 DZ and 272 MZ and 638 DZ respectively with mean age of testing of 16.4 ($SD = 0.7$).

Zygosity was determined objectively in same-sex twin pairs with an extremely high probability of correct assignment ($> 99.9\%$) (M. J. Wright & Martin, 2004), using the AmpFISTR Profiler Plus Amplification kit, ABI, which examines 10 independent DNA markers; nine short tandem repeat (STR) loci and one homologous region on the X and Y chromosomes (M. J. Wright & Martin, 2004; Wainwright et al., 2005). For all participants under the age of 18, informed consent was signed from both participants and their parents before starting any testing. In the initial

phase of the data collection, back in early 2000, some participants where excluded if parents indicated in the report either a history of significant head injury, psychiatric or neurological illness, use of medication with known effects on the central nervous system or substance abuse or dependence (M. J. Wright & Martin, 2004). Moreover, all participants were admitted only if they had normal or corrected-to-normal vision. Ethical approval was given to the Brisbane Longitudinal Twin studies the QIMR Berghofer Human Research Ethics Committee (Lazaroo et al., 2019).

Measures and Procedures

We had many tests in our database and many interesting aspects could have been investigated, but for the purpose of this project we decided to focus on a more stringent number of cognitive tests. In support of our decision, parallel factor analysis on the cognitive tests showed a number of 6 factors and 3 components. A graphical representation of those subgroups is given in the Figure 1, showing a correlation matrix with 3 sub-groups of tests highly correlated and 2 tests not correlated to the others. The choice was based on the parallel factor analysis and the correlation matrix. To investigate the first hypothesis on Academic Achievement, Comprehend and Collect (CC) was chosen for being the most correlated with the other sub-tests of the QCST and Comprehend and Collect was chosen to investigate Creativity. For the second hypothesis both Verbal IQ and Performance IQ were chosen, they both correlated with the MAB sub-tests and tested well the hypothesis of a lower Verbal IQ due to the nature of the test involving reading ability. For Choice Reaction Time (CRT), 3 different tasks were available, CRT_8 was chosen for being the most correlated with the other two tasks. Lastly Inspection Time (IT), and Digit Symbol were chosen for not being related to other Cognitive Tests and being part, together with CRT, of the third hypothesis stating that the POlygenic Score should predict slower Information Processing Speed.

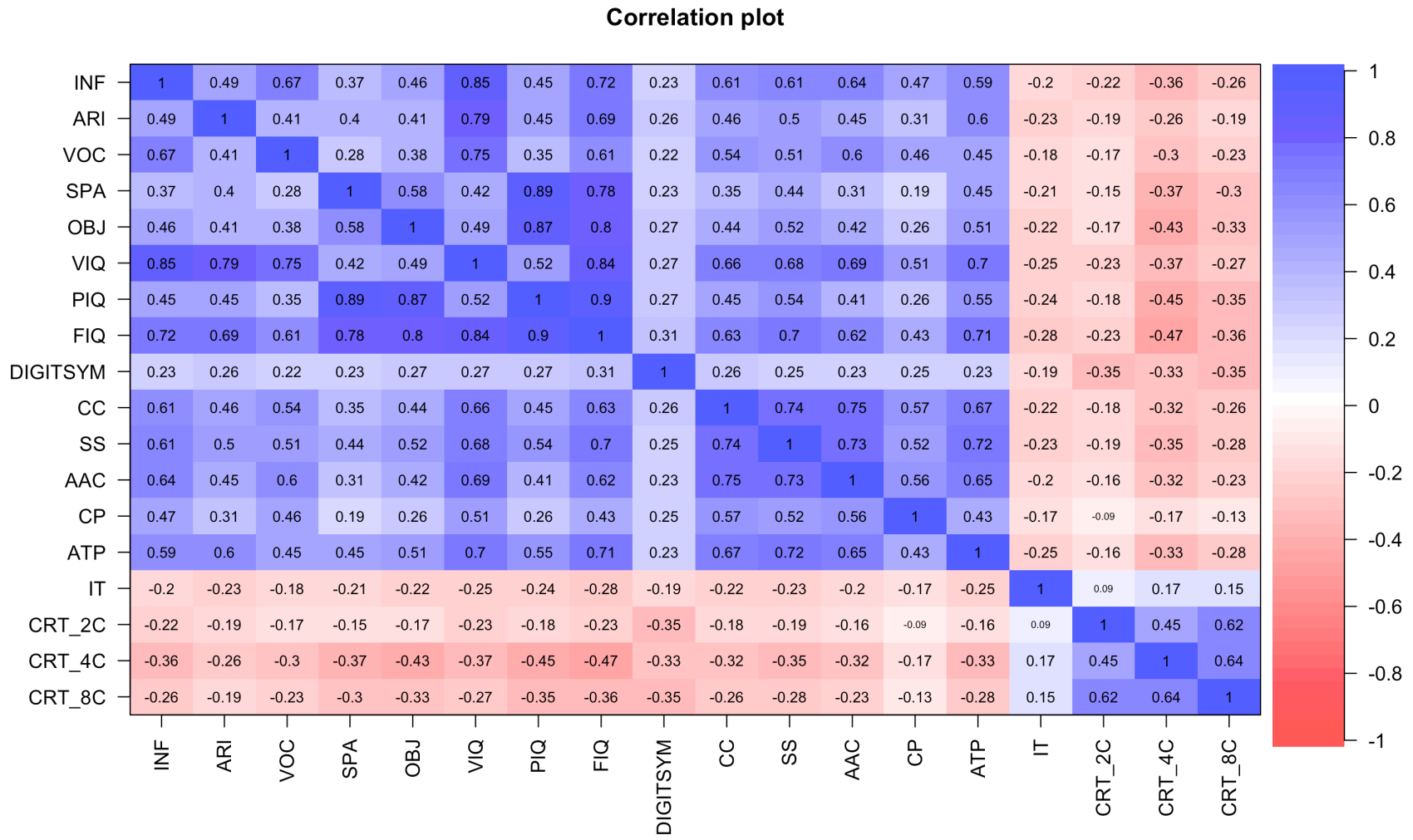

Figure 1: Correlation matrix pairwise of all 18 variables (n. range min: 919, max: 1,641). Blue colors show a positive correlation and red colors show a negative correlation. Visually 3 sub-groups can be individuated and 2 single sub-tests stand out fro not being related to the other cognitive tests.

Multidimensional Aptitude Battery (MAB) and Digit Symbol

The MAB is a general intelligence test with multiple choice answers based on the revised Wechsler Adult Intelligence Scale (WAIS-R) (Wechsler, 1981; Jackson, 1984). The test aims to assess cognitive abilities in adults and adolescents aged ≥ 16. The sub-tests administered were five and they can be grouped in two subgroups, those assessing Verbal IQ (VIQ): Vocabulary (VOC), Information (INF), Arithmetic (ARI); and those assessing Performance IQ (PIQ): Spatial (SPA), Object Assembly (OBJ). (Jackson, 1984) found a correlation between the MAB and WAIS-R of 0.94 for the verbal sub-scales and 0.79 for the performance sub-scales. Moreover, a previous study conducted on a sub-sample of twins used in our research reported test-retest reliability of 0.89 and 0.97 for VIQ and PIQ respectively (Luciano, Wright, et al., 2001). From the WAIS-R the Digit Symbol sub-test will be used. For the purpose of this project the two subgroups VIQ and PIQ will be used as estimators of the Verbal component and the Performance component of Intelligence and Digit Symbol as a test investigating processing speed.

Queensland Core Skills Test (QCST)

The QCST is a standardised test which must be sat by all those students aiming to enter tertiary education in their final year of high-school in Queensland (Lazaroo et al., 2019). The test assesses students' academic achievement by investigating their higher-order scholastic skills that are thought in a wide range of courses but it is not subject specific, in this way it avoids penalising students based on the curriculum they chose (Matters & Gray, 1995). Nevertheless the QCST can be considered a valid estimator of academic achievement with historical data from 1992 to 2007 showing a correlation of the QCST of $0.70 - 0.75$ with a within-school ranking of students and with their school grades (Authority, 2008). The QCST is composed of four exam papers; a written task, two multiple choice papers and a collection

of short response questions. Students completed these papers over two consecutive days during the third term of their final year. The content of the test falls within five predetermined academic factors:

The Comprehend and Collect factor assesses the students' ability to comprehend, manipulate and interpret data from a wide range of stimuli, including written, mathematical and visual stimuli. SS assesses the ability to select, sort and organise information and detect complex patterns and relationships. Analyse, Assess and Conclude evaluates induction and deduction of relationships, evaluating the merit of text and drawing conclusions from text. Create and Present evaluates a students' ability to structure and use written language to communicate ideas effectively. Lastly, Apply Techniques and Procedures requires to make calculations and evaluates the ability to solve mathematical problems.

After the permission was given from both participants and their guardians, the database with the scores was forwarded from the QCAA (Lazaroo et al., 2019). Annual statistics where provided for means, standard deviations and score ranges. This was necessary to allow to standardise the scores using the mean and standard deviation since the maximum score attainable for each factor varied every year.

For this research the focus will be on Comprehend and Collect (CC), which was the most correlated with the other sub-tests in the correlation table (1) to not have redundant measures, and Create and Present (CP) to test our hypothesis on creativity.

Inspection Time (IT)

IT, a component of Processing Speed, was tested by a line discrimination task, which was presented as a pseudo-computer game where the participant had to choose

the longer of two worms to go fishing (Luciano, Smith, et al., 2001). Participants were told that the lines represented the worms burrowing into the ground and that their task was to identify the longer worm to catch the most fish by pressing the corresponding left or right arrow key on the computer's keyboard. For every five correct judgements a fish would appear at the lower left side of the screen. The aim was not to achieve a low reaction time but accuracy in giving the right command and it was made clear to all participants before initiating the task.

Choice Reaction Time (CRT)

CRT was evaluated in the study by (Luciano, Wright, et al., 2001) with a task that was presented to the participants in the visual form of dripping taps. All participants were instructed to quickly press the appropriate computer key in order to stop a tap from dripping. They had to align and rest their fingers on the keys Z, X, C, and V (left hand, index finger on V) and the keys M, comma (,), period (.), and slash (/) (right hand, index finger on M) of a standard QWERTY keyboard. Different coloured taps corresponded to the same fingers on both hands to aid tap and finger alliance, for example, the taps matching the index fingers were both red.

To familiarise the participants with the response keys, they were initially presented with the eight taps and were required to respond to each tap in a left-to-right sequence. Two-, four-, and eight-choice conditions had a number of trials of 96, 48, and 96, respectively. In all conditions, eight taps appeared on the monitor; those taps in use for the two- and four-choice conditions were had brightening their colours to make them salient. The output measure was the mean RT (in ms) of correct responses for each choice condition. Being the scores on two-, four-, and eight-choice condition highly correlated, a decision was made to use only eight-choice in our analysis.

Description of the variables' scores distributions

In the figure below (2) there are the graphic representations of the scores for each test and sub-test and for the Polygenic Risk Score for Dyslexia. More in detail, it can be seen how the Dyslexia Polygenic Risk Score has a normal distribution with a mean of -0.04 and $SD = 0.97$, the distribution of the scores has a Skewness of 0.06, indicating that it is centred close to 0, and Kurtosis of 2.99, indicating sharpness of the peak in the distribution. Moreover, Verbal IQ has as well a normal distribution with 110.27 and $SD = 11.23$, Skewness of 0.18 and Kurtosis of 2.76. Performance IQ has a mean of 112.74 and $SD = 15.8$ with Skewness of -0.21 and Kurtosis of 2.52. For Digit Symbol the mean score is of 59.38 and $SD = 10.75$, Skewness of 0.08 and Kurtosis of 2.9. Then, Create and Present and Comprehend and Collect have a mean of 0.24 and 0.21, $SD = 0.94$ and $SD = 0.92$ respectively, Skewness of -0.21 and Kurtosis of 2.66 for CC and Skewness of -0.03 and Kurtosis 3.46 for CP. Inspection Time is the test with the more Skewed distribution, with mean of 91.58 and $SD = 0.07$, Skewness is of 3.06 and Kurtosis of 13.78. The last task analysed in our project, Choice Reaction Time, shows a mean of 2.76 and $SD = 0.07$, with Skewness of 0.02 and Kurtosis of 3.91.

Data Analysis

The model used to analyse the data is a multivariate multilevel model (the formula of the model can be found in 2). The predictive part of the model is always the same: a fixed effect (the Polygenic Score) and a random effect (Family ID, which takes into account the zygosity of the twins); the dependent variables are the cognitive tests. All the tests scores have been adjusted before running the model for sex, age and ancestry to have less noisy measures using a linear regression model and keeping the residual part for the next analysis.

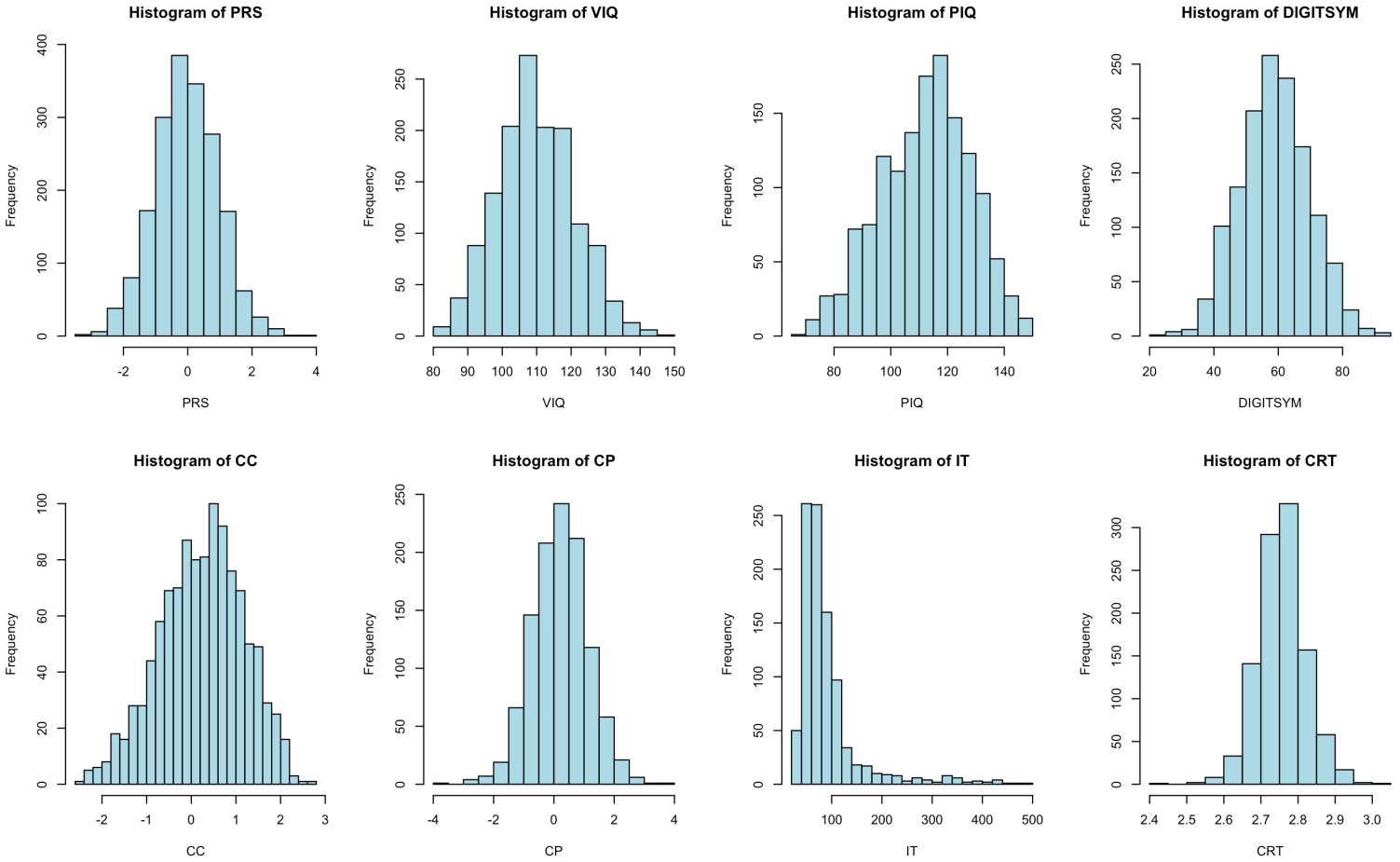

Figure 2: Histograms of the distribution of the scores for the Polygenic Risk(n=1,877), VIQ(n=1,406), PIQ(n=1,404), Digit Symbol(n=1,371), Comprehend and Collect(n=1,110), Create and Present(n=1,110), Inspection Time (n=967) and Choice Reaction Time (n=1,040)

Due to the multivariate nature of the model, the final number of participants on which the analysis was based is of 701 because only the subjects which had measures on all test scores were included. Despite the loss in sample size, a multivariate approach allows to have more reliable estimates of the parameters. The model was fitted using 'brms' R-package (Bürkner, 2018). The parameters of the model were estimated with the Markov chain Monte Carlo (MCMC) method, which samples data based on the distribution of the collected data and when specified implements the priors defined in the model, using Stan as a compiler (Carpenter et al., 2017). Posterior distribution were estimated for each parameter with 4 MCMC chains, each of them running at least 4,000 replications. In the section specific to each dependent variable, the posterior prediction will be graphically visualised with Posterior Predictive Checks (with the 'pp_check' function from the package 'brms'); the graphic visualisation of the data allows us to see how much the curve of the collected data overlaps with the posterior draws which are randomly extracted based on the parameters of the model, the closer the dark-blue line is to the light-blue lines, the better the model predicts new data (see Figure 3) and if the model is a good fit, it should be able to generate a dataset that resembles the observed data. The graphical representation of Conditional Effects (through the function 'conditional_effects' with the 'brms' package) will be given for each equation of the model, representing the effects of the Dyslexia Polygenic Score on the Cognitive Test of interest, with the grey band showing the level of uncertainty of where the blue line actually lies (see Figure 3, right panel).

Initially, a model comparison between our target model (M1) and the null model (Table 2), i.e. a model with only random effect (M0), is made to compare the predictive capability of the two models, with the goal of identifying the model that best predict scores in the Cognitive Tests. The main interest of this model comparison is to see if the Fixed Effect (the Dyslexia Polygenic Score) gives a significant contribu-

M0	M1
VIQ ~ (1 | FID)	VIQ ~ PRS + (1 | FID)
PIQ ~ (1 | FID)	PIQ ~ PRS + (1 | FID)
DIGITSYM ~ (1 | FID)	DIGITSYM ~ PRS + (1 | FID)
CC ~ (1 | FID)	CC ~ PRS + (1 | FID)
CP ~ (1 | FID)	CP ~ PRS + (1 | FID)
IT ~ (1 | FID)	IT ~ PRS + (1 | FID)
CRT ~ (1 | FID)	CRT ~ PRS + (1 | FID)

Table 2: The table shows the null model (M0) which has only the Random Effect, and the model containing the Dyslexia Polygenic Score (M1). Each model contains the same 7 dependent variables which are the Cognitive Tests that were selected based on the Parallel Factor Analysis and the graphical representation of the Correlation Matrix

tion in the amount of variance explained. The result of a model comparison using model weights is always a sum of 1 between the weights of the models, and if those two values are put in relation, the result is the probability that the first model will better predict new data compared to the second model. The best model will initially be identified based on the result of the Model Weights, using Leave-One-Out cross-validation (LOO-CV) (Vehtari, Gelman, & Gabry, 2017) with the stacking method which combines all models by maximising the leave-one-out predictive density of the combination distribution (Yao, Vehtari, Simpson, & Gelman, 2018). Then, for each dependent variable the Bayesian R^2 of the two models will be compared (Gelman, Goodrich, Gabry, & Vehtari, 2019) in order to find how much of the variance is explained by the Fixed Effect. Credible Intervals (CI) are provided for R^2 and Estimates to give a measure of uncertainty around effect estimates (Hespanhol, Vallio, Costa, & Saragiotto, 2019). CIs can be interpret similarly to Confidence Intervals but they should not be dichotomised in statistically significant or non-statistically significant. The correct interpretation is that there is a 95% probability that the true (unknown) estimate would lie within the interval, given the evidence provided by the observed data (Hespanhol et al., 2019).

In the next chapter the results of the model comparison will be presented. The

best model will be analysed with the support of the statistics explained in the last paragraph. Graphical representations of the posterior predictive distribution and the conditional effects will be explained in each section of the model.

Chapter 4

Results

In the previous Chapter Hypothesis have been expressed in great detail and an overview of the Cognitive Tests involved in the study has been given with a description of the sample for each one of them. Moreover, the scores' distributions for each Cognitive Test have been graphically represented with histograms and their characteristics have been presented. The Data Analysis has been explained and the statistics that will be used have been presented. Now the target model (M1) will be compared to a null model (M0) first and then it will been analysed in great detail providing model parameters and R2, with Credible Intervals, Posterior Predictive Checks and Conditional Effects.

Model Comparison

Initially the model was compared using Model Weights to a null model which had only the Random Effect. Model weights are used as an estimate of the probability that the model will make the best predictions on new data, conditional on the set of models considered (Burnham & Anderson, 2002). Within the package 'brms', the function 'loo_model_weights' was used compare the two models described in Table

2 in an attempt to identify the better model. The model (M1) containing both a random and a fixed effect had a weight of 0.56 and the null model (M0) with only the random effect had a weight of 0.44. If the two weights are put in relation the number obtained is an estimate of how well the first model predicts new data compared to the second model. Therefore, the model with the Polygenic Score is 1.28 times more plausible of M0. The result shows a very low relative evidence, meaning that the model does not predict well the data. From the results of the model weights, keeping in consideration the low relative evidence of M1, the decision to use M1 was made. In the next section each dependent variable will be analysed individually and a comparison with the null model will be carried out regarding the R^2.

VIQ

For Verbal IQ – VIQ ~ PRS + (1 | FID) – the Estimate had a value of -0.13 ($95\%CI = [-0.2, -0.06]$), since the Credible Interval does not include the 0 it can be stated that most likely the true value of the estimate is of negative sign. The posterior probability of the Estimate being negative is very close to 1. Moreover, looking at the effect size of the model, R^2 is used to estimate the variance explained by the model that has been chosen and will be compared to the one of the null model. The regression with the Polygenic Score for Dyslexia and Family ID had an R^2 of $0.118(95\%CI = [0.06, 0.18])$, whereas the model for VIQ from M0, with only the random effect, had an R^2 of 0.10 ($95\%CI = [0.050.16]$), from those two values it can easily be deducted how the actual variance explained by the model is mostly relative to the random effect, rather than to the Dyslexia Polygenic Score. As it was first hypothesized the Verbal IQ was negatively affected by the Dyslexia Polygenic Score for Dyslexia, indicating slightly lower scores on VIQ for those subjects with a high Polygenic Score. Another aspect investigated in the project is how the model predicts the distribution of new simulated data based on the model's parameters

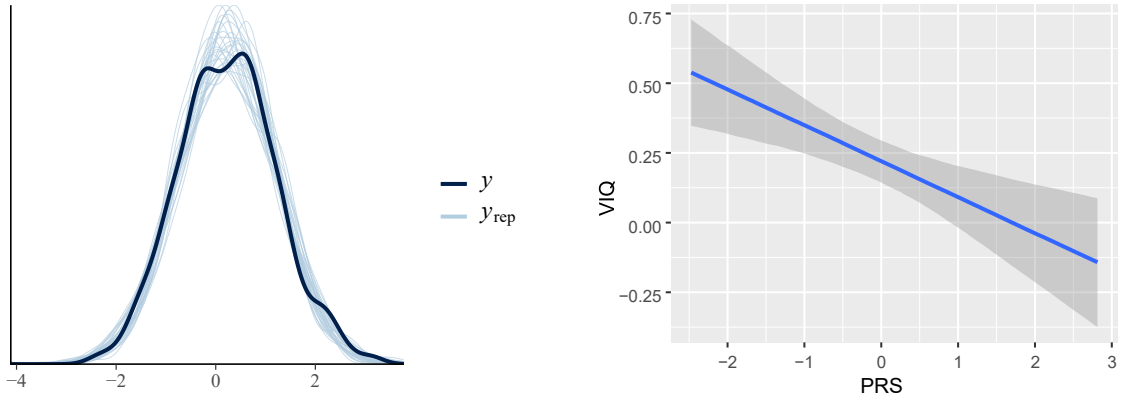

Figure 3: Left: Posterior Predictive Check of M1 for Verbal IQ; the dark-blue line (y) represents the density of observed data, the light-blue lines (yrep) represent the posterior predictive distributions obtained from 30 simulated samples. Right: Conditional Effect of the Dyslexia Polygenic Score on Verbal IQ; the blue line represents the regression line with the grey band representing the level of uncertainty.

with a posterior predictive check. With the package 'brms', which was used to run the multivariate multilevel model, the function 'pp_check' was used compare the actual data to the posterior predictive distribution. As Figure 3 shows on the left, this model actually predicts well the simulated data with the y and y_{rep} mostly overlapping, showing a good predictive ability of M1. In the Figure (3) on the right the graphic representation of the Conditional Effect shows a negative direction of the effect of the Dyslexia Polygenic Score on VIQ with a narrow uncertainty band, confirming the result of the Estimate and ultimately supporting the initial hypothesis.

PIQ

Moving on to Performance IQ – PIQ ~ PRS + (1 | FID) – , the model estimated a value of -0.05 ($95\% CI = [-0.13, 0.04]$), in this case it is very likely that the true value is around 0, therefore it can not be said that the Polygenic Score for Dyslexia has a negative influence on PIQ. The posterior probability of the PIQ's Estimate being of negative sign is of 87%. To investigate even further the effect

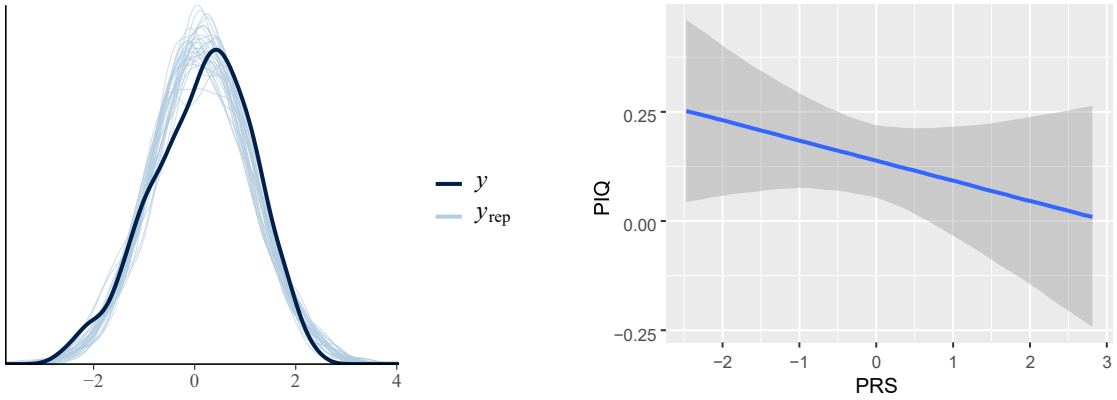

Figure 4: Left: Posterior Predictive Check of M1 for Performance IQ; the dark-blue line (y) represents the density of observed data, the light-blue lines (yrep) represent the posterior predictive distributions obtained from 30 simulated samples. Right: Conditional Effect of the Dyslexia Polygenic Score on Performance IQ; the blue line represents the regression line with the grey band representing the level of uncertainty.

size, R^2 will be compared to find it to be of 0.24 ($95\%CI = [0.18, 0.31]$) for M1 and of 0.24 ($95\%CI = [0.17, 0.3]$) for M0, indicating no variance explained by the Dyslexia Polygenic Score. Once again the Hypothesis is supported. In line with the diagnosis of Dyslexia stating that IQ is not affected, the Dyslexia Polygenic Score for dyslexia does not predict lower levels of Performance IQ, which relies less on verbal instruction compared to Verbal IQ. The model, as it shows in the Figure 4 (left panel), predicts well the simulated data, with the y and y_{rep} mostly overlapping, showing a good predictive ability of the model. In the Figure (4) on the right the graphic representation of the Conditional Effect shows a small but negative direction of the effect of the Dyslexia Polygenic Score on PIQ with a wider uncertainty band, confirming the result of the Estimate in line with the initial hypothesis.

Digit Symbol

Digit Symbol – DIGITSYM ~ PRS + (1 | FID) – had an estimated value of -0.11 ($95\%CI = [-0.19, -0.03]$), showing a negative influence of the Polygenic Score for Dyslexia on the task from the WAIS-R sub-test Digit Symbol, with support of the

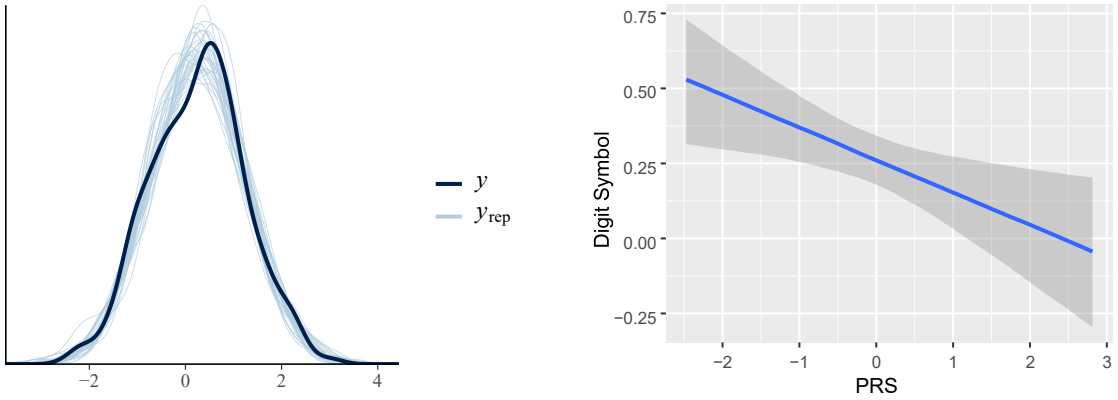

Figure 5: Left: Posterior Predictive Check of M1 for Digit Symbol; the dark-blue line (y) represents the density of observed data, the light-blue lines (yrep) represent the posterior predictive distributions obtained from 30 simulated samples. Right: Conditional Effect of the Dyslexia Polygenic Score on Digit Symbol; the blue line represents the regression line with the grey band representing the level of uncertainty.

Credible Interval not including the 0 in the range, making it very likely that the true value falls into a negative range. The Estimate has a posterior probability to fall in the negative range close to 1. In this case, the Effect Size calculated with the R^2 was of 0.241 ($95\%CI = [0.16, 0.32]$) and 0.221 ($95\%CI = [0.13, 0.3]$) for M1 and M0 respectively, showing once more little variance explained by the fixed effect in M1. According to the third hypothesis, the Digit Symbol task was expected to be negatively affected and results are in support of it. Once again a predictive check was replicated in Figure 5 (left panel) to see how simulated data from the posterior distribution were representative of the actual data and it can be seen a good overlap of y_{rep} on y, showing a good predictive ability of M1. In the Figure (5) on the right the graphic representation of the Conditional Effect shows a negative direction of the effect of the Dyslexia Polygenic Score on Digit Symbol with a narrow uncertainty band, confirming the result of the Estimate and ultimately supporting the initial hypothesis.

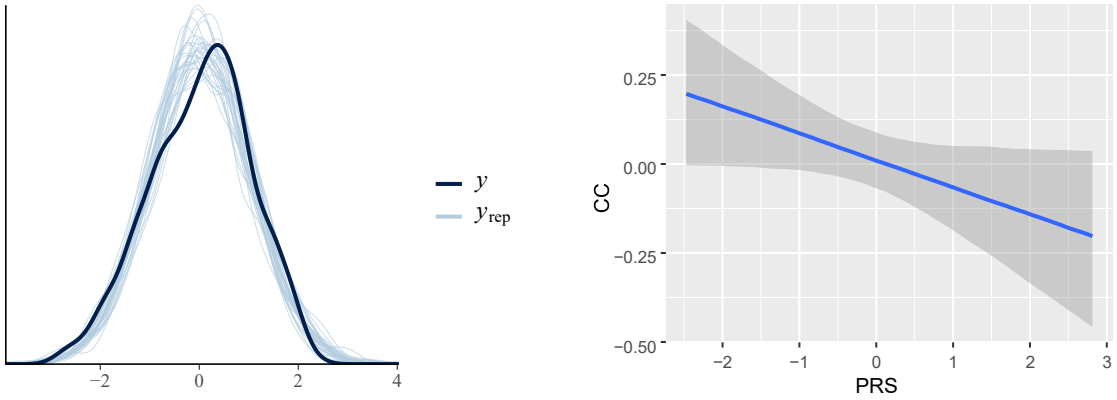

Figure 6: Left: Posterior Predictive Check of M1 for Comprehend and Collect; the dark-blue line (y) represents the density of observed data, the light-blue lines (yrep) represent the posterior predictive distributions obtained from 30 simulated samples. Right: Conditional Effect of the Dyslexia Polygenic Score on Comprehend and Collect; the blue line represents the regression line with the grey band representing the level of uncertainty.

Comprehend and Collect

In relation to the task from the QCST Comprehend and Collect – CC ~ PRS + (1 | FID) – lower academic achievement was expected for those with a high Dyslexia Polygenic Score. The estimate of CC in the model had a value of -0.08 ($95\%CI = [-0.16, 0.00]$), in this case the value for the upper part of the Credible Interval is 0, with a chance of getting an Estimate of negative sign based on the posterior probability distribution of 97%, so it could be said that the effect of the Dyslexia Polygenic Score on Academic Achievement is little, with an R^2 of 0.08 ($95\%CI = [0.03, 0.14]$) and 0.08 ($95\%CI = [0.03, 0.13]$) for M1 and M0 respectively that is quite small and also almost no variance explained by the Dyslexia Polygenic Score. Once more the initial hypothesis is supported with a slightly lower Academic Achievement for those with a high Polygenic Score for Dyslexia but with no support of the R^2 comparison of M1 and M0. For this part of the model another predictive check was replicated to see how simulated data was graphically overlapping with the actual data and it can be seen in Figure 6 (left panel) that the simulated data (y_{rep}) overlaps well with the actual data (y). In the Figure (6) on the right the

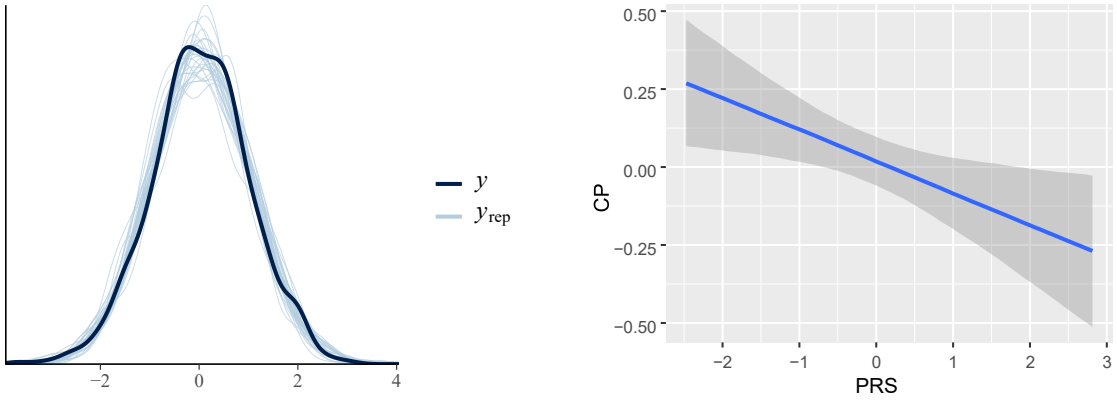

Figure 7: Left: Posterior Predictive Check of M1 for Create and Present; the dark-blue line (y) represents the density of observed data, the light-blue lines (yrep) represent the posterior predictive distributions obtained from 30 simulated samples. Right: Conditional Effect of the Dyslexia Polygenic Score on Create and Present; the blue line represents the regression line with the grey band representing the level of uncertainty.

graphic representation of the Conditional Effect shows a small negative effect of the Dyslexia Polygenic Score on Comprehend and Collect with the uncertainty band being neither too small nor too wide, confirming the result of the Estimate and ultimately supporting the initial hypothesis.

Create and Present

The second task from the QCST that we wanted to analyse in our model is Create and Present – CP ~ PRS + (1 | FID) – , which is an estimate of creative writing. The Estimate had a value of -0.10 ($95\%CI = [-0.18, -0.02]$), with the Credibility Interval not including the 0 and therefore supporting the hypothesis that the value could be negative and that the Polygenic Score for Dyslexia affects negatively creativity. The posterior probability shows a 99% chance of getting a negative Estimate. R^2 for this part of the model was of 0.06 ($95\%CI = [0.01, 0.12]$) and 0.05 ($95\%CI = [0.0, 0.12]$) for M1 and M0 respectively, showing little variance explained by the Dyslexia Polygenic Score. According to the first hypothesis a negative asso-

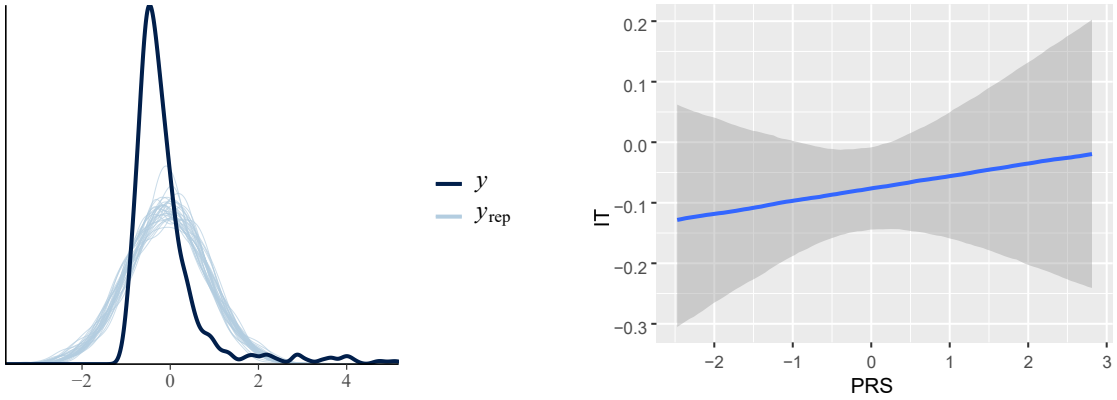

Figure 8: Left: Posterior Predictive Check of M1 for Inspection Time; the dark-blue line (y) represents the density of observed data, the light-blue lines (yrep) represent the posterior predictive distributions obtained from 30 simulated samples. Right: Conditional Effect of the Dyslexia Polygenic Score on Inspection Time; the blue line represents the regression line with the grey band representing the level of uncertainty.

ciation of Creativity and the Polygenic Score was expected. Results in this study weakly supported the findings of (Ritchie et al., 2013) where, instead of a positive association, a negative association between Dyslexia and Creativity was found. Also for Create and Present it can be seen how the simulated data (y_{rep}) in Figure 7 on the left overlaps well with the actual data (y), showing a good predictive ability of the model. In the Figure (7) on the right the graphic representation of the Conditional Effect shows a small negative effect of the Dyslexia Polygenic Score on Create and Present with the uncertainty band being neither too small nor too wide, confirming the result of the Estimate and ultimately supporting the initial hypothesis.

Inspection Time

Moving towards the end of the model, Inspection Time – IT ~ PRS + (1 | FID)– will be analyzed. In this case an Estimate of 0.02 ($95\% CI = [-0.05, 0.09]$) was found, having the 0 inside the Credible Interval and only a 30% chance in the posterior probability distribution of getting a negative result. For this part of the model, the

R^2 was 0.05 ($95\%CI = [0.00, 0.13]$) and 0.05 ($95\%CI = [0.0, 0.14]$) for M1 and M0, showing no variance explained by the Polygenic Score for Dyslexia. For the first time the initial hypothesis has not been supported, the Dyslexia Polygenic Score did not predict poor performance on the Inspection Time Task. In this case Figure 8 (left panel) shows a very bad ability of the model to predict simulated data based on the parameters of the M1 for Inspection Time, y is completely different in shape from y_{rep} with no overlap at all, showing a bad predictive ability of M1 and suggesting the need of a more complex model, probably not linear, to analyse those scores. In the Figure (8) on the right, the graphic representation of the Conditional Effect shows a small but positive effect of the Dyslexia Polygenic Score on Inspection Time with the uncertainty band being very wide, confirming the result of the Estimate but not supporting the initial hypothesis.

Choice Reaction Time

Lastly, this study aimed at investigating if Choice Reaction Time – CRT ~ PRS + (1 | FID) – was negatively predicted by the Polygenic Risk Score. The Estimate for CRT was of -0.01 ($95\%CI = [-0.05, 0.09]$), showing basically no effect of the Dyslexia Polygenic Score on CRT supported by a posterior probability of getting a negative result equal to 64%. The variance explained by the M1 was of 0.27 ($95\%CI = [0.18, 0.35]$), as the one of M0 – 0.27 ($95\%CI = [0.18, 0.35]$) –, both calculated with the R^2. Once more the Hypothesis could not be supported, by the results of this analysis it can not be said that there is a negative performance in Choice Reaction Time Task for those with a high Polygenic Score for Dyslexia. The last part of the model shows a good predictive quality, with the simulated data (y_{rep}) almost completely overlapping with the actual data (y) collected for this research. In the Figure (9) on the right, the graphic representation of the Conditional Effect shows a small but negative effect of the Dyslexia Polygenic Score on Choice

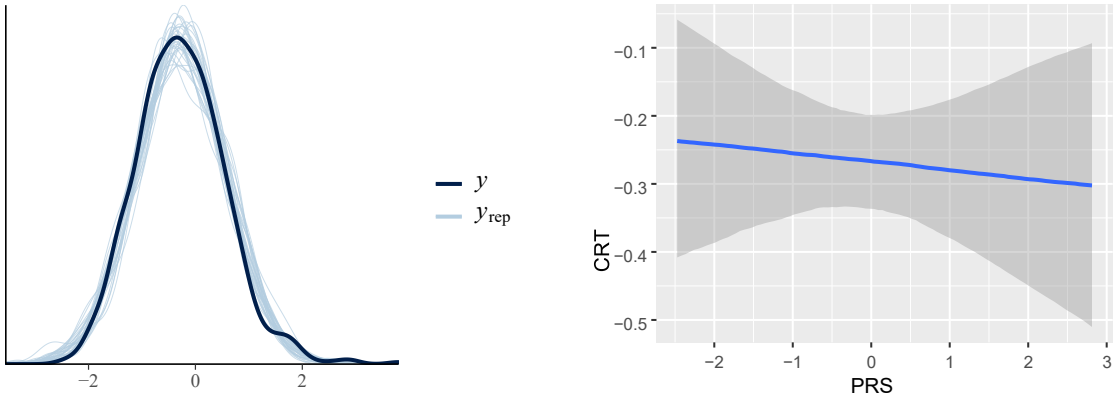

Figure 9: Left: Posterior Predictive Check of M1 for Choice Reaction Time; the dark-blue line (y) represents the density of observed data, the light-blue lines (yrep) represent the posterior predictive distributions obtained from 30 simulated samples. Right: Conditional Effect of the Dyslexia Polygenic Score on Choice Reaction Time; the blue line represents the regression line with the grey band representing the level of uncertainty.

Reaction Time with the uncertainty band being very wide, confirming the result of the Estimate and not supporting the initial hypothesis.

Discussion and Conclusion

The primary purpose of this thesis was to determine if the Dyslexia Polygenic Score accounts for some variance in cognitive abilities evaluated through some cognitive tests. It was hypothesized that the Polygenic Score would have a negative association with Academic Achievement, in particular with the sub-test from the Queensland Core Skill Test (QCST) Create and Present (CP), aimed at evaluating creative writing, a negative association with Verbal IQ and that it would be less strong for Performance IQ (PIQ), since it relied less on verbal instructions and verbal production. Lastly, high scores on the Polygenic Score, indicating higher risk for Dyslexia, were expected to have a negative impact on Inspection Time (IT) and Choice Reaction Time (CRT) tasks and Digit Symbol.

The first step has been to decide which model to use to analyze the data. Due to the nature of the research question, there was only one predictor, a random effect and many dependent variables. The choice of a Multivariate Multilevel Model was made to help increase the reliability of the parameters, having many dependent variables with the same predictive part, the risk was to overestimate the effect size of the models. A Parallel Factor Analysis was conducted on all the cognitive tests available to reduce the number of variables in the study. From the graphical representation of the correlation matrix, 3 main sub-groups of tests and 2 single

tests were intuitively visible. The choice was made to use: VIQ, PIQ, Digit Symbol, Comprehend and Collect, Create and Present, Inspection Time and Choice Reaction Time.

The first statistic adopted aimed at making sure that the model was the best model. To test it, the model (M1) was compared to a null model (M0) through a model comparison, using model weights, and results showed that M1 was 1.28 times more plausible of M0. M1 proved to be only slightly better than M0, with the addition of the Polygenic Risk Score not increasing significantly the goodness of the model. Most of the variance in the model is explained by the Family ID, meaning that the relation between the twins explains most of the variance in the scores. The decision to move forward with M1 was made, taking into consideration the weakness of the evidence of M1 being a better predictor of M0, and then each part of the model was analysed in subsections.

Not all the hypothesis have been confirmed. Regarding Academic Achievement and Creativity the initial hypothesis has been confirmed, the Dyslexia Polygenic Score predicted lower scores on Comprehend and Collect and Create and Present, confirming the results of Ritchie et al.(2013), who found a negative association between Dyslexia and Creativity, confuting the compensatory theory. Then a negative association with Verbal IQ was found and for Performance IQ the negative Estimate was not significant, confirming the second Hypothesis. Lastly, the Polygenic Score was expected to be affecting negatively those cognitive tests inspecting Information Processing Speed, this hypothesis has not been confirmed for Inspection Time and Choice Reaction Time, with both Estimates including the 0 in their Credibility Interval, not allowing to reject the null hypothesis, but a negative direction was found for Digit Symbol WAIS-R sub-test, confirming the hypothesis.

Some limitations must be acknowledged: the sample on which the analysis was carried out and also the cohort on which the Polygenic Score for Dyslexia were mostly from higher socio-economic status, not being representative of the total population and being a possible bias (M. J. Wright & Martin, 2004; Wainwright et al., 2005; Doust et al., 2022). Then, from the model comparison a weakness of our target model emerged, requiring to be cautious with the results of the analysis. Another limitation is that for each section the R^2, an indicator of the variance explained was reported both for M0 and M1, allowing to estimate the marginal variance explained, the part accountable by the Fixed Effect, resulting in little to no variance explained by the Dyslexia Polygenic Score, in line with past studies were the maximum variance explained was 6%.

For future research, a better randomisation of the participants is needed and in this specific project, a bigger sample size should be included in order to have more reliable results. With the great amount of data available, many more aspects could be investigated and a future goal should be to create a more powerful Dyslexia Polygenic Score. In next analysis, this research could be used to develop informative priors to develop a full Bayesian Approach. This research contributes to the scientific area of interest by providing for the first time an insight on the relation between a Polygenic Score for Dyslexia and Specific Cognitive Abilities, informing on the direction of the relation and giving the basis for the next studies in the same area.

The aim of this study was to investigate the relation between the Dyslexia Polygenic Score and Specific Cognitive Tests to better understand the influence of a propensity for reading difficulty on cognitive abilities. Polygenic scores have the great potential to become a valuable tool to help identify children with a propensity for dyslexia, allowing learning support before the development of reading skills (Doust et al., 2022). The final aim is to be able one day to support those children

allowing them to close the gap as much as possible with their peers in order to have better chances of succeeding at life.

References

Andreola, C., Mascheretti, S., Belotti, R., Ogliari, A., Marino, C., Battaglia, M., & Scaini, S. (2021). The heritability of reading and reading-related neurocognitive components: A multi-level meta-analysis. *Neuroscience & Biobehavioral Reviews*, *121*, 175–200.

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Washington, DC: Autor.

Authority, Q. S. (2008). Queensland curriculum, assessment and reporting (qcar) framework. *Retrieved February*, *13*, 2008.

Betjemann, R. S., Willcutt, E. G., Olson, R. K., Keenan, J. M., DeFries, J. C., & Wadsworth, S. J. (2008). Word reading and reading comprehension: Stability, overlap and independence. *Reading and Writing*, *21*(5), 539–558.

Brown Waesche, J. S., Schatschneider, C., Maner, J. K., Ahmed, Y., & Wagner, R. K. (2011). Examining agreement and longitudinal stability among traditional and rti-based definitions of reading disability using the affected-status agreement statistic. *Journal of learning disabilities*, *44*(3), 296–307.

Burnham, K., & Anderson, D. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Verlag.

Byrne, B., Coventry, W. L., Olson, R. K., Samuelsson, S., Corley, R., Willcutt, E. G., … DeFries, J. C. (2009). Genetic and environmental influences on

aspects of literacy and language in early childhood: Continuity and change from preschool to grade 2. *Journal of Neurolinguistics*, *22*(3), 219–236.

Byrne, B., Olson, R. K., Samuelsson, S., Wadsworth, S., Corley, R., DeFries, J. C., & Willcutt, E. (2006). Genetic and environmental influences on early literacy. *Journal of Research in Reading*, *29*(1), 33–49.

Bürkner, P.-C. (2018). Advanced bayesian multilevel modeling with the r package brms. *The R Journal*, *10*(1), 395–411. Retrieved from `https://doi.org/10.32614/RJ-2018-017` doi: 10.32614/RJ-2018-017

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science*, *343*(6178), 1478–1485.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*(1).

Castles, A. (2006). The dual route model and the developmental dyslexias. *London Review of Education*.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*(2), 149–180.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). Drc: a dual route cascaded model of visual word recognition and reading aloud. *Psychological review*, *108*(1), 204.

Court, M. H. (2007). A pharmacogenomics primer. *The Journal of Clinical Pharmacology*, *47*(9), 1087–1103.

Dale, P. S., Harlaar, N., & Plomin, R. (2005). Telephone testing and teacher assessment of reading skills in 7-year-olds: I. substantial correspondence for a sample of 5544 children and for extremes. *Reading and Writing*, *18*(5), 385–400.

de Oliveira, D. G., da Silva, P. B., Dias, N. M., Seabra, A. G., & Macedo, E. C.

(2014). Reading component skills in dyslexia: word recognition, comprehension and processing speed. *Frontiers in psychology*, *5*, 1339.

Dietrich, A. (2007). Who's afraid of a cognitive neuroscience of creativity? *Methods*, *42*(1), 22–27.

Dietrich, A., & Kanso, R. (2010). A review of eeg, erp, and neuroimaging studies of creativity and insight. *Psychological bulletin*, *136*(5), 822.

Doust, C., Fontanillas, P., Eising, E., Gordon, S. D., Wang, Z., Alagöz, G., … others (2022). Discovery of 42 genome-wide significant loci associated with dyslexia. *Nature Genetics*, 1–9.

Everatt, J., Steffert, B., & Smythe, I. (1999). An eye for the unusual: Creative thinking in dyslexics. *Dyslexia*, *5*(1), 28–46.

Falconer, D., & Mackay, T. (1960). Introduction to quantitative genetics edinburgh. *Oliver and*.

Fawcett, A. J., & Nicolson, R. I. (1995). The dyslexia early screening test. *The Irish Journal of Psychology*, *16*(3), 248–259.

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. E., Michaels, R., & Shaywitz, S. E. (2015). Achievement gap in reading is present as early as first grade and persists through adolescence. *The Journal of pediatrics*, *167*(5), 1121–1125.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2018). *Learning disabilities: From identification to intervention*. Guilford Publications.

Friend, A., Pennington, B. F., Smith, S. D., & Gilger, J. W. (2013). Dyslexia and related communication disorders. In *Emery and rimoin's principles and practice of medical genetics* (pp. 1–22). Elsevier Ltd.

Gaddes, W. H. (2013). *Learning disabilities and brain function: A neuropsychological approach*. Springer Science & Business Media.

Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for bayesian regression models. *The American Statistician*.

Gialluisi, A., Andlauer, T. F., Mirza-Schreiber, N., Moll, K., Becker, J., Hoffmann, P., … others (2019). Genome-wide association scan identifies new variants associated with a cognitive predictor of dyslexia. *Translational psychiatry*, *9*(1), 1–15.

Gialluisi, A., Andlauer, T. F., Mirza-Schreiber, N., Moll, K., Becker, J., Hoffmann, P., … others (2021). Genome-wide association study reveals new insights into the heritability and genetic correlates of developmental dyslexia. *Molecular psychiatry*, *26*(7), 3004–3017.

Gialluisi, A., Newbury, D. F., Wilcutt, E. G., Olson, R. K., DeFries, J. C., Brandler, W. M., … others (2014). Genome-wide screening for dna variants associated with reading and language traits. *Genes, Brain and Behavior*, *13*(7), 686–701.

Harlaar, N., Spinath, F. M., Dale, P. S., & Plomin, R. (2005). Genetic influences on early word recognition abilities and disabilities: A study of 7-year-old twins. *Journal of Child Psychology and Psychiatry*, *46*(4), 373–384.

Hawke, J. L., Wadsworth, S. J., & DeFries, J. C. (2006). Genetic influences on reading difficulties in boys and girls: the colorado twin study. *Dyslexia*, *12*(1), 21–29.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, *103*(6), 2052–86.

Hensler, B. S., Schatschneider, C., Taylor, J., & Wagner, R. K. (2010). Behavioral genetic approach to the study of dyslexia. *Journal of Developmental & Behavioral Pediatrics*, *31*(7), 525–532.

Hernandez, D. J. (2011). Double jeopardy: How third-grade reading skills and poverty influence high school graduation. *Annie E. Casey Foundation*.

Hespanhol, L., Vallio, C. S., Costa, L. M., & Saragiotto, B. T. (2019). Understanding and interpreting confidence and credible intervals around effect estimates. *Brazilian journal of physical therapy*, *23*(4), 290–301.

Jackson, D. N. (1984). *Mab, multidimensional aptitude battery: manual*. Research Psychologists Press.

Lambert, S. A., Abraham, G., & Inouye, M. (2019). Towards clinical utility of polygenic risk scores. *Human molecular genetics*, *28*(R2), R133–R142.

Lazaroo, N. K., Bates, T. C., Hansell, N. K., Wright, M. J., Martin, N. G., & Luciano, M. (2019). Genetic structure of iq, phonemic decoding skill, and academic achievement. *Frontiers in genetics*, *10*, 195.

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., … others (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature genetics*, *50*(8), 1112–1121.

Light, J. G., & DeFries, J. C. (1995). Comorbidity of reading and mathematics disabilities: Genetic and environmental etiologies. *Journal of learning disabilities*, *28*(2), 96–106.

Light, J. G., Defries, J. C., & Olson, R. K. (1998). Multivariate behavioral genetic analysis of achievement and cognitive measures in reading-disabled and control twin pairs. *Human Biology*, 215–237.

Lonigan, C. J., Purpura, D. J., Wilson, S. B., Walker, P. M., & Clancy-Menchetti, J. (2013). Evaluating the components of an emergent literacy intervention for preschool children at risk for reading difficulties. *Journal of experimental child psychology*, *114*(1), 111–130.

Luciano, M. (2017). Making reading easier: how genetic information can help. *Policy Insights from the Behavioral and Brain Sciences*, *4*(2), 147–154.

Luciano, M., Evans, D. M., Hansell, N. K., Medland, S. E., Montgomery, G. W., Martin, N. G., … Bates, T. C. (2013). A genome-wide association study for reading and language abilities in two population cohorts. *Genes, Brain and Behavior*, *12*(6), 645–652.

Luciano, M., Smith, G. A., Wright, M. J., Geffen, G. M., Geffen, L. B., & Martin,

N. G. (2001). On the heritability of inspection time and its covariance with iq: A twin study. *Intelligence*, *29*(6), 443–457.

Luciano, M., Wright, M., Smith, G., Geffen, G., Geffen, L., & Martin, N. (2001). Genetic covariance among measures of information processing speed, working memory, and iq. *Behavior genetics*, *31*(6), 581–592.

Luciano, M., Wright, M. J., Geffen, G. M., Geffen, L. B., Smith, G. A., & Martin, N. G. (2004). A genetic investigation of the covariation among inspection time, choice reaction time, and iq subtest scores. *Behavior genetics*, *34*(1), 41–50.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of dyslexia*, *53*(1), 1–14.

Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Petersen, A. (1996). On the bases of two subtypes of development dyslexia. *Cognition*, *58*(2), 157–195.

Matters, G., & Gray, K. (1995). The queensland core skills test: implications for the mathematical sciences [abridged paper presented to the national conference on assessment in the mathematical sciences (2nd: 1993).]. *Unicorn (Carlton, Vic)*, *21*(4), 74–89.

Meaburn, E. L., Harlaar, N., Craig, I., Schalkwyk, L., & Plomin, R. (2008). Quantitative trait locus association scan of early reading disability and ability using pooled dna and 100k snp microarrays in a sample of 5760 children. *Molecular psychiatry*, *13*(7), 729–740.

Michail, K. (2010). *Dyslexia: The experiences of university students with dyslexia* (Unpublished doctoral dissertation). University of Birmingham.

Ott, P. (1997). *How to detect and manage dyslexia: A reference and resource manual*. Heinemann.

Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of

kindergarten children: implications for early screening and theories of reading. *Developmental science*, *20*(5), e12471.

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The lancet*, *379*(9830), 1997–2007.

Price, K. M., Wigg, K. G., Feng, Y., Blokland, K., Wilkinson, M., He, G., … others (2020). Genome-wide association study of word reading: Overlap with risk genes for neurodevelopmental disorders. *Genes, Brain and Behavior*, *19*(6), e12648.

Rack, J. P., & Olson, R. K. (1993). Phonological deficits, iq, and individual differences in reading disability: Genetic and environmental influences. *Developmental Review*, *13*(3), 269–278.

Reiss, M., & Brooks, G. (2004). *Developmental dyslexia in adults: a research review.* National Research and Development Centre for Adult Literacy and Numeracy.

Richardson, J. T., & Wydell, T. N. (2003). The representation and attainment of students with dyslexia in uk higher education. *Reading and Writing*, *16*(5), 475–503.

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in bioinformatics*, *3*(2), 119–133.

Ritchie, S. J., Luciano, M., Hansell, N. K., Wright, M. J., & Bates, T. C. (2013). The relationship of reading ability to creativity: Positive, not negative associations. *Learning and Individual Differences*, *26*, 171–176.

Rocha, D., Gut, I., Jeffreys, A. J., Kwok, P.-Y., Brookes, A. J., & Chanock, S. J. (2006). Seventh international meeting on single nucleotide polymorphism and complex genome analysis:'ever bigger scans and an increasingly variable genome'. *Human genetics*, *119*(4), 451–456.

Scerri, T. S., & Schulte-Körne, G. (2010). Genetics of developmental dyslexia. *European child & adolescent psychiatry*, *19*(3), 179–197.

Selzam, S., Dale, P. S., Wagner, R. K., DeFries, J. C., Cederlöf, M., O'Reilly, P. F., …

Plomin, R. (2017). Genome-wide polygenic scores predict reading performance throughout the school years. *Scientific Studies of Reading*, *21*(4), 334–349.

Snowling, M. J., Hulme, C., & Nation, K. (2020). Defining and understanding dyslexia: past, present and future. *Oxford Review of Education*, *46*(4), 501–513.

Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin*, *142*, 498 - 545.

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of educational psychology*, *89*(1), 114.

Stevenson, J. (1991). Which aspects of processing text mediate genetic effects? In *Reading disabilities* (pp. 61–81). Springer.

Stevenson, J., Graham, P., Fredman, G., & Mcloughli, V. (1987). A twin study of genetic influences on reading and spelling ability and disability. *Journal of child psychology and psychiatry*, *28*(2), 229–247.

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., … Kattan, M. W. (2010). Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass.)*, *21*(1), 128.

Truong, D. T., Adams, A. K., Paniagua, S., Frijters, J. C., Boada, R., Hill, D. E., … others (2019). Multivariate genome-wide association study of rapid automatised naming and rapid alternating stimulus in hispanic american and african–american youth. *Journal of medical genetics*, *56*(8), 557–566.

Uffelmann, E., Huang, Q. Q., Munung, N. S., De Vries, J., Okada, Y., Martin, A. R., … Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, *1*(1), 1–21.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing*, *27*(5),

1413–1432.

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1998). The case for early intervention in diagnosing specific reading disability. *Journal of School Psychology*, *36*(4), 367–397.

Wagner, R. K., Zirps, F. A., Edwards, A. A., Wood, S. G., Joyner, R. E., Becker, B. J., … Beal, B. (2020). The prevalence of dyslexia: A new approach to its estimation. *Journal of learning disabilities*, *53*(5), 354–365.

Wainwright, M. A., Wright, M. J., Geffen, G. M., Luciano, M., & Martin, N. G. (2005). The genetic basis of academic achievement on the queensland core skills test and its shared genetic variance with iq. *Behavior genetics*, *35*(2), 133–145.

Wainwright, M. A., Wright, M. J., Luciano, M., Geffen, G. M., & Martin, N. G. (2008). Genetic covariation among facets of openness to experience and general cognitive ability. *Twin research and human genetics*, *11*(3), 275–286.

Wechsler, D. (1981). *Wais-r manual: Wechsler adult intelligence scale-revised*. Psychological Corporation.

Wolf, M., & Bowers, P. (1999). The question of naming-speed deficits in developmental reading disabilities: An introduction to the double-deficit hypothesis. *Journal of Educational Psychology*, *19*(1), 1–24.

Wolf, M., & Denckla, M. B. (2005). *Ran/ras: Rapid automatized naming and rapid alternating stimulus tests*. Pro-ed Austin, TX.

Wright, M., De Geus, E., Ando, J., Luciano, M., Posthuma, D., Ono, Y., … others (2001). Genetics of cognition: outline of a collaborative twin study. *Twin Research and Human Genetics*, *4*(1), 48–56.

Wright, M. J., & Martin, N. G. (2004). Brisbane adolescent twin study: outline of study methods and research projects. *Australian Journal of Psychology*, *56*(2), 65–78.

Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. (2018). Using stacking to average

bayesian predictive distributions (with discussion). *Bayesian Analysis*, *13*(3), 917–1007.