MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS

WWW.MPI.NL

Why do we change how we speak?

Multivariate genetic analyses of language and related traits across development and disorder

ELLEN VERHOEF

Why do we change how we speak?

Multivariate genetic analyses of language and related traits across development and disorder

Ellen Verhoef

The educational component of the doctoral training was provided by the International Max Planck Research School (IMPRS) for Language Sciences. The graduate school is a joint initiative between the Max Planck Institute for Psycholinguistics and two partner institutes at Radboud University – the Centre for Language Studies, and the Donders Institute for Brain, Cognition and Behaviour. The IMPRS curriculum, which is funded by the Max Planck Society for the Advancement of Science, ensures that each member receives interdisciplinary training in the language sciences and develops a well-rounded skill set in preparation for fulfilling careers in academia and beyond.

More information can be found at www.mpi.nl/imprs.

ISBN: 978-94-92910-25-7

Cover design and lay-out by Ellen Verhoef
Printed and bound by Ipskamp Drukkers, Nijmegen

The research reported in this thesis was conducted at the Max Planck Institute for Psycholinguistics, in Nijmegen, the Netherlands.

Why do we change how we speak?

Multivariate genetic analyses of language and related traits across development and disorder

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 5 maart 2021
om 10.30 uur precies

door
Ellen Verhoef
geboren op 10 maart 1993
te Enschede

Promotor
Prof. dr. S.E. Fisher

Copromotor
Dr. B. St Pourcain (MPI)

Manuscriptcommissie
Prof. dr. J.K. Buitelaar
Prof. dr. B. Müller-Myhsok (MPI, München, Duitsland)
Dr. M.E. Hayiou-Thomas (University of York, Verenigd Koninkrijk)

Table of contents

Chapter 1

General introduction

1.1. Outline

Language is a complex human capacity[1–4] that allows us to acquire knowledge, share thoughts, convey feelings, and report experiences. According to current views of linguistics, its main components are phonology, morphology, syntax, semantics and pragmatics[5]. The first three of these components all relate to form: phonology entails the use of individual sound units, morphology the use of the smallest units of language that have meaning, and syntax the organisation of words into phrases or sentences. Semantics refers to the meaning of language, at both the single word and the word combination level. Finally, pragmatics integrates the skills mentioned above in order to use language in a social context. These five interrelated components all play a role in spoken language (listening and speaking) as well as literacy (reading and writing).

Acquiring the components of language is not a simple task[4,6] and is influenced by multiple biological, cognitive, psychosocial and environmental factors. Specifically, language acquisition involves a combination of neural commitment (commitment of neuronal networks to patterns that reflect natural language input), social skills, computational abilities, and pattern detection[7]. Together these processes enable us to recognise and, subsequently, produce speech sounds, combine them to make meaningful words, and create sentences, all within a conversational context[4]. Note that children who acquire sign language will pass through similar developmental stages, suggesting that comparable psychological, linguistic and neuronal mechanisms are involved[8]. In this thesis, I focused on spoken language development.

Despite the complexity of the task, most children acquire language rapidly and effortlessly[1–4]. However, there exist large individual differences in language abilities during the first few years of life[9], which are predictive of future language and literacy skills[10–14], as well as academic achievement[15]. At least part of these individual differences can be explained by genetic factors. The main aim of the present thesis was to gain deeper knowledge about the genetic factors influencing language abilities during the first three years of life, including their roles in subsequent language, literacy and cognitive development, as well as genetic links between language and related traits with childhood-onset neurodevelopmental disorders.

1.2. Language development in typically developing children

Language development already starts *in utero*, though the first vocalisations can only be detected after birth[12]. To assess the language abilities of infants and toddlers, researchers often study lexical development with measures of expressive and receptive vocabulary[1,2]. These constructs relate to the ability to produce and understand language, respectively. In addition to building a lexicon, children also acquire more complex

linguistic constructions in their preschool years, including aspects of morphology and syntax[16].

Important processes during the first year of life that contribute to later spoken language development include auditory processing[17] and word segmentation from a continuous speech stream[18,19], which are both related to speech perception. The ability to process and discriminate rapid auditory signals starts to develop *in utero* and enables hearing children to discriminate linguistically-relevant contrasts in speech when they are around six months of age[7,20]. During the second half of the first-year, children become able to successfully identify words within utterances (word segmentation), another important skill for building a lexicon[18]. Indeed, between six and nine months of age children know the meanings of several common words, the start of their receptive vocabulary[21] (Figure 1).

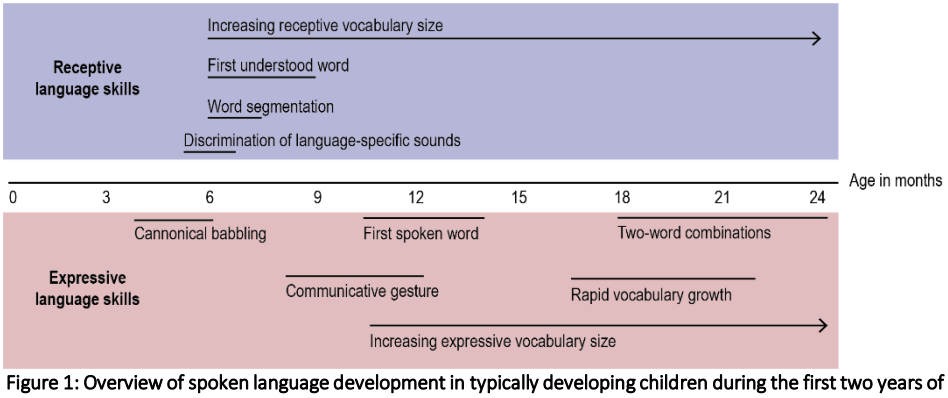


Figure 1: Overview of spoken language development in typically developing children during the first two years of life.

One of the first precursors of language production in typically developing children is canonical babbling[22] and the use of gesture[23] (Figure 1). Canonical babbling emerges from the age of four to six months[1] and is followed by the production of the first spoken word between 10 and 15 months of age[2], a few months later than infants understand the first words (Figure 1). With progressing development, the number of words produced increases, reaching ~50 words at 12-18 months of age. This is often trailed by a period of rapid growth of the lexicon around 16-22 months of age[24] and a steady increase after that, resulting in an expressive vocabulary size of ~500 words at 30 months[25]. In addition to an expanding vocabulary size, the complexity of language use increases during early development. Whereas the first words are usually spoken in isolation[2], children start to produce two-word combinations from the age of 18 to 24 months[26,27] onwards, followed by the use of more complex grammatical structures[26,27]. The transition from the single-word to the two-word combination phase is also seen as the onset of grammar[26,27].

When children are still limited in their verbal communication using words, they extend their communicative abilities via the use of gestures, starting between 8 and 12 months of age[28] (Figure 1). The use of gesture during the first few years of life is predictive for later language development[29]. For example, a meta-analysis of 25 studies and 734 children showed that a more frequent use of gesture at 9-24 months of age, especially declarative pointing, was related to a larger vocabulary at the same or later developmental timepoints (9-54 months)[30]. Declarative pointing is thought to be an intentional communication behaviour that a child uses to direct someone's attention to an event or object in the world[30]. Assuming that such behaviour indicates that a child understands other's intentions, the use of declarative pointing forms an early manifest of social understanding[31], possibly reflecting joint attention[30]. Furthermore, children that use gesture-word combinations early on also tend to be the first ones producing two-word combinations[32].

An important point to note about early language development is the existence of large individual differences[9]. These are apparent from infancy to early childhood and have been observed for different language aspects, including vocabulary size[33], grammatical competence[34,35] and pragmatic development[36]. Individual differences in vocabulary size exist within age windows, but children also differ in their developmental trajectory. For example, at 24 months of age, the number of words a child produces may vary between zero to over 600, with a median of ~300 words (Figure 2).

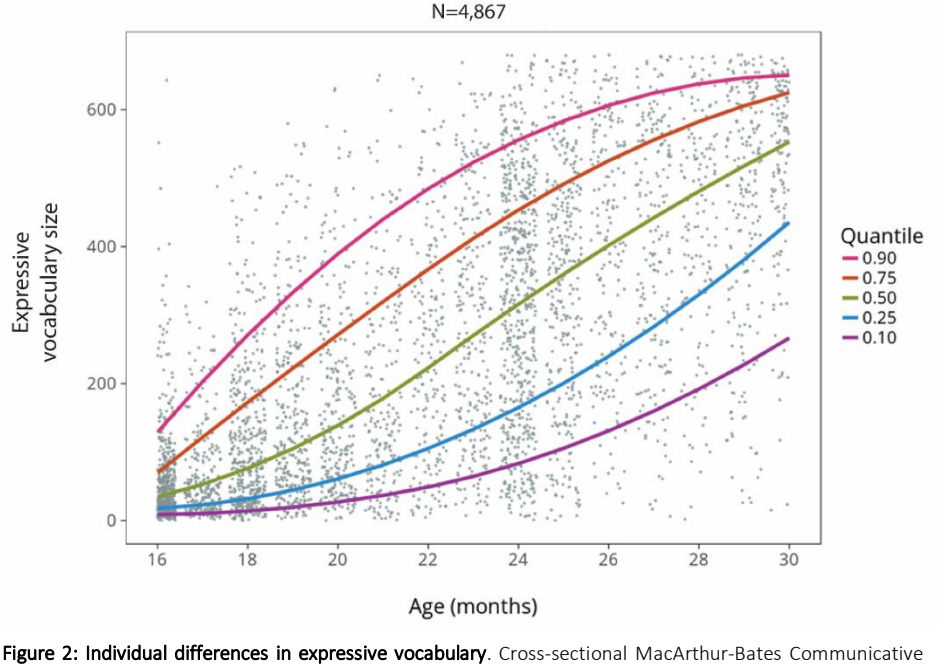


Figure 2: Individual differences in expressive vocabulary. Cross-sectional MacArthur-Bates Communicative Development Inventory expressive vocabulary data (N=4,687) based on English-speaking children between 16-30 months of age. Data were downloaded from the Wordbank[31] on 22 May 2020.

After a period of rapid development during the first years of life, language abilities continue to advance more steadily during mid-childhood and early adolescence. A major developmental milestone during this period, which is crucially shaped by language proficiency, is the acquisition of literacy[37]. According to the "simple view of reading" theory, reading comprehension is the product of decoding (recognition of printed words) and oral language comprehension[38]. Thus, in order to learn to read, children need to be able to decode written text and comprehend oral language[38]. Early vocabulary has a central role in the development of both these abilities[39]. Decoding is substantially influenced by phonological awareness (i.e. the awareness of sound structures of speech)[40], which develops during the preschool period and relates to vocabulary size[41–43]. Listening comprehension (i.e. the understanding of spoken language) necessarily begins with vocabulary comprehension[43].

The proposed relationships between language skills during infancy and the acquisition of language and literacy abilities in childhood have been supported by studies of individual differences[10–14]. For example, vocabulary size at 16-24 months is predictive of vocabulary size, as well as performance on tests of phonological awareness, reading accuracy and reading comprehension assessed five years later[44]. Similarly, a larger expressive vocabulary at 24 months has been associated with larger vocabulary size and better decoding, word recognition, and passage comprehension skills assessed up to primary school[10]. Thus, individual differences in language skills during the preschool period are predictive of individual differences in language and literacy abilities later in life, suggesting shared underlying aetiologies.

1.3. Language and literacy development in children diagnosed with a neurodevelopmental disorder

Children with a neurodevelopmental disorder often suffer from impairments in various developmental domains, including language and cognition[45,46]. This includes disorders that are characterised by a primary deficit in speech and/or language abilities, such as developmental dyslexia (also known as reading disability) and specific language impairment (also known as developmental language disorder), although cohort sizes for neurobiological and genetic studies of such disorders are currently rather low. However, also children diagnosed with neurodevelopmental disorders that are characterised by different primary deficits, such as Attention-Deficit/Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD), may show problems with language and communication[46], and for those disorders there already are large sample sizes available for neurogenetic investigations. In this thesis, I studied genetic overlap of ADHD and ASD with language and literacy skills and, in extension, educational attainment and cognitive abilities.

Attention-Deficit/Hyperactivity Disorder

ADHD is a complex childhood-onset neurodevelopmental condition that affects about 5% of the general population[47] and is characterised by hyperactive, inattentive and impulsive symptoms[48]. In addition to these symptoms, children diagnosed with ADHD often experience difficulties with mastering language and literacy skills[49–51]. For example, poor language skills at three years of age were found to be predictive of inattention and hyperactive symptoms two years later in life[52]. A small-scale observational study (N=21) suggested that language problems at the age of two years may predict an ADHD diagnosis in children aged between six and seven years[53]. Later on in life, ADHD has been linked to reading impairments, with up to 40% of the children diagnosed with ADHD also suffering from reading disability (also known as developmental dyslexia) and vice versa[54]. In addition, abilities related to syntax[55,56], phonology[55,56], writing[57,58] and spelling[59,60] might be impaired in children with ADHD. Thus, a variety of language and literacy skills that are acquired from infancy to adolescence may be impaired in children diagnosed with ADHD.

Autism Spectrum Disorder

ASD is a neurodevelopmental condition that affects about 1-1.5% of the general population[61]. It is a term for a group of pervasive disorders that are characterised by symptoms of repetitive and restrictive behaviour, as well as social and communication impairments[62]. Children with ASD may show communication and language difficulties already very early in life. Several studies reported evidence suggesting that problems related to joint attention and the use of gesture can be observed in children with ASD as young as 8 to 24 months of age[63,64]. Individuals at high risk for ASD, identified based on an Autistic Disorder diagnosis of an older full biological sibling, have been shown to have an impaired vocabulary growth compared to low-risk individuals[65], and a substantial proportion of children diagnosed with ASD (~30%) has little or no spontanous spoken language by the time they reach school age[66]. However, the phenotypic spectrum of patients with ASD is highly heterogeneous, ranging from individuals with severe cognitive impairments to individuals with an above-average intelligence quotient and high academic functioning. These 'high-functioning' ASD patients often experience less severe deficits, particularly in the domain of language[67].

1.4. Language abilities: A multifactorial and complex aetiology

Language abilities are multifactorial traits that are influenced by environmental and genetic factors, as well as their interplay. Environmental factors include, for example, the quantity and quality of language input a child receives in their environment[68], both related to family socio-economic status[69]. Genetically, individual

differences in language skills are thought to be shaped by multiple genetic factors that form a polygenic architecture[70] (Figure 3). A polygenic trait is influenced by many genetic variants that each account for a very small proportion of the trait variance, but together explain a substantial proportion. Genetic variants may affect a trait via an additive mode of action, via non-additive interaction (epistasis)[71], and/or via interactions with the environment[72]. In this thesis, I focused on additive genetic variation, implying that the combined effect of multiple genetic variants equals the sum of their individual effects (Figure 3).

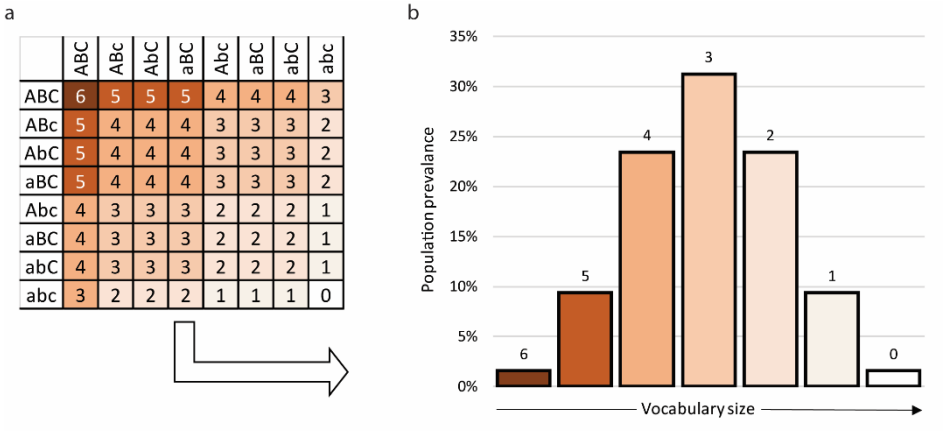


Figure 3: Polygenic inheritance for vocabulary size. Polygenic inheritance for vocabulary size assuming three underlying genetic variants, resulting in 64 possible genotypes. **(a)** Table depiciting the number of risk alleles per possible genotype. Risk alleles are depicted with capitals and are associated with a smaller vocabulary size. Genotypes that result in a similar number of risk alleles have the same color. **(b)** Frequency plot of the risk allele distribution in the population based on (a).

The proportion of trait variance that can be attributed to genetic variance (differences in the genetic make-up between individuals) is defined as heritability. Twin studies estimate heritability based on phenotypic correlations in monozygotic and dizygotic twin pairs (twin-h^2) and have suggested that between 10% and 25% of the variation in early expressive vocabulary is due to genetic factors when investigated within a community-based sample of children from the UK[73,74]. Similarly, a community-based twin study in the US reported a twin-h^2 of ~30% for variation in early receptive vocabulary[75]. Studies of population-based samples including unrelated individuals have furthermore shown that a substantial proportion of this genetic variance can be attributed to common genetic variation (minor allele frequency > 0.01), as tagged by bi-allelic markers on commercial genotyping arrays[76]. The proportion of total genetic variance that can be explained by additive effects as captured by common genetic variants, typically single-nucleotide polymorphisms (SNPs), is known as SNP-heritability (SNP-h^2).

The development of the genome-wide association study (GWAS) design has facilitated the hypothesis-free identification of SNPs that are associated with a trait[77]. Due to the small effect sizes for individual SNPs, and the need to make a stringent correction for multiple testing of millions of genetic variants, a GWAS requires several thousands of participants to have sufficient statistical power. To achieve such large sample sizes researchers often perform a meta-GWAS, a study design that includes a meta-analysis across multiple GWASs from individual cohorts. There is only one meta-GWAS available for early language abilities to date, which focused on expressive vocabulary[76].

Recent research showed that more than half of the loci in the genome are associated with at least one trait[78]. The vast majority of these loci (~90%), however, is related to multiple traits (with associations often identified in different studies)[78]. This complex phenomenon where a gene or genetic variant is correlated with more than one trait is known as pleiotropy[79]. A quantitative parameter that is thought to capture pleiotropic effects and that assesses the genetic relationship between two traits at the polygenic level is genetic correlation. A related concept, so-called bivariate heritability, quantifies the amount of phenotypic covariance between two traits that can be attributed to shared genetic influences. For example, twin studies on reading deficits and ADHD reported evidence for bivariate heritabilities up to 95%[80–82], suggesting a role for pleiotropy in the co-occurence of reading impairments and ADHD. However, both ADHD and reading abilities are also genetically related to educational attainment (EA)[83–85], which could thus be a third common factor that plays a role in the reported pleiotropy between reading skills and ADHD. To disentangle such complex relationships among multiple genetically correlated traits sophisticated methodologies are required, as adjusting for heritable covariates may lead to biased findings, also known as collider bias[86].

1.5. Recent advances in the field of behaviour genetics

The wide availability of high-throughput genotyping arrays, the advances in genomic methodologies such as GWASs, and the increase in data sharing during the last decade make it possible to carry out molecular genetic analyses of language development based on samples of unrelated individuals. Especially the development of GWASs and subsequent sharing of the relevant outputs, so-called GWAS summary statistics, represent a major advantage for the field of behaviour genetics. Researchers can now search for SNPs that are associated with a certain trait with sufficient statistical power, by combining data from several cohorts. This has resulted in the availability of very large datasets for disorders with well-defined diagnostic criteria such as ADHD[85] and ASD[61], as well as for broad cognition-related constructs such as general intelligence[87] and

EA[88]. However, the sample sizes of GWASs on developmental speech, language and/or reading disorders are currently still small, compared to GWAS efforts on ADHD[85] and ASD[61], and thus have, limited statistical power[89]. For example, the largest meta-GWAS on ADHD to date was based on 53,293 European individuals[85], compared to 3,468 individuals for a meta-GWAS on word reading, spelling, decoding skills, phoneme awareness, verbal short-term memory and naming speed in nine cohorts of European individuals with reading impairments and typically developing participants[90].

Recently developed methodologies in genetic epidemiology now allow the study of genetic links between traits and disorders based on phenotypes assessed in different cohorts[91], using methods such as polygenic scoring[92] and Linkage Disequilibrium SCore (LDSC) correlation[93]. However, in order for LDSC analyses to have sufficient statistical power, a minimal sample size of 5,000 individuals is recommended, limiting the possibilities to study genetic overlap early language abilities with developmental speech, language and/or reading disorders based on summary statistics. Finally, during the past few years, statistical methodologies have become available that allow for the study of genetic relationships between more than two traits, such as genetic-relationship-matrix structural equation modelling (GSEM)[94], providing the opportunity to further elucidate the complex genetic architecture underlying language abilities during different developmental stages.

Taken together, the broad availability of genome-wide genetic data together with methodological advances provide a great opportunity to study the complex genetic components underlying language abilities during development. In order to improve our understanding of the involved genetic mechanisms, we need to (i) assess genetic architectures underlying language and related abilities (i.e. literacy) in a developmental context, and (ii) disentangle the complex genetic mechanisms underlying the genetic overlap of language and related abilities with neurodevelopmental disorders.

1.6. Aims of this thesis

This thesis was focussed on the developmental genetic architecture underlying language development as captured by genome-wide variation on genotyping chips. Here, I investigated genetic mechanisms and variance compositions of vocabulary skills assessed from infancy to early childhood, and studied their genetic relationships with mid-childhood/early-adolescent language and literacy abilities. Using powerful proxy measures, such as mid-childhoood language and literacy skills, as well as educational attainment, I studied the complex mechanisms underlying genetic overlap with several neurodevelopmental disorders. To this end, I investigated population-based cohorts and community twin samples with genome-wide genotyping data, and phenotype information on language and literacy skills. I also analysed existing summary statistics of

powerful GWAS studies on general intelligence and EA, as well as neurodevelopmental disorders, such as ASD and ADHD.

In chapter 2, I carried out a review of the literature on genetic factors underlying language development during infancy and early childhood. I provided an overview of heritability estimates for language abilities assessed before the age of four years based on both twin collections and samples of unrelated individuals. Next, I focused on the role of common genetic variation in early language abilities and described individual genetic variants that have previously been associated with vocabulary scores during infancy. I also discussed the role of genetic factors in the link between early vocabulary skills (0-4 years) and subsequent language- and literacy-related abilities, assessed from mid-childhood to early adolescence. Finally, I outlined current challenges for studying the genetic architecture underlying early language development.

In chapter 3, I investigated multivariate genetic relationships between expressive and receptive vocabulary in three-year-old children and their later language and literacy skills, as assessed between 7 and 13 years of age in a large birth cohort, the Avon Longitudinal Study of Parents and Children (ALSPAC). Applying a structural equation modelling approach based on directly genotyped genome-wide data (GSEM), I investigated the developmental origins of genetic factors for a wide range of mid-childhood/early-adolescent language and literacy abilities related to reading, spelling, phonemic awareness, listening comprehension, non-word repetition and verbal intelligence.

In chapter 4, I examined the genetic architecture of expressive and receptive vocabulary across early developmental stages in detail (15-38 months). Subsequently, I assessed the emergence of genetic associations with mid-childhood reading, verbal and non-verbal intelligence (7-8 years). For this, I adopted a similar approach to that described in chapter 3 based on assessments from the ALSPAC cohort. This study design allowed me to (i) identify whether vocabulary size at different developmental timepoints during the first four years of life is affected by the same or different genetic factors, (ii) study to what extent expressive and receptive vocabulary share genetic sources, and (iii) determine which of the identified early genetic factors are important for literacy and/or cognition later on in life.

In chapter 5, I performed a meta-analysis of GWASs on vocabulary size assessed during the first three years of life. This is the largest meta-GWAS available to date and extends a previous effort[76] by containing a wider age range, a larger number of samples, and measures of both expressive and receptive vocabulary. In order to identify SNPs that are associated with vocabulary size, I performed meta-analyses across two developmental phases, an early phase (15-18 months) and a late phase (24-38 months), including data from seven independent population- and community-based cohorts. Finally, I augmented the statistical power of GWAS meta-analyses by implementing a multivariate analysis approach that combines data across developemental stages and

traits. Using the derived summary statistics, I investigated evidence for shared genetic factors between early vocabulary and, for example, cognition-related later life outcomes, infant and childhood antropometric traits, as well as several neurodevelopmental disorders.

In chapter 6, I used a polygenic scoring approach to investigate the genetic associations between ADHD and thirteen language- and literacy-related skills assessed from mid-childhood to early adolescence. Using multivariable regression analyses, I examined whether shared genetic influences between ADHD and language and literacy skills were due to a third factor, genetically predicted EA. This approach allowed me to disentangle relationships between multiple genetically correlated traits while controlling for bias[95]. These analyses were carried out using genotype and mid-childhood/early-adolescent language and literacy data from ALSPAC, as well as GWAS summary statistics for ADHD and EA using a multivariate analysis approach.

In chapter 7, I examined genetic links between ADHD, ASD and EA in a multivariate context. This study included EA as a powerful genetic proxy measure for literacy. Applying multivariable regression analyses based on GWAS summary statistics, I gained insight into the complex pleiotropic mechanisms underlying the genetic overlap between ADHD, ASD and EA. More specifically, I (i) studied the genetic mechanisms underlying the discordant polygenic association pattern between EA, and both ASD and ADHD risk, (ii) characterised the risk variants that contributed most strongly to this polygenic association pattern, and (iii) assessed the specificity of the observed polygenic association pattern by studying genetic overlap for other psychiatric disorders.

Finally, in chapter 8, I summarised and reviewed the main findings of chapters 3 to 7. I discussed the contributions of my findings with respect to scientifc challenges as described in chapter 2 and outlined future research perspectives.

References

1. Kennison, S. M. *Introduction to language development*. (SAGA, 2014).
2. Clark, E. V. *First language acquisition*. (Cambridge University Press, 2016).
3. Saxton, M. *Child language*. (SAGE, 2010).
4. Rowland, C. *Understanding child language acquisition*. (Routledge, 2014).
5. Gleason, J. B. & Ratner, N. B. *The Development of Language*. (Pearson, 2016).
6. Barrett, M. D. *The Development of Language*. (Psychology Press, 1999).
7. Kuhl, P. K. Early language acquisition: cracking the speech code. *Nat Rev Neurosci* **5**, 831–43 (2004).
8. Emmorey, K. *Language, Cognition, and the Brain: Insights From Sign Language Research*. (Psychology Press, 2001).
9. Kidd, E., Donnelly, S. & Christiansen, M. H. Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences* **22**, 154–169 (2018).
10. Lee, J. Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics* **32**, 69–92 (2011).
11. Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev Sci* **15**, 233–49 (2012).
12. Bornstein, M. H., Hahn, C.-S., Putnick, D. L. & Suwalsky, J. T. D. Stability of Core Language Skill from Early Childhood to Adolescence: A Latent Variable Approach. *Child Development* **85**, 1346–1356 (2014).
13. Harlaar, N., Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. Why Do Preschool Language Abilities Correlate With Later Reading? A Twin Study. *J Speech Lang Hear Res* **51**, 688–705 (2008).
14. Bleses, D., Makransky, G., Dale, P. S., Højen, A. & Ari, B. A. Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics* **37**, 1461–1476 (2016).
15. Hohm, E., Jennen-Steinmetz, C., Schmidt, M. H. & Laucht, M. Language development at ten months. Predictive of language outcome and school achievement ten years later? *Eur Child Adolesc Psychiatry* **16**, 149–56 (2007).
16. *Language Development: The Essential Readings*. (Wiley-Blackwell, 2001).
17. Chonchaiya, W. *et al.* Developmental trends in auditory processing can provide early predictions of language acquisition in young infants. *Developmental Science* **16**, 159–172 (2013).
18. Kidd, E. & Donnelly, S. Individual Differences in First Language Acquisition. *Annu. Rev. Linguist.* **6**, 319–340 (2020).
19. Junge, C., Kooijman, V., Hagoort, P. & Cutler, A. Rapid recognition at 10 months as a predictor of language development. *Dev Sci* **15**, 463–473 (2012).
20. Bailey, P. J. & Snowling, M. J. Auditory processing and the development of language and literacy. *Br Med Bull* **63**, 135–146 (2002).
21. Bergelson, E. & Swingley, D. At 6-9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3253–3258 (2012).
22. Morgan, L. & Wren, Y. E. A Systematic Review of the Literature on Early Vocalizations and Babbling Patterns in Young Children: *Communication Disorders Quarterly* (2018) doi:10.1177/1525740118760215.
23. Bates, E. & Hammel, E. A. *The Emergence of Symbols: Cognition and Communication in Infancy*. (Academic Press, 2015).
24. Goldfield, B. A. & Reznick, J. S. Early lexical acquisition: rate, content, and the vocabulary spurt. *J Child Lang* **17**, 171–183 (1990).
25. Fenson, L. *et al. User's Guide and Technical Manual for the MacArthur Communicative Development Inventories*. (Singular Publishing, 1993).
26. Fenson, L. *et al.* Variability in early communicative development. *Monogr Soc Res Child Dev* **59**, 1–173; discussion 174-85 (1994).
27. Hoff, E. *Language Development*. (Cengage Learning, 2013).
28. Bates, E. *Language and context: The acquisition of pragmatics*. (Academic Press, 1976).
29. Goldin-Meadow, S. Gesture as a window onto communicative abilities: Implications for diagnosis and intervention. *Perspect Lang Learn Educ* **22**, 50–60 (2015).

30. Colonnesi, C., Stams, G. J. J. M., Koster, I. & Noom, M. J. The relation between pointing and language development: A meta-analysis. *Developmental Review* **30**, 352–366 (2010).
31. Camaioni, L., Perucchini, P., Bellagamba, F. & Colonnesi, C. The Role of Declarative Pointing in Developing a Theory of Mind. *Infancy* **5**, 291–308 (2004).
32. Iverson, J. M. & Goldin-Meadow, S. Gesture paves the way for language development. *Psychol Sci* **16**, 367–371 (2005).
33. Frank, M. C., Braginsky, M., Yurovsky, D. & Marchman, V. A. Wordbank: an open repository for developmental vocabulary data. *J Child Lang* **44**, 677–694 (2017).
34. Kidd, E. Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics* **33**, 393–418 (2012).
35. Dabrowska, E. & Divjak, D. *Handbook of Cognitive Linguistics*. (De Gruyter Mouton, 2015). doi:10.1515/9783110292022.
36. O'Neill, D. K. The language use inventory for young children: a parent-report measure of pragmatic language development for 18- to 47-month-old children. *J. Speech Lang. Hear. Res.* **50**, 214–228 (2007).
37. Berman, R. A. *Language Development Across Childhood and Adolescence*. (John Benjamins Publishing, 2004).
38. Gough, P. B. & Tunmer, W. E. Decoding, Reading, and Reading Disability. *Remedial and Special Education* **7**, 6–10 (1986).
39. Jong, P. F. de & Leij, A. van der. Effects of Phonological Abilities and Linguistic Comprehension on the Development of Reading. *Scientific Studies of Reading* **6**, 51–77 (2002).
40. Hogan Tiffany P., Catts Hugh W. & Little Todd D. The Relationship Between Phonological Awareness and Reading. *Language, Speech, and Hearing Services in Schools* **36**, 285–293 (2005).
41. Edwards Jan, Beckman Mary E. & Munson Benjamin. The Interaction Between Vocabulary Size and Phonotactic Probability Effects on Children's Production Accuracy and Fluency in Nonword Repetition. *Journal of Speech, Language, and Hearing Research* **47**, 421–436 (2004).
42. Walley, A. C., Metsala, J. L. & Garlock, V. M. Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing* **16**, 5–20 (2003).
43. Dickinson, D. K. & Neuman, S. B. *Handbook of Early Literacy Research*. (Guilford Press, 2007).
44. Duff, F. J., Reen, G., Plunkett, K. & Nation, K. Do infant vocabulary skills predict school-age language and literacy outcomes? *J Child Psychol Psychiatry* **56**, 848–856 (2015).
45. Jeste, S. S. Neurodevelopmental behavioral and cognitive disorders. *Continuum (Minneap Minn)* **21**, 690–714 (2015).
46. Mastergeorge, A. M. Speech and language disorders in childhood: A neurodevelopmental perspective. in *Autism and other neurodevelopmental disorders* 257–282 (American Psychiatric Publishing, Inc., 2013).
47. Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J. & Rohde, L. A. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am J Psychiatry* **164**, 942–948 (2007).
48. Thapar, A. & Cooper, M. Attention deficit hyperactivity disorder. *The Lancet* **387**, 1240–1250 (2016).
49. Geurts, H. M. & Embrechts, M. Language profiles in ASD, SLI, and ADHD. *J Autism Dev Disord* **38**, 1931–1943 (2008).
50. Helland, W. A., Posserud, M.-B., Helland, T., Heimann, M. & Lundervold, A. J. Language Impairments in Children With ADHD and in Children With Reading Disorder. *Journal of Attention Disorders* 1087054712461530 (2012) doi:10.1177/1087054712461530.
51. Germanò, E., Gagliano, A. & Curatolo, P. Comorbidity of ADHD and dyslexia. *Dev Neuropsychol* **35**, 475–493 (2010).
52. Peyre, H. *et al.* Relationship between early language skills and the development of inattention/hyperactivity symptoms during the preschool period: Results of the EDEN mother-child cohort. *BMC Psychiatry* **16**, (2016).
53. Miniscalco, C., Nygren, G., Hagberg, B., Kadesjö, B. & Gillberg, C. Neuropsychiatric and neurodevelopmental outcome of children at age 6 and 7 years who screened positive for language problems at 30 months. *Developmental Medicine & Child Neurology* **48**, 361–366 (2006).
54. Maughan, B. & Carroll, J. Literacy and mental disorders. *Curr Opin Psychiatry* **19**, 350–354 (2006).
55. Green, B. C., Johnson, K. A. & Bretherton, L. Pragmatic language difficulties in children with hyperactivity and attention problems: an integrated review. *Int J Lang Commun Disord* **49**, 15–29 (2014).

56. Hawkins, E., Gathercole, S., Astle, D. & Holmes, J. Language Problems and ADHD Symptoms: How Specific Are the Links? *Brain Sci* **6**, (2016).
57. Mayes, S. D. & Calhoun, S. L. Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child Neuropsychol* **13**, 469–493 (2007).
58. Mayes, S. D. & Calhoun, S. L. Frequency of reading, math, and writing disabilities in children with clinical disorders. *Learning and Individual Differences* **16**, 145–157 (2006).
59. Czamara, D. *et al.* Children with ADHD Symptoms Have a Higher Risk for Reading, Spelling and Math Difficulties in the GINIplus and LISAplus Cohort Studies. *PLOS ONE* **8**, e63859 (2013).
60. Åsberg, J., Kopp, S., Berg-Kelly, K. & Gillberg, C. Reading comprehension, word decoding and spelling in girls with autism spectrum disorders (ASD) or attention-deficit/hyperactivity disorder (AD/HD): performance and predictors. *International Journal of Language & Communication Disorders* **45**, 61–71 (2010).
61. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics* **51**, 431 (2019).
62. Lord, C., Elsabbagh, M., Baird, G. & Veenstra-Vanderweele, J. Autism spectrum disorder. *The Lancet* **392**, 508–520 (2018).
63. Veness, C., Prior, M., Eadie, P., Bavin, E. & Reilly, S. Predicting autism diagnosis by 7 years of age using parent report of infant social communication skills. *J Paediatr Child Health* **50**, 693–700 (2014).
64. Mitchell, S. *et al.* Early language and communication development of infants later diagnosed with autism spectrum disorder. *J Dev Behav Pediatr* **27**, S69-78 (2006).
65. Iverson, J. M. *et al.* Early Gesture and Vocabulary Development in Infant Siblings of Children with Autism Spectrum Disorder. *J Autism Dev Disord* **48**, 55–71 (2018).
66. Tager-Flusberg, H. *et al.* Handbook of autism and pervasive developmental disorders. (2005).
67. Ozonoff, S., South, M. & Miller, J. N. DSM-IV-Defined Asperger Syndrome: Cognitive, Behavioral and Early History Differentiation from High-Functioning Autism. *Autism* **4**, 29–46 (2000).
68. Hoff, E. How social contexts support and shape language development. *Developmental Review* **26**, 55–88 (2006).
69. Schwab, J. F. & Lew-Williams, C. Language learning, socioeconomic status, and child-directed speech. *Wiley Interdiscip Rev Cogn Sci* **7**, 264–275 (2016).
70. Korf, B. R. Chapter 16 - Introduction to Human Genetics∗. in *Clinical and Translational Science (Second Edition)* (eds. Robertson, D. & Williams, G. H.) 281–311 (Academic Press, 2017). doi:10.1016/B978-0-12-802101-9.00016-8.
71. Phillips, P. C. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet* **9**, 855–867 (2008).
72. Mascheretti, S. *et al.* Complex effects of dyslexia risk factors account for ADHD traits: evidence from two independent samples. *J Child Psychol Psychiatr* n/a-n/a (2016) doi:10.1111/jcpp.12612.
73. Dionne, G., Dale, P. S., Boivin, M. & Plomin, R. Genetic Evidence for Bidirectional Effects of Early Lexical and Grammatical Development. *Child Development* **74**, 394–412 (2003).
74. Dale, P. S., Dionne, G., Eley, T. C. & Plomin, R. Lexical and grammatical development: a behavioural genetic perspective. *Journal of Child Language* **27**, 619–642 (2000).
75. Reznick, J. S., Corley, R. & Robinson, J. A longitudinal twin study of intelligence in the second year. *Monogr Soc Res Child Dev* **62**, i–vi, 1–154; discussion 155-60 (1997).
76. St Pourcain, B. *et al.* Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature communications* **5**, 4831 (2014).
77. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* **101**, 5–22 (2017).
78. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nature Genetics* 1–10 (2019) doi:10.1038/s41588-019-0481-0.
79. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* **14**, 483–495 (2013).

80. Martin, N. C., Levy, F., Pieka, J. & Hay, D. A. A Genetic Study of Attention Deficit Hyperactivity Disorder, Conduct Disorder, Oppositional Defiant Disorder and Reading Disability: Aetiological overlaps and implications. *International Journal of Disability, Development and Education* **53**, 21–34 (2006).
81. Willcutt, E. G., Pennington, B. F. & DeFries, J. C. Twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **96**, 293–301 (2000).
82. Willcutt, E. G., Pennington, B. F., Olson, R. K. & DeFries, J. C. Understanding comorbidity: A twin study of reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **144B**, 709–714 (2007).
83. Selzam, S. *et al.* Genome-Wide Polygenic Scores Predict Reading Performance Throughout the School Years. *Sci Stud Read* 1–16 (2017) doi:10.1080/10888438.2017.1299152.
84. Luciano, M. *et al.* Single Nucleotide Polymorphisms Associated with Reading Ability Show Connection to Socio-Economic Outcomes. *Behav Genet* **47**, 469–479 (2017).
85. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics* 1 (2018) doi:10.1038/s41588-018-0269-7.
86. Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am. J. Hum. Genet.* **96**, 329–339 (2015).
87. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics* **50**, 912–919 (2018).
88. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* **50**, 1112–1121 (2018).
89. Deriziotis, P. & Fisher, S. E. Speech and Language: Translating the Genome. *Trends in Genetics* **33**, 642–656 (2017).
90. Gialluisi, A. *et al.* Genome-wide association scan identifies new variants associated with a cognitive predictor of dyslexia. *Transl Psychiatry* **9**, (2019).
91. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* **20**, 567–581 (2019).
92. Choi, S. W., Mak, T. S.-H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nature Protocols* 1–14 (2020) doi:10.1038/s41596-020-0353-1.
93. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236–41 (2015).
94. St Pourcain, B. *et al.* Developmental changes within the genetic architecture of social communication behaviour: A multivariate study of genetic variance in unrelated individuals. *Biological Psychiatry* **83**, 598–606 (2017).
95. Burgess, S. *et al.* Dissecting Causal Pathways Using Mendelian Randomization with Summarized Genetic Data: Application to Age at Menarche and Risk of Breast Cancer. *Genetics* genetics.300191.2017 (2017) doi:10.1534/genetics.117.300191.

Chapter 2

The genetic architecture of language abilities in infancy and early childhood: A review

Abstract

The genetic architecture underlying language abilities in infancy and early childhood is polygenic and genetically complex. Weak to modest heritability estimates for expressive language traits, including vocabulary and grammar, have been observed both in studies of twins and unrelated individuals. These findings suggest that the majority of genetic variance that contributes to individual differences in early expressive vocabulary can be attributed to common genetic variation, as captured by commercial genotyping chips. At the polygenic level, there is evidence for both developmental genetic stability and change during language and literacy development from infancy to early adolescence. At the level of single nucleotide polymorphisms, only few significant associations are known, but the limited evidence thus far is consistent with developmental specificity. In particular, a genetic variant robustly associated with early vocabulary size showed an attenuation of association effects with respect to language and literacy skills assessed in later childhood. Current challenges for studying the genetic architecture underlying early language development concern low statistical power, pleiotropy, population-based phenomena, and the narrow phenotype definition of early language in large-scale molecular analyses.

2.1. Aims of this chapter

In this chapter, I provide an overview of studies estimating the heritability of language abilities during the first three years of life in community-based samples of twins and population-based samples of unrelated individuals. In addition, I describe evidence for genetic links between language and literacy abilities assessed during different developmental stages and discuss efforts aiming to identify single-nucleotide polymorphisms (SNPs) that contribute to variation in early language skills. Finally, I discuss current challenges for studying the genetic architecture underlying early language development.

2.2. The heritability of language abilities in infancy and early childhood

There is a large body of evidence, as reported by both twin and molecular studies, suggesting that variation in language abilities observed during the first three years of life is heritable[1–5] (Table 1). Twin studies assess heritability indirectly, based on differences in phenotypic correlations between monozygotic and dizygotic twins[6] (twin-h^2), and have reported modest contributions of genetic factors to variation in language skills assessed between 14 and 36 months of age[1–5]. A study of expressive vocabulary in 5,733 twin pairs from the UK estimated a twin-h^2 of 20% for individual differences at 24 months of age[1]. This estimate is in line with those obtained by studying fewer individuals of the same sample[2,3] (Table 1). At 36 months of age, 10%-14% of the variation in expressive vocabulary was attributable to genetic factors in these twins (N≥1,049 twin pairs)[2]. Using directly assessed genotype information from unrelated individuals, a recent meta-analysis across genome-wide association studies (GWASs) reported highly consistent results, with SNP-h^2 estimates of 13% and 14% for expressive vocabulary assessed at 15-18 months (N=8,022) and 24-30 months of age (N=9,966), respectively[1].

In addition to expressive vocabulary, genetic influences underlying early grammatical skills[2,3], a combination of expressive vocabulary and grammar[4], as well as receptive language skills[5] have been investigated in twin samples. Twin-h^2 estimates for grammatical abilities were moderate and slightly higher than the heritability of expressive vocabulary measures[2,3] (Table 1). For example, twin-h^2 of sentence complexity in two-year-old toddlers was estimated at 39% based on 2,898 twin pairs from the UK[3]. Consistent estimates were obtained studying only same-sex twin pairs of the same sample, split by birth cohort[2]. A construct for expressive language that consisted of both vocabulary and grammar, created in the same UK sample, was modestly heritable, with 28% and 25% of the variance explained by additive genetic influences at 24 and 36 months of age[4]. Finally, a study of receptive verbal abilities in

378 twin pairs from the US reported a modest heritability at 14 months (twin-h^2=28%), but no evidence for a substantial influence of genetic factors at 20 and 24 months of age[5] (Table 1). Compared to early expressive language abilities, genetic studies on receptive language skills are scarce, resulting in limited knowledge about the underlying genetic factors.

Table 1: Heritability estimates for language abilities during the first four years of life

Language ability	Sample collection	Psychological instrument	Age (months)	Heritability estimate (SE)	N (individuals)	Reference
Expressive vocabulary	Twin pairs‡	MCDI[10]	24	0.20 (0.01)	11,466	St Pourcain et al. 2014[1]
				0.25 (0.04)	5,796	Dale et al. 2000[3]
	Twin pairs‡ (same sex)	MCDI[10]	24	0.21 (0.03) 0.17 (0.03)	3,010 2,098	Dionne et al. 2003[2#]
			36	0.10 (0.03) 0.14 (0.03)	3,010 2,098	
	Unrelated individuals	CDI:WG[53], N-CDI-2A[77]	15-18	0.13 (0.05)	8,022	St Pourcain et al. 2014[1*]
		CDI:WS[54], MCDI[10], LDS[52]	24-30	0.14 (0.05)	9,966	
Expressive language	Twin pairs‡ (same sex)	Grammar and vocabulary (MCDI[10])	24	0.28 (0.08)	7,124	Hayiou-Thomas et al. 2012[4]
			36	0.25 (0.03)	6,364	
Grammar	Twin pairs‡	MCDI[10]	24	0.39 (0.07)	5,796	Dale et al. 2000[3]
	Twin pairs‡ (same sex)	MCDI[10]	24	0.42 (0.03) 0.40 (0.09)	3,010 2,098	Dionne et al. 2003[2#]
			36	0.34 (0.04) 0.29 (0.09)	3,010 2,098	
Receptive language	Twin pairs (same sex)	Construct based on SICD[78] and BSID[79]	14	0.28 (0.17)	756	Reznick et al. 1997[5]
			20	0.13 (0.16)	684	
			24	0.18 (0.16)	591	

Non-comprehensive overview of heritability estimates for language abilities assessed during the first four years of life using parental and/or observer reports. Heritability estimates were derived based on samples of twins or unrelated individuals. The reported sample size (N) represents the total number of individuals included in each analysis. ‡ Analyses based on different subsets of the Twins Early Development Study. # Heritability estimates were derived for a split sample based on birth cohort (1994 or 1995). * Heritability estimates were derived by meta-analyses across samples. Abbreviations: BSID, Bayley Scale of Infant Development; CDI:WG, Communicative Development Inventories: Words and Gestures; CDI:WS, Communicative Development Inventories: Words and Sentences; LDS, Language Development Survey; MCDI, MacArthur Communicative Development Inventory; N-CDI-2A, Dutch translation of the CDI; SICD, Sequenced Inventory of Communication Development

2.3. Genetic links among language abilities in infancy and early childhood

Developmental changes in language abilities may correspond to changes in the underlying genetic architecture. The same trait could be influenced by distinct genetic factors at different developmental stages and the trait variation explained by a certain genetic factor may vary during development[7]. There is evidence for both stability and change in the genetic factors that contribute to language abilities during early development[1–3,8]. Moderate genetic correlations among expressive vocabulary assessments between 15 and 36 months suggest developmental genetic stability[1,2]. For example, based on molecular research in 5,739 unrelated UK children, expressive vocabulary assessed at 15 months had a genetic correlation with scores at 24 months of 0.69[1]. Likewise, expressive vocabulary at 24 months shows weak to moderate genetic correlations with measurements 12 months later, with estimates ranging between 0.48 and 0.68, as reported by research of up to 1,505 UK twin pairs[2]. Evidence for genetic stability was also observed for longitudinally assessed grammar skills, at 24 and 36 months in the same twins, although genetic correlations were only weak (r_g=0.33)[2]. However, 95%-confidence intervals overlapped with genetic correlations estimated for expressive vocabulary scores that were collected at the same time for the same individuals[2], suggesting similar developmental genetic stability. Grammar and vocabulary skills may, furthermore, share underlying genetic mechanisms, as indicated by moderate to strong genetic correlations at 24 and 36 months of age (r_g=0.61-0.89)[2,3]. At two-years of age, shared genetic influences accounted for up to 28% of the observed phenotypic covariance between grammar and vocabulary[3]. These findings support the hypothesis of genetic stability and the presence of shared genetic influences during early language development. However, imperfect genetic correlations between vocabulary and grammar also suggest some developmental heterogeneity. Indeed, a twin study of the same UK sample that modelled a latent factor underlying both early vocabulary and grammar skills estimated that 3%-28% of the total variation was explained by age-specific additive genetic variance, while ~28% of the variation in expressive language skills at 24, 36 and 48 months of age could be accounted for by shared additive genetic influences[8]. However, detailed knowledge of the developmental genetic architecture underlying expressive and, especially receptive language skills, during the first few years of life is limited and solely based on twin research.

2.4. Genetic loci associated with early language

In 2014, the first meta-GWAS investigating language abilities during infancy was published[1]. This study aimed to identify SNPs related to individual differences in early

expressive vocabulary and included two developmental age-windows. For the first age-window, spanning 15 to 18 months of age, data from two independent cohorts were meta-analysed, resulting in a total sample size of 8,889 individuals[1]. This meta-GWAS provided evidence for a genome-wide significant association of rs7642482, a SNP located near *ROBO2*[1] (see below). The minor allele G, which has an allele-frequency of 0.19 in European (non-Finnish, due to linkage disequilibrium structure differences) individuals according to the GnomAD database[9] (v2.1.1), was associated with lower expressive vocabulary[1]. Variation in the sequence of rs7642482 accounted for only 0.34%-0.35% of the phenotypic variance in expressive vocabulary, consistent with the hypothesised polygenic architecture underlying language abilities, which implies the involvement of many genetic variants that each account for a very small proportion of trait variance.

To explore whether rs7642482 was related to other linguistic skills than expressive vocabulary at 15-18 months, the authors studied associations with six additional language- and literacy-related traits[1]. These included the emergence of first single-word utterances (12 months), later expressive vocabulary (24-30 months), phonological memory (8 years), verbal intelligence (9 years), reading speed (10 years), and reading comprehension (10 years) (Figure 1). There was no evidence for association across all six language and literacy skills reported, apart from nominal evidence for association of rs7642482 with reading speed (β=-0.07(SE=0.03), P=0.009, Figure 1)[1]. Furthermore, 95%-confidence intervals of the estimated effect on expressive vocabulary at 15-18 months did not overlap with associations for the other language and literacy skills, except for first single-word utterances at 12 months (Figure 1). This suggests specificity to the early phase of language acquisition, reflecting a developmental period that is characterised by a slow accumulation of single words[10,11].

Since the publication of the meta-GWAS on early vocabulary in 2014, many large-scale GWASs became publically available. To identify additional associations of rs7642482, I performed a screen across 2,986 publicly available GWASs, including basic association analyses performed within the UK Biobank sample[12]. There was little vidence for association of rs7642482 with any other trait at the phenome-wide level (P<1x10^{-5}), adjusted for the number of traits studied. The strongest association was observed for treatment with senokot 7.5mg tablet, a constipation remedy, in 337,159 individuals from the UK Biobank (P=4x10^{-5})[12]. Although this association did not meet the multiple testing threshold, it may suggest an underlying pleiotropic effect of rs7642482. Previous studies have reported an association between *ROBO2* mutations and vesicoureteral reflux[13,14], a condition that is characterised by the backward flow of urine from the bladder into the ureters of the kidney, which has been linked to constipation[15].

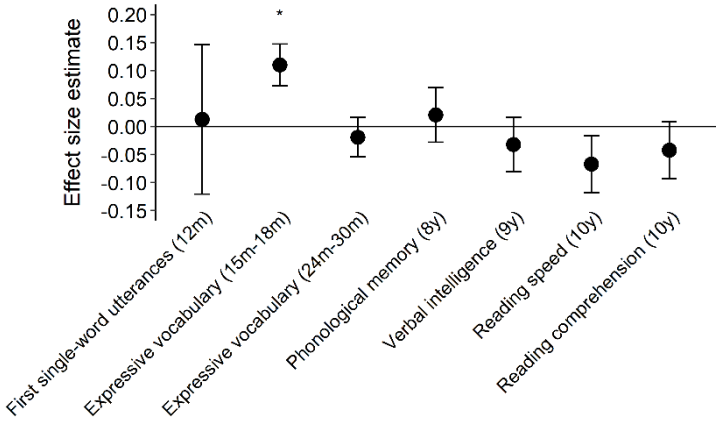


Figure 1: Associations of rs7642482 with language- and literacy-related skills. Associations of rs7642482 with language- and literacy-related skills as reported by St Pourcain et al[1]. Analysis for the emergence of first single-word utterances (12m) was based on the Northern Finnish Birth Cohort 1966 sample and the effect size estimate was derived from the odds ratio. Effect size estimates for expressive vocabulary (15m-18m and 24m-30m) were derived from the meta-analysis performed by St Pourcain et al[1]. Effect sizes for language- and literacy-related skills in mid-childhood were all based on association analyses performed within the Avon Longitudinal Sample of Parents and Children, as reported by St Pourcain et al[1]. * Evidence for association at the genome-wide significance level (P<5x10^{-8}). Abbreviations: m, months; y, years

rs7642482 is located within an intergenic region on the short arm of chromosome 3 (3p12.3), about 19 kb 3' of *ROBO2* (OMIM: 602431). *ROBO2* encodes *Homo sapiens* roundabout homologue 2 (Drosophila), an axon guidance receptor that binds to secreted SLIT ligands[16,17]. *ROBO2* is highly expressed in the human central nervous system, across different brain regions, but also in other tissues, such as the lung, according to the Genotype-Tissue Expression project [18] (GTEx, v8). During the course of development, its expression peaks in the first trimester[1]. Robo receptors and SLIT ligands are highly conserved from fly to human[17,19,20] and Genomic Evolutionary Rate Profiling[21] scores indicated that also the sequence of rs7642482 is highly conserved, with values above three[1]. In both Drosophila and mice Robo2 cooperates with Robo1, encoded by its neighbouring gene *ROBO1,* in axon guidance during brain development[17,22]. Recently, a study showed that selective silencing of Robo1 and Robo2 expression in pre-cerebellar neurons in mice did not lead to migration anomalies, suggesting non-cell autonomous effects[23].

Although there was no evidence that rs7642482 related to protein-coding variation in *ROBO2* in the initial publication[1], analyses of gene expression and regulatory chromatin states indicated that variation at rs7642482 might be related to regulatory mechanisms in embryonic cell types[1]. However, no *cis* expression quantitative trait loci (eQTL) effects within ±1 Mb in postnatal derived cell types or adult brain tissue were reported for rs7642482[1]. To further investigate potential eQTL effects of rs7642482, I carried out additional analyses based on the latest SNP-gene expression data sets as included in PhenoScanner[24,25] (v2) and available via the Genotype-Tissue Expression

project[18] (GTEx, v8). Studying association with gene expression across a variety of tissues, including multiple brain regions, did not provide evidence for eQTL effects, consistent with the initial report[1]. Finally, I did not find evidence for eQTL effects of rs7642482 in fetal human brain tissue ($P \geq 0.09$) based on gene expression in brain tissue derived from the second trimester of gestation (N=120)[26].

Previous studies of dyslexia and related traits provide additional support for potential roles of *ROBO2* in language and communication processes. The 3p12-p13 region has been linked to quantitative dyslexia traits[27], speech-sound disorder traits and reading abilities[28]. Although no coding variants in *ROBO2* have been reported to be associated with developmental dyslexia[29], *ROBO1* is considered as a candidate gene for this disorder, based initially on investigations of co-segregating alleles in a multigenerational family[29,30]. Common genetic variants located within *ROBO1* have been associated with reading disability[31] and non-word repetition[32]. However, the implicated genetic variants have failed to replicate in recent GWAS studies of dyslexia-related traits[33] and no evidence was observed for a link between early expressive vocabulary and *ROBO1* SNPs[1].

A meta-analysis of expressive vocabulary assessed between 24 and 30 months, based on 10,819 individuals, did not result in the identification of associated SNPs at the genome-wide significant level[1]. At time of writing, no GWAS on early receptive vocabulary or grammatical skills has been published.

2.5. Genetic links with later language- and literacy-related abilities

Compared to the heritability estimates reported for language abilities during the first three years of life (see section 2.1 and Table 1), the contribution of genetic influences to language and literacy skills assessed from mid-childhood to adolescence is larger, with twin-h^2 estimates of 47%-72% estimated based on up to 7,179 UK twin pairs[4,8]. This increase in heritability during development has been reported for many cognitive skills[34–36], and may involve different processes that are referred to as genetic innovation and amplification[37]. Innovation involves an increase in heritability due to previously unrelated genetic variation that becomes associated with a trait over time[34], implying developmental genetic heterogeneity. In contrast, amplification refers to an increase in heritability due to the same genetic variation that explains more phenotypic variation over time[34], consistent with genetic stability.

A UK twin study examined this developmental paradigm by investigating genetic links between early expressive language (24-48 months) and skills related to vocabulary, semantics, syntax and pragmatics (7-12 years). They reported moderate genetic correlations (r_g=0.46-0.54), which could explain only about a third of the phenotypic relationship between early and middle childhood language latent factors[4]. A similar

developmental pattern was observed in the same sample for early expressive language and mid-childhood reading abilities[8]. These findings suggested that processes of innovation, rather than amplification, may account for the observed increase in heritability for language and literacy abilities during the transition from early to middle childhood[4]. The developmental origins of these stable genetic factors are, however, little understood beyond latent factor twin analyses[4,8].

The meta-GWAS on early expressive vocabulary[1] found little support for association of rs7642482 with language and literacy skills later in life, except for a nominal association with reading speed at 10 years of age[1] (Figure 1). Despite several GWAS efforts aiming to identify SNPs related to language- and/or reading-related abilities assessed from childhood to adolescence[33,38–43], evidence for association at the genome-wide significance level has so far only been reported for measures of rapid automatised naming[33,43]. In rapid automatised naming tasks participants are asked to name visually presented items as quickly and accurately as possible. This ability is related to reading performance, and rapid automatised naming skills in kindergarten are predictive of reading fluency throughout elementary school[44]. A genome-wide association screen analysing rapid automatised naming of letters in 2,563 individuals of European descent reported association with rs17663182 (P=4.7x10^{-9})[33]. This variant is located within the non-coding *micro-RNA 924 host gene* (*MIR924HG*) for which currently no regulatory role is known[33]. Multivariate analyses of rapid automatised naming across letters, objects and numbers in 1,331 Hispanic American and African-American youth identified an association with rs1555839 (P=2.2x10^{-8}), located in an intergenic region near the long non-coding RNA *ribosomal protein L7 pseudogene 34* (*RPL7P34*) that has an unknown function[43]. This latter association signal was also related to measures of word reading[43]. However, neither variant was among the top associated signals (P<1x10^{-4}) for early expressive vocabulary (15-18 months) or later expressive vocabulary (24-30 months)[1], suggesting that their effects on expressive vocabulary during early development are limited. Larger and more powerful samples are warranted to deepen the knowledge of genetic mechanisms underlying early language performance.

2.6. Challenges for studying the genetic architecture underlying early language development

Statistical power

One of the major challenges for studying the genetic architecture underlying early language development concerns statistical power. This issue relates to both the low heritability of early language skills (Table 1) and the sample size of current large-scale efforts. Assuming SNP-h^2 of 13% for expressive vocabulary, as observed in previous efforts[1] (Table 1), individual-level genotype data from at least 6,900 unrelated individuals

is required for 80% study power to detect evidence for an underlying genetic variance component[45]. For methods that estimate heritability based on summary statistics, far larger sample sizes are needed, as these methods generally have larger standard errors than methods based on individual-level genotype data[46,47]. The summary statistics derived from the meta-GWAS on early expressive vocabulary[1] are based on 8,889 and 10,819 children for expressive vocabulary size at 15-18 months and 24-30 months, respectively. Still, their power to detect genetic association at the genome-wide significant level is only moderate, with 39% and 60% (assuming an additive model and an increaser allele frequency of 0.1 explaining 0.3% trait variation, with complete LD with marker and genetic risk variant)[48].

The application of multivariate methodologies is an approach to increase statistical power without the need to increase the absolute sample size[47]. There is a wide range of methods available aiming to combine GWAS summary statistics of correlated traits to improve the power of detecting SNP-trait associations (for review see e.g. [49,50]). Among them, multi-trait analysis of GWAS (MTAG) has been introduced, a method that takes advantage of the genetic relationships among traits and provides a generalised estimate of inverse-variance-weighted meta-analysis by integrating GWAS summary statistics of different traits, while allowing for overlapping samples[51]. The increase in statistical power of multi-trait analyses compared to single-trait analyses can be interpreted in terms of increase in sample size for single-trait association studies. For example, analyses of three genetically correlated traits (r_g~0.70), based on samples with overlapping individuals, resulted in an increase of up to 55% in the single-trait association sample size[51]. Considering the moderate genetic correlation patterns among language abilities during early development, including both vocabulary and grammar (see section 2.2), multivariate analyses could be a relatively easy way to increase the statistical power.

Assessing language abilties from infancy to early childhood

Language abilities in large samples of young children, that can be used to study the role of common genetic variation, are often assessed using parental questionnaires such as the MacArthur Communicative Development Inventory (CDI)[10] or Language Development Survey (LDS)[52]. The LDS was developed to identify language delays in children of 24 months and only evaluates expressive language[52]. CDI forms were developed to assess language acquisition in typically developing children[10], with the MacArthur CDI:Words & Gestures[53] (CDI-WG) being created for children between 8 and 18 months, whereas the MacArthur CDI:Words & Sentences[54] (CDI-WS) can be used to capture language development in children aged between 16 and 30 months. Although parental judgements might be associated with higher random error rates compared to direct assessments, affecting heritability estimations[55], moderate to strong correlations between parental judgements and direct assessments of a child's vocabulary suggest

that parent reports have sufficient instrument validity[56,57]. However, assessments of early language are often restricted to expressive vocabulary only. While CDI-WG includes an assessment of receptive vocabulary, both the CDI-WS and LDS do not. In general, parents are thought to be poorer at judging their child's language comprehension compared to language production[58], as assessing receptive language skills requires that parents notice their children's non-verbal responses to words and is therefore deemed more subjective than assessing expressive language skills. Indeed, comparison of parent-reported receptive vocabulary size in children between 12 and 24 months of age, using an adapted form of the CDI-WG measuring infant performance with a preferential looking task, showed that parents tend to underestimate their children's receptive vocabulary[59]. Furthermore, the CDI-WS excluded the receptive scale as it was thought to be too complex for parents to accurately assess this at times of rapid vocabulary growth. However, a study of 25-month-old children showed that parents are able to reliably assess receptive vocabulary, with a correlation of 0.55 between parent report and child task performance, highlighting the feasibility of studies investigating early receptive language skills[56].

Potential sources of bias in genetic associations

Estimations of genetic associations in population-based samples of unrelated individuals, both heritability and genetic correlations, may be biased due to violations of methodological assumptions[60,61]. This includes population phenomena such as dynastic effects, assortative mating and population stratification[62,63]. Dynastic effects imply that genetic variants have an influence on offspring phenotype due to their effects on parental phenotypes that shape children's environment[62]. For example, if the number of books available in a household is influenced by genetic variants associated with increased education in the parents, a child that grows up in an environment that stimulates language development is also likely to inherit education-associated variants. In that case, a child's language development is influenced by both genetic variants related to education as well as a stimulating environment, which is influenced by parental phenotype and consequently parental genetic variation. The latter is a specific form of gene-environment correlation and may result in inflated heritability and genetic correlation estimates[62], especially for cognition- and socioeconomic-related traits[62,63].

Other population phenomena that may bias genetic associations include assortative mating and population stratification[62]. Assortative mating refers to mate selection based on phenotypic characteristics, resulting in non-random pairing of spouses. If there is a genetic component underlying these phenotypes, this results in an increased genetic similarity for spouses compared to the general population, leading to inflated heritability estimates[64]. Population stratification refers to the induction of associations between genotype and phenotype due to systematic differences in allele frequencies between populations arising from different ancestries. Studies of siblings or

families are robust to this effect and could be used to assess the influence of population stratification. However, as the phenotypic and genetic similarity among relatives is increased compared to unrelated individuals, such studies require even larger samples to achieve sufficient statistical power[62]. Currently, researchers adjust for population stratification using principal components, designed to capture ancestry-related genetic differences, thus, minimising bias[62].

Pleiotropy

Genetic correlations are abundant across a variety of human traits[65], including psychiatric disorders and brain phenotypes[66] and reflect pleiotropic effects at the polygenic level. Pleiotropy may arise due to different underlying mechanisms (Figure 2) and can be defined at different genetic levels, including genetic variants or genes. Biological pleiotropy encompasses mechanisms where genetic variants have a direct biological influence on more than one phenotype (Figure 2a). Mediated pleiotropy refers to the indirect association between a genetic variant and a further phenotype that arises due to causal associations between the two phenotypes[67] (Figure 2b), while spurious pleiotropy involves multiple sources of bias that cause a false association between a genetic variant or gene and multiple phenotypes[67]. For example, different causal risk variants in high LD could be located in different genes and affect different phenotypes, but will be, due to their LD structure, captured by the same GWAS marker (known as co-localising variants)[67] (Figure 2c). Furthermore, two phenotypes might be genetically related to each other due to a third common factor[67] (Figure 2d). For example, twin studies on reading deficits and ADHD reported evidence for bivariate heritabilities up to 95%[68–70], suggesting a role for pleiotropy in the co-occurence of reading impairments and ADHD. However, both ADHD and reading abilities are also genetically related to educational attainment[71–74], which could thus reflect shared genetic associations with a third common factor that genetically links reading skills to ADHD. Consequently, the identification of causal relationships between phenotypes using genetic analyses is challenging, especially as the assumptions of causal modelling approaches, such as Mendelian Randomization, are often not fulfilled[75]. Thus, I interpreted in this thesis relationships between phenotypes due to shared genetic variation as genetic associations only. However, even genetic correlation patterns between phenotypes can be complex, as genetic correlations describe the average pleiotropic effect across the genome[47]. Consequently, genetic correlation patters at the regional level may deviate from the average pattern, and strong positive and negative regional correlations have been observed even in the absence of genome-wide genetic correlations[76]. Understanding the complex genetic mechanisms underlying trait overlap may lead to a better understanding of shared biological mechanisms[47].

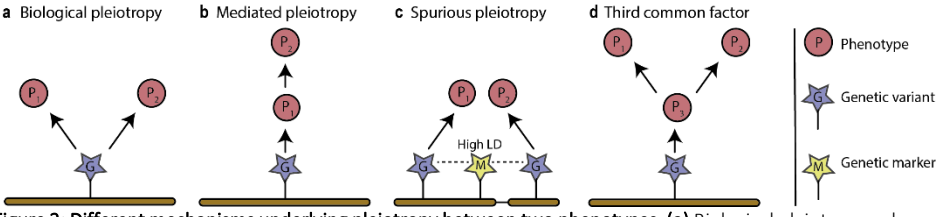


Figure 2: Different mechanisms underlying pleiotropy between two phenotypes. (a) Biological pleiotropy, where the same genetic factor has a direct influence on two phenotypes. **(b)** Mediated pleiotropy reflects a causal relationship between two phenotypes. **(c)** Spurious pleiotropy is attributable to bias. For example, a single genetic marker tags two different genetic factors (in different genes) that each relate to a different phenotype through co-localisation. **(d)** Pleiotropy between two phenotypes due to a third common factor. Abbreviations: LD, linkage disequilibrium.

2.7. Conclusions

The heritability of expressive language abilities assessed from infancy to early childhood is low to moderate, indicating that genetic factors, including common variation, play a role in the observed individual differences during language acquisition processes. During the course of development, there is evidence for both developmental stability and heterogeneity in the genetic factors underlying language and literacy skills. However, knowledge of the composition of shared and specific genetic factors underlying language and literacy development is limited, beyond findings from latent factor twin analyses. Specifically, the role of receptive vocabulary and the emergence of genetic links between vocabulary skills assessed from infancy to early childhood and later language and literacy performance, has not yet been characterised based on genome-wide information. Finally, low statistical power, the identification of pleiotropic mechanisms and bias through population phenomena are challenges that future studies will need to address.

References

1. St Pourcain, B. *et al.* Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature communications* **5**, 4831 (2014).
2. Dionne, G., Dale, P. S., Boivin, M. & Plomin, R. Genetic Evidence for Bidirectional Effects of Early Lexical and Grammatical Development. *Child Development* **74**, 394–412 (2003).
3. Dale, P. S., Dionne, G., Eley, T. C. & Plomin, R. Lexical and grammatical development: a behavioural genetic perspective. *Journal of Child Language* **27**, 619–642 (2000).
4. Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev Sci* **15**, 233–49 (2012).
5. Reznick, J. S., Corley, R. & Robinson, J. A longitudinal twin study of intelligence in the second year. *Monogr Soc Res Child Dev* **62**, i–vi, 1–154; discussion 155-60 (1997).
6. Mayhew, A. J. & Meyre, D. Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities. *Curr Genomics* **18**, 332–340 (2017).
7. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era — concepts and misconceptions. *Nature Reviews Genetics* **9**, 255–266 (2008).
8. Harlaar, N., Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. Why Do Preschool Language Abilities Correlate With Later Reading? A Twin Study. *J Speech Lang Hear Res* **51**, 688–705 (2008).
9. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
10. Fenson, L. *et al.* Variability in early communicative development. *Monogr Soc Res Child Dev* **59**, 1–173; discussion 174-85 (1994).
11. Clark, E. V. *First language acquisition*. (Cambridge University Press, 2016).
12. Howrigan, D. Details and considerations of the UK Biobank GWAS. *Details and considerations of the UK Biobank GWAS* http://www.nealelab.is/blog/2017/9/11/details-and-considerations-of-the-uk-biobank-gwas.
13. Bertoli-Avella, A. M. *et al.* ROBO2 Gene Variants Are Associated with Familial Vesicoureteral Reflux. *J Am Soc Nephrol* **19**, 825–831 (2008).
14. Lu, W. *et al.* Disruption of ROBO2 Is Associated with Urinary Tract Anomalies and Confers Risk of Vesicoureteral Reflux. *Am J Hum Genet* **80**, 616–632 (2007).
15. O'Regan, S., Schick, E., Hamburger, B. & Yazbeck, S. Constipation associated with vesicoureteral reflux. *Urology* **28**, 394–396 (1986).
16. Fricke, C., Lee, J.-S., Geiger-Rudolph, S., Bonhoeffer, F. & Chien, C.-B. astray, a Zebrafish roundabout Homolog Required for Retinal Axon Guidance. *Science* **292**, 507–510 (2001).
17. Simpson, J. H., Kidd, T., Bland, K. S. & Goodman, C. S. Short-Range and Long-Range Guidance by Slit and Its Robo Receptors: Robo and Robo2 Play Distinct Roles in Midline Guidance. *Neuron* **28**, 753–766 (2000).
18. Aguet, F. *et al.* The GTEx Consortium atlas of genetic regulatory effects across human tissues. *bioRxiv* 787903 (2019) doi:10.1101/787903.
19. Seeger, M., Tear, G., Ferres-Marco, D. & Goodman, C. S. Mutations affecting growth cone guidance in Drosophila: genes necessary for guidance toward or away from the midline. *Neuron* **10**, 409–426 (1993).
20. Kidd, T. *et al.* Roundabout controls axon crossing of the CNS midline and defines a novel subfamily of evolutionarily conserved guidance receptors. *Cell* **92**, 205–215 (1998).
21. Davydov, E. V. *et al.* Identifying a High Fraction of the Human Genome to be under Selective Constraint Using GERP++. *PLOS Computational Biology* **6**, e1001025 (2010).
22. López-Bendito, G. *et al.* Robo1 and Robo2 cooperate to control the guidance of major axonal tracts in the mammalian forebrain. *J. Neurosci.* **27**, 3395–3407 (2007).
23. Dominici, C., Rappeneau, Q., Zelina, P., Fouquet, S. & Chédotal, A. Non-cell autonomous control of precerebellar neuron migration by Slit and Robo proteins. *Development* **145**, (2018).
24. Kamat, M. A. *et al.* PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).
25. Staley, J. R. *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* **32**, 3207–3209 (2016).

26. O'Brien, H. E. *et al.* Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. *Genome Biol* **19**, (2018).
27. Fisher, S. E. *et al.* Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nat. Genet.* **30**, 86–91 (2002).
28. Stein, C. M. *et al.* Pleiotropic Effects of a Chromosome 3 Locus on Speech-Sound Disorder and Reading. *Am J Hum Genet* **74**, 283–297 (2004).
29. Massinen, S. *et al.* Genomic sequencing of a dyslexia susceptibility haplotype encompassing ROBO1. *Journal of Neurodevelopmental Disorders* **8**, 4 (2016).
30. Hannula-Jouppi, K. *et al.* The Axon Guidance Receptor Gene ROBO1 Is a Candidate Gene for Developmental Dyslexia. *PLOS Genetics* **1**, e50 (2005).
31. Tran, C. *et al.* Association of the ROBO1 gene with reading disabilities in a family-based analysis. *Genes Brain Behav.* **13**, 430–438 (2014).
32. Bates, T. C. *et al.* Genetic variance in a component of the language acquisition device: ROBO1 polymorphisms associated with phonological buffer deficits. *Behav. Genet.* **41**, 50–57 (2011).
33. Gialluisi, A. *et al.* Genome-wide association scan identifies new variants associated with a cognitive predictor of dyslexia. *Transl Psychiatry* **9**, (2019).
34. Briley, D. A. & Tucker-Drob, E. M. Explaining the increasing heritability of cognitive ability across development: a meta-analysis of longitudinal twin and adoption studies. *Psychol Sci* **24**, 1704–1713 (2013).
35. Haworth, C. *et al.* The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* **15**, 1112–1120 (2010).
36. Plomin, R. & Deary, I. J. Genetics and intelligence differences: five special findings. *Mol Psychiatry* **20**, 98–108 (2015).
37. Plomin, R. & DeFries, J. C. The origins of individual differences in infancy; the Colorado adoption project. *Science* **230**, 1369+ (1985).
38. Eicher, J. D. *et al.* Genome-wide association study of shared components of reading disability and language impairment. *Genes Brain Behav* **12**, 792–801 (2013).
39. Gialluisi, A. *et al.* Genome-wide screening for DNA variants associated with reading and language traits. *Genes Brain Behav* **13**, 686–701 (2014).
40. Luciano, M. *et al.* A genome-wide association study for reading and language abilities in two population cohorts. *Genes Brain Behav.* **12**, 645–652 (2013).
41. Field, L. L. *et al.* Dense-map genome scan for dyslexia supports loci at 4q13, 16p12, 17q22; suggests novel locus at 7q36. *Genes Brain Behav.* **12**, 56–69 (2013).
42. Meaburn, E. L., Harlaar, N., Craig, I. W., Schalkwyk, L. C. & Plomin, R. Quantitative trait locus association scan of early reading disability and ability using pooled DNA and 100K SNP microarrays in a sample of 5760 children. *Mol. Psychiatry* **13**, 729–740 (2008).
43. Truong, D. T. *et al.* Multivariate genome-wide association study of rapid automatised naming and rapid alternating stimulus in Hispanic American and African-American youth. *J. Med. Genet.* **56**, 557–566 (2019).
44. Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D. & Foorman, B. R. Kindergarten Prediction of Reading Skills: A Longitudinal Comparative Analysis. *Journal of Educational Psychology* **96**, 265–282 (2004).
45. Visscher, P. M. *et al.* Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples. *PLOS Genetics* **10**, e1004269 (2014).
46. Ni, G., Moser, G., Wray, N. R. & Lee, S. H. Estimation of Genetic Correlation via Linkage Disequilibrium Score Regression and Genomic Restricted Maximum Likelihood. *Am J Hum Genet* **102**, 1185–1194 (2018).
47. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* **20**, 567–581 (2019).
48. Purcell, S., Cherny, S. S. & Sham, P. C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003).
49. Porter, H. F. & O'Reilly, P. F. Multivariate simulation framework reveals performance of multi-trait GWAS methods. *Scientific Reports* **7**, 38837 (2017).

50. Shriner, D. Moving toward System Genetics through Multiple Trait Analysis in Genome-Wide Association Studies. *Front Genet* **3**, 1 (2012).
51. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics* **50**, 229–237 (2018).
52. Rescorla Leslie. The Language Development Survey. *Journal of Speech and Hearing Disorders* **54**, 587–599 (1989).
53. Fenson, L. *et al.* Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics* **21**, 95–116 (2000).
54. Fenson, L. *et al. User's Guide and Technical Manual for the MacArthur Communicative Development Inventories*. (Singular Publishing, 1993).
55. Walters, R., Churchhouse, C. & Hosking, R. Heritability 101: What is 'heritability'? http://www.nealelab.is/blog/2017/9/13/heritability-101-what-is-heritability (2019).
56. Ring, E. D. & Fenson, L. The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language* **20**, 141–159 (2000).
57. Sachse, S. & Von Suchodoletz, W. Early identification of language delay by direct language assessment or parent report? *J Dev Behav Pediatr* **29**, 34–41 (2008).
58. Tomasello, M. & Mervis, C. B. The Instrument Is Great, but Measuring Comprehension Is Still a Problem. *Monographs of the Society for Research in Child Development* **59**, 174–179 (1994).
59. Houston-Price, C., Mather, E. & Sakkalou, E. Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *J Child Lang* **34**, 701–724 (2007).
60. Evans, L. M. *et al.* Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nature Genetics* **50**, 737–745 (2018).
61. Young, A. I. *et al.* Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics* **50**, 1304–1310 (2018).
62. Morris, T. T., Davies, N. M., Hemani, G. & Smith, G. D. Why are education, socioeconomic position and intelligence genetically correlated? *bioRxiv* 630426 (2019) doi:10.1101/630426.
63. Selzam, S. *et al.* Comparing Within- and Between-Family Polygenic Score Prediction. *The American Journal of Human Genetics* **105**, 351–363 (2019).
64. Falconer, D. S. & Mackay, T. F. C. *Quantitative genetics*. vol. fourth (Pearson, 1996).
65. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236–41 (2015).
66. Brainstorm Consortium *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**, (2018).
67. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* **14**, 483–495 (2013).
68. Martin, N. C., Levy, F., Pieka, J. & Hay, D. A. A Genetic Study of Attention Deficit Hyperactivity Disorder, Conduct Disorder, Oppositional Defiant Disorder and Reading Disability: Aetiological overlaps and implications. *International Journal of Disability, Development and Education* **53**, 21–34 (2006).
69. Willcutt, E. G., Pennington, B. F. & DeFries, J. C. Twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **96**, 293–301 (2000).
70. Willcutt, E. G., Pennington, B. F., Olson, R. K. & DeFries, J. C. Understanding comorbidity: A twin study of reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **144B**, 709–714 (2007).
71. Selzam, S. *et al.* Genome-Wide Polygenic Scores Predict Reading Performance Throughout the School Years. *Sci Stud Read* 1–16 (2017) doi:10.1080/10888438.2017.1299152.
72. Luciano, M. *et al.* Single Nucleotide Polymorphisms Associated with Reading Ability Show Connection to Socio-Economic Outcomes. *Behav Genet* **47**, 469–479 (2017).
73. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics* 1 (2018) doi:10.1038/s41588-018-0269-7.
74. Gialluisi, A. et al. Genome-wide association study reveals new insights into the heritability and genetic correlates of developmental dyslexia. Mol. Psychiatry 1–14 (2020) doi:10.1038/s41380-020-00898-x.
75. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* **23**, R89–R98 (2014).

76. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *Am. J. Hum. Genet.* **101**, 737–751 (2017).
77. Zink, I. & Lejaegere, M. N-CDI's: korte vormen, Aanpassing en hernormering van de MacArthur Short Form Vocabulary Checklist van Fenson et al.
78. Hedrick, D. L., Prather, E. M. & Tobin, A. R. *Sequenced Inventory of Communication Development*. (University of Washington Press).
79. Bayley, N. & Psychological Corporation. *Bayley scales of infant development.* (Psychological Corp., 1969).

Chapter 3

The developmental origins of genetic factors influencing language and literacy: Associations with early-childhood vocabulary

This chapter has been published as:

Verhoef, E., Shapland, C.Y., Fisher, S.E., Dale, P.S., St Pourcain, B. The developmental origins of genetic factors influencing language and literacy: Associations with early-childhood vocabulary. *Journal of Child Psychology and Psychiatry* (2020, in press). doi: 10.1111/jcpp.13327

Abstract

Background: The heritability of language and literacy skills increases from early childhood to adolescence. The underlying mechanisms are little understood, and may involve (i) the amplification of genetic influences contributing to early language abilities and/or (ii) the emergence of novel genetic factors (innovation). Here, we investigate the developmental origins of genetic factors influencing mid-childhood/early-adolescent language and literacy. We evaluate evidence for the amplification of early-childhood genetic factors for vocabulary, in addition to genetic innovation processes.

Methods**:** Expressive and receptive vocabulary scores at 38 months, thirteen language- and literacy-related abilities and non-verbal cognition (7-13 years) were assessed in unrelated children from the Avon Longitudinal Study of Parents and Children (ALSPAC, $N_{individuals} \leq 6{,}092$). We investigated the multivariate genetic architecture underlying early-childhood expressive and receptive vocabulary, and each of 14 mid-childhood/early-adolescent language, literacy or cognitive skills with trivariate structural equation (Cholesky) models as captured by genome-wide genetic relationship matrices. The individual path coefficients of the resulting structural models were finally meta-analysed to evaluate evidence for overarching patterns.

Results**:** We observed little support for the emergence of novel genetic sources for language, literacy or cognitive abilities during mid-childhood or early adolescence. Instead, genetic factors of early-childhood vocabulary, especially those unique to receptive skills, were amplified and represented the majority of genetic variance underlying many of these later complex skills (≤99%). The most predictive early genetic factor accounted for 29.4%(SE=12.9%) to 45.1%(SE=7.6%) of the phenotypic variation in verbal intelligence and literacy skills, but also for 25.7%(SE=6.4%) in performance intelligence, while explaining only a fraction of the phenotypic variation in receptive vocabulary (3.9%(SE=1.8%)).

Conclusions**:** Genetic factors contributing to many complex skills during mid-childhood and early adolescence, including literacy, verbal cognition and non-verbal cognition, originate developmentally in early childhood and are captured by receptive vocabulary. This suggests developmental genetic stability and overarching aetiological mechanisms.

Keywords: ALSPAC, behavioural genetics, language and literacy development

3.1. Introduction

Individual differences in vocabulary during the preschool period are predictive of many later language- and literacy-related skills[1–4], an important component of academic achievement[5]. For example, a latent factor consisting of expressive and receptive vocabulary size at 16-24 months predicted vocabulary size, as well as performance on tests of phonological awareness, reading accuracy and reading comprehension in children five years later[3]. Similarly, infants with a larger expressive vocabulary at 24 months showed a larger vocabulary as well as better decoding, word recognition, and passage comprehension skills when assessed up to primary school[4].

Associations between infant vocabulary and language and literacy skills during later life may arise due to shared underlying aetiologies. According to the "simple view of reading" theory, reading comprehension is the product of printed word recognition (decoding) and oral language comprehension[6]. Early vocabulary is a central component of both these abilities[7]. Decoding is substantially based on phonological awareness (i.e. the awareness of sound structures of speech), which develops during the preschool period and has been shown to be related to vocabulary size[7]. Listening comprehension (i.e. the understanding of spoken language), particularly bottom-up processing, necessarily begins with vocabulary comprehension[8]. Spelling performance is also closely related to phonological awareness and other phonological abilities[9]. However, the biological processes that underlie these complex developmental interrelationships are only partially understood.

Variation in expressive and receptive language skills, assessed during the first four years of life, is modestly heritable, while genetic influences on language and literacy skills assessed from mid-childhood to early adolescence are moderate to strong[10–13]. Specifically, longitudinal twin studies, assessing heritability indirectly, i.e. based on differences in phenotypic correlations between monozygotic and dizygotic twins (twin-h^2)[14], have reported heritability estimates of 22%-28% for a combined language measure including expressive vocabulary at 2, 3 and 4 years of age[11]. A considerable part of this twin-h^2 can be attributed to common genetic variation, as estimated from directly assessed genotype information in population-based samples of unrelated children. Single-nucleotide polymorphism (SNP)-h^2 estimates range between 13% and 14% for expressive vocabulary at 15-18 and 24-30 months of age respectively[12]. In contrast, the heritability for language and literacy skills assessed from mid-childhood onwards is larger, with twin-h^2 estimates from 47% to 72% [10,11] and SNP-h^2 estimates from 32% to 54%[13]. Twin-based genetic correlations, reflecting the extent to which genetic variation is shared between two traits, are moderate between early childhood and later developmental stages and imply some genetic stability[10,11].

The increase in heritability from early childhood to adolescence has been reported for many cognitive skills[15,16], suggesting overarching aetiological mechanisms

that may involve processes of genetic innovation and amplification [17]. Innovation refers to novel genetic factors emerging during development (i.e. previously unrelated genetic variation becomes associated with a trait over time). In contrast, amplification refers to genetic influences that are associated with a trait throughout development, explaining increasingly more variation with progressing age[15]. A meta-analysis of twin studies on cognitive abilities suggested that novel genetic influences predominate during the transition from early to middle childhood, supporting innovation[15]. From eight years of age onwards (mid-childhood), there was evidence for enhanced genetic stability with dominance of amplification processes[15]. A similar pattern of amplification and innovation processes was reported by twin studies examining genetic links between early language (including expressive vocabulary and syntax skills between 2-4 years of age) and both mid-childhood/adolescent language[11] and reading abilities[10], based on latent factor models. Thus, innovation and to a lesser degree amplification processes may account for the observed increase in heritability of language and literacy skills during the transition from early to mid-childhood.

Beyond latent factor twin analyses[10,11], the developmental origins of genetic variation contributing to child and adolescent language, literacy and cognition are little characterised. In particular, genetic relationships with early-childhood receptive vocabulary are unknown and the spectrum of interrelated later-life skills that are genetically related to early-childhood language abilities is only partially understood. Furthermore, evidence for amplification and innovation processes has not yet been established beyond twin research. Here, we use SNP information from directly genotyped markers and structural equation models to seek evidence for innovation and/or amplification processes during language and literacy development within a sample of unrelated children from the Avon Longitudinal Study of Parents And Children (ALSPAC, N≤6,092). Specifically, we study expressive and receptive vocabulary at 38 months and a wide range of mid-childhood/early-adolescent language- and literacy-related skills, including reading, spelling, phonemic awareness, listening comprehension, non-word repetition and verbal intelligence, as well as non-verbal intelligence (7-13 years).

3.2. Methods

Participants

All participants were drawn from ALSPAC, a UK population-based longitudinal pregnancy-ascertained birth cohort (estimated birth date: 1991-1992, Appendix S1)[18,19]. The ALSPAC Ethics and Law Committee and the Local Research Ethics Committees provided ethical approval for the study. Consent for biological samples has been collected in accordance with the Human Tissue Act (2004). Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

ALSPAC participants were genotyped using the Illumina HumanHap550 quad chip genotyping platforms. Standard genomic quality control was performed using PLINK (v1.07)[20] (Appendix S2). After quality control, 465,740 SNPs and ≤6,092 individuals with high quality genetic and phenotypic data remained.

Measures

Early-childhood vocabulary: Expressive and receptive vocabulary were assessed at 38 months using parent reports and age-specific defined word lists adapted from the MacArthur Communicative Development Inventory Words & Sentences (CDI)[21]. Parents were asked whether their child was able to (i) say, (ii) understand or (iii) both say and understand a word from a list of 123 words. Expressive vocabulary size reflects the number of words a child produces, regardless of whether they also understand these words and was defined as the sum of words a child (i) says and (iii) says and understands. Receptive vocabulary size reflects the number of words a child understands, regardless of whether children are able to produce these words and was defined as the sum of words a child (ii) understands and (iii) says and understands. CDI expressive vocabulary scores have high validity, showing correlations with direct assessments of over 0.70[22,23]. The correlation between parental and direct assessment of receptive vocabulary is 0.55[23]. In total, 6,092 children had both early vocabulary and genome-wide genetic data available (Table 1).

Mid-childhood/early-adolescent language- and literacy-related abilities: Thirteen language- and literacy-related abilities (LRAs) capturing reading, spelling, phonemic awareness, listening comprehension, non-word repetition and verbal intelligence were assessed from mid-childhood to early adolescence (7-13 year, N≤5,749) using both standardised and ALSPAC-specific instruments (Table 1, Appendix S3). Word reading accuracy and comprehension (age 7 years) was measured using the basic reading subtest of the Wechsler Objective Reading Dimensions (WORD) assessment. Word and non-word reading accuracy scores were assessed using an ALSPAC-specific measure (Appendix S3), in addition to passage reading accuracy and speed with the revised Neale Analysis of Reading Ability (NARA II), all at age 9 years. Word and non-word reading

speed (age 13 years) were captured with the Test of Word Reading Efficiency (TOWRE). Spelling accuracy (age 7 and 9 years) was assessed with an ALSPAC-specific measure (Appendix S3). Phonemic awareness (age 7 years) was measured with the Auditory Analysis Test (AAT) and listening comprehension, non-word repetition and verbal intelligence quotient (VIQ) scores (all age 8 years) were assessed with a subset of the Wechsler Objective Language Dimensions (WOLD) test, an adaptation of the Children's Test of Nonword Repetition (CNRep) and the Wechsler Intelligence Scale for Children (WISC-III) respectively. A detailed description of each instrument, including reliability, validity and references, is available in Table 1 and Appendix S3.

Table 1: Language, literacy and cognitive abilities in ALSPAC children

Measure	Mean Score (SE)	Score range	Mean Age (SE) in years	N (%males)
Expressive voc 38m (CDI)	113.33 (17.44)	0-123	3.21 (0.10)	6,092 (51.4)
Receptive voc 38 m (CDI)	109.75 (23.75)	0-123	3.21 (0.10)	6,092 (51.4)
Reading a/c 7 (WORD)	28.52 (9.25)	0-50	7.53 (0.31)	5,723 (50.9)
Reading a 9 (NBO)	7.58 (2.42)	0-10	9.87 (0.32)	5,574 (49.6)
Reading a 9 (NARA II)	104.27 (13.58)	69-131	9.88 (0.32)	5,048 (49.4)
Reading s 9 (NARA II)	105.60 (12.47)	69-131	9.88 (0.32)	5,037 (49.3)
Reading s 13 (TOWRE)	82.69 (10.26)	18-104	13.83 (0.20)	4,131 (48.5)
NW reading a 9 (NBO)	5.25 (2.47)	0-10	9.87 (0.31)	5,569 (49.5)
NW reading s 9 (TOWRE)	50.91 (9.34)	4-63	13.83 (0.20)	4,121 (48.4)
Spelling a 7 (NB)	7.92 (4.39)	0-15	7.53 (0.31)	5,637 (50.5)
Spelling a 9 (NB)	10.30 (3.42)	0-15	9.87 (0.31)	5,564 (49.5)
PhonAware 7 (AAT)	20.29 (9.52)	0-40	7.53 (0.31)	5,749 (50.9)
Listening c 8 (WOLD)	7.52 (1.97)	2-15	8.63 (0.30)	5,324 (50.1)
NW repetition 8 (CNRep)	7.29 (2.50)	0-12	8.63 (0.30)	5,315 (50.1)
VIQ 8 (WISC-III)	108.04 (16.74)	50-155	8.64 (0.31)	5,305 (49.9)
PIQ 8 (WISC-III)	100.24 (16.95)	46-147	8.64 (0.31)	5,296 (49.9)

Measures were assessed in unrelated ALSPAC participants with phenotypic and genotype information (genetic relationship<0.05). PIQ was assessed for sensitivity analyses only. Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CDI, Communicative Development Inventory (Toddler); CNRep, Children's Test of Nonword Repetition; m, months; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; PIQ, performance intelligence quotient; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

Mid-childhood performance intelligence: We studied performance intelligence quotient (PIQ) scores (age 8 years), assessed using the WISC-III (Table 1, Appendix S3), as part of sensitivity analyses.

Phenotype transformation: Early-childhood vocabulary and mid-childhood/early-adolescent LRA and PIQ scores were rank-transformed to achieve normality and to allow for comparisons of genetic effects across different psychological instruments. All measures were residualised for sex, age (unless measures were derived using age-specific norms) and the two most significant ancestry-informative principal components, calculated using EIGENSOFT (v6.1.4)[24]. In addition, vocabulary scores were residualised for age squared, as vocabulary develops rapidly during early childhood[25].

Analyses

Phenotypic correlations: Phenotypic correlations (r_p) were calculated for untransformed and rank-transformed scores using Spearman rank-correlation and Pearson correlation coefficients respectively. Patterns were highly similar for untransformed and transformed scores (Figure S1).

Genome-wide Complex Trait Analysis: SNP-h^2 was estimated using Restricted Maximum Likelihood (REML) analyses as implemented in Genome-wide Complex Trait Analysis (GCTA, v1.26.0, https://cnsgenomics.com/software/gcta/) software[26]. This method examines unrelated individuals, pair by pair, and predicts phenotypic similarity by genetic similarity. Genetic interrelatedness between individuals is captured by a genetic-relationship matrix (GRM)[26], i.e. a matrix with as many columns and rows as individuals. Here, the GRM matrix was created with PLINK (v1.9)[20] using individuals with a genetic relationship <0.05 ($N_{individuals}$≤6,092) and directly genotyped SNPs only (N_{SNPs}=465,740). Genetic correlations (r_g), reflecting the extent to which two measures are influenced by the same genetic factors, were estimated using bivariate REML within GCTA[27] and the GRM as described above.

Multivariate genetic analyses: To investigate the developmental origins of genetic factors influencing mid-childhood/early-adolescent language, literacy and cognition, we studied the genetic variance/co-variance structures of vocabulary at 38 months and thirteen mid-childhood/early-adolescent LRAs using Genetic-relationship-matrix Structural Equation Modelling (GSEM, https://gitlab.gwdg.de/beate.stpourcain/gsem)[28]. Multivariate trait variances were modelled using a saturated Cholesky decomposition model and maximum likelihood estimation (Appendix S4). The fitted path models are analogous to twin research methodologies. However, like GCTA, GSEM relies on GRMs to estimate genetic variance/co-variance structures between unrelated individuals (Appendix S4)[28]. GSEM models were fitted using all available observations for children across development (R:gsem library, version 0.1.5), allowing for missing data, similar to GCTA[26]. In addition, we estimated SNP-h^2, genetic correlations, factorial co-heritability (the proportion of total genetic variance explained by a specific genetic factor) and bivariate heritability (the contribution of genetic factors to the observed phenotypic covariance between two measures) with GSEM (Appendices S5-S6).

Due to computational constrains, it was not possible to include all measures of interest into one large structural equation model. Consequently, our data analysis strategy followed a two-step procedure: First, we fitted 13 trivariate Cholesky decomposition models, each consisting of expressive and receptive vocabulary at 38 months and one of the 13 LRAs (in this order, termed “forward” GSEM, Figure S2a, S3a). Second, we carried out a meta-analysis of absolute GSEM path coefficients for these 13 models across pre-defined domains including (i) reading-related measures, (ii) spelling-related measures, and (iii) all LRA outcomes (Table S1), accounting for interrelatedness

between LRAs (R:metafor library, Rv3.2.0, http://www.metafor-project.org/doku.php)(Appendix S7). As Cholesky decompositions are sensitive to the order of modelled traits, the order of the two vocabulary measures at 38 months was reversed within the 13 trivariate Cholesky decomposition models (termed "reverse" GSEM, Figure S4a, S5a) as part of sensitivity analyses. Finally, to compare LRA genetic covariance patterns with non-verbal cognitive abilities, we studied expressive and receptive vocabulary at 38 months together with PIQ at 8 years.

Experiment-wide significance threshold: The effective number of phenotypes (N=9) was calculated based on phenotypic correlations among early-childhood vocabulary and LRAs using matrix Spectral Decomposition (matSpD, https://gump.qimr.edu.au/general/daleN/matSpD/)[29]. This corresponds to an experiment-wide significance threshold of 0.005 (0.05/9).

3.3. Results

Phenotypic and genetic descriptives

Expressive and receptive vocabulary assessed at 38 months are modestly heritable as tagged by common genotyping information, with GCTA-SNP-h^2 estimates of 18% (GCTA-SNP-h^2: 0.18(SE=0.06)) and 12% (GCTA-SNP-h^2: 0.12(SE=0.06)) respectively (Figure 1a, Table S2). GCTA-SNP-h^2 estimates for LRAs assessed during mid-childhood, including reading abilities (comprehension, accuracy and speed), spelling abilities (accuracy), phonemic awareness, listening comprehension, non-word repetition and VIQ, as well as early-adolescent reading speed, were moderate, reaching up to 54% (GCTA-SNP-h^2:0.54(SE=0.07)) (Figure 1a, Table S2), as previously reported[13].

Consistent with phenotypic correlations between expressive and receptive vocabulary at 38 months (r_p=0.63, Figure 1b), bivariate genetic correlations were strong (r_g=0.86(SE=0.15), *P*=0.004, Figure 1c) and shared genetic influences accounted for ~20% of the phenotypic overlap (bivariate heritability: 0.19(SE=0.07)). Both vocabulary measures were also phenotypically correlated with language and literacy skills later in life (Figure 1b). Phenotypic correlations of LRAs with receptive vocabulary ranged between 0.14 and 0.26, and with expressive vocabulary between 0.12 and 0.18 (Figure 1b). At the genetic level, receptive vocabulary was moderately to strongly linked with the entire spectrum of LRAs, with genetic correlations ranging from 0.58 (SE=0.21, *P*=0.001) to 0.95 (SE=0.23, *P*=$1x10^{-8}$) (Figure 1c). In contrast, expressive vocabulary was genetically correlated with VIQ at 8 years only (r_g=0.38(SE=0.14), *P*=0.003, Figure 1c).

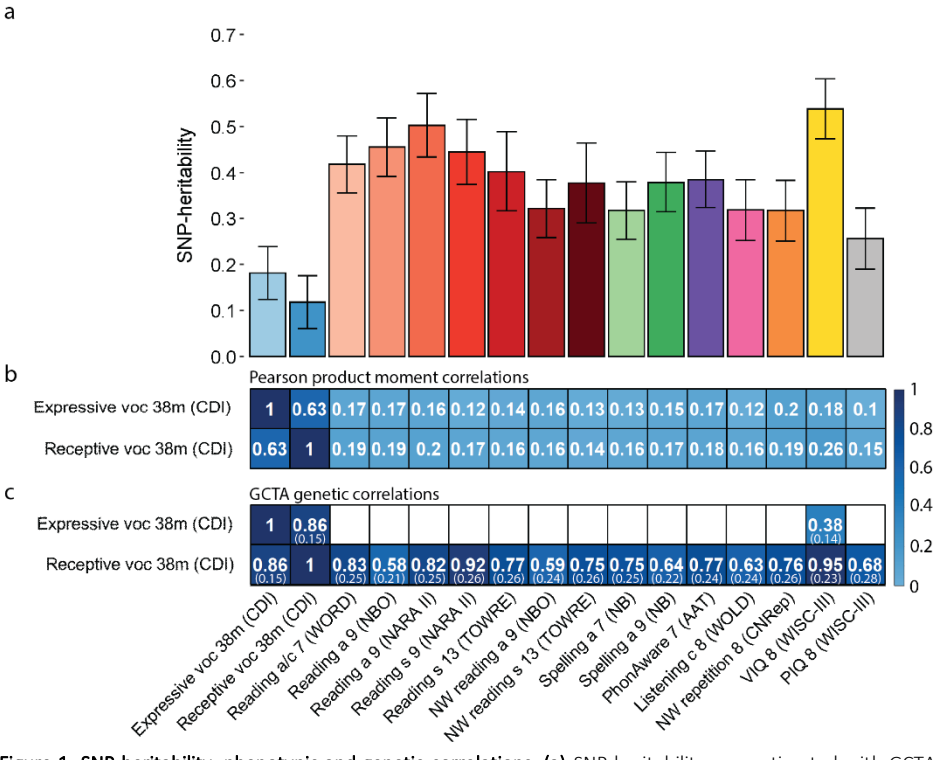


Figure 1: SNP-heritability, phenotypic and genetic correlations. (a) SNP-heritability was estimated with GCTA software, based on directly genotyped SNPs. Bars represent standard errors. **(b,c)** Bivariate correlations among rank-transformed measures passing the experiment-wide significance threshold ($P \leq 0.005$) are shown and were estimated with **(b)** Pearson correlation coefficients at the phenotype level and **(c)** GCTA at the genetic level. Standard errors for genetic correlations are shown in brackets. Trait abbreviations are described in Table 1.

Structural equation modelling

Next, we modelled multivariate genetic variances between expressive and receptive vocabulary at 38 months and, in turn, each of the 13 mid-childhood/early-adolescent LRAs using GSEM. Within each forward GSEM model, the estimated path coefficients link to shared and unique genetic variance components through structural equations (Appendix S4). SNP-h^2 estimates were consistent between GCTA and GSEM (Table S2).

Squared path coefficients for the first genetic factor (A1) fully reflect the genetic variance in expressive vocabulary at 38 months (a_{11}) and genetic variance contributions to receptive vocabulary (a_{21}) and later LRAs (a_{31}, Figure S2a). The first two path coefficients (a_{11}, a_{21}) were nearly identical across all 13 models (Figure S2) and are here reported for the model including VIQ at 8 years (Figure 2a). The first genetic factor explained 17.7%(SE=5.7%) of the phenotypic variance in expressive vocabulary (path-coefficient a_{11}=-0.42(SE=0.06), P=4x10^{-10}, Figure 2a-b), corresponding to the SNP-h^2. It also accounted for 9.0%(SE=4.5%) of the phenotypic variance in receptive vocabulary

(path-coefficient a_{21}=-0.30(SE=0.08), P=7x10^{-5}, Figure 2a-b), capturing approximately two thirds of its total genetic variance (factorial co-heritability: 0.67(SE=0.17)). In addition, this early genetic factor explained 8.6%(SE=6.0%) of the phenotypic variance in VIQ at 8 years (path-coefficient a_{31}=-0.29(SE=0.10), P=0.004, Figure 2a-b), but not other LRAs (Figure S2).

Squared path coefficients for the second genetic factor (A2) reflect the unique genetic variance in receptive vocabulary at 38 months, independent of expressive vocabulary (a_{22}, Figure S2a), as well as genetic variance contributions to later LRAs (a_{32}, Figure S2a). Nearly identically across the 13 GSEM models, the second genetic factor described a further 4.5%(SE=2.0%) of the phenotypic variance in receptive vocabulary at 38 months (path coefficient a_{22}=-0.21(SE=0.05), P=4x10^{-6}, Figure 2a-b). Thus, about a third of the genetic variance in receptive vocabulary is unique (factorial co-heritability=0.33(SE=0.17)). Importantly, this small proportion of genetic variance was amplified and accounted for the majority of genetic influences in subsequent VIQ, reading and spelling abilities (path-coefficient a_{32}, Figure 2, Figure S2-S3, Table S3). For example, this genetic factor accounted for 45.1%(SE=7.6%) of the phenotypic variance in VIQ at 8 years (path-coefficient a_{32}=-0.67(SE=0.06), P<1x10^{-10}, Figure 2a-b). Similarly, for literacy-related traits, the second genetic factor explained 38.2%(SE=6.0%) of the phenotypic variance in reading accuracy/comprehension at 7 years of age (path-coefficient a_{32}=-0.62(SE=0.05), P<1x10^{-10}, Figure 2c-d), entailing nearly the entire SNP-h^2 of the measure (factorial co-heritability: 0.94(SE=0.08), Table S3). Comparable patterns were observed for reading accuracy at 9 years (assessed with NARA II), reading speed at 9 years, reading and non-word reading speed at 13 years and spelling accuracy at 7 years, with ≥29.4% of phenotypic variation explained by genetic variance unique to receptive vocabulary (Figure S2-S3).

Squared path coefficients for the third genetic factor (A3) account for unique genetic variance in the studied LRAs, independent of genetic factors contributing to expressive and/or receptive vocabulary at 38 months (a_{33}, Figure S2a). We found little evidence for novel genetic LRA influences arising after early childhood (Figure 2, Figure S2).

At the level of individual LRAs, forward GSEMs identified two highly related developmental association patterns. The first pattern, observed for VIQ only, includes shared genetic variation with both expressive (a_{31}) and receptive (a_{32}) vocabulary (Figure 2a). The second pattern includes, primarily, an amplification of genetic influences for receptive vocabulary (a_{32}) that relate to multiple literacy skills, including reading accuracy/comprehension at 7 years (Figure 2c), reading accuracy at 9 years (assessed with NARA II), reading speed at 9 years, reading and non-word reading speed at 13 years and spelling accuracy at 7 years (Figure S2d-i).

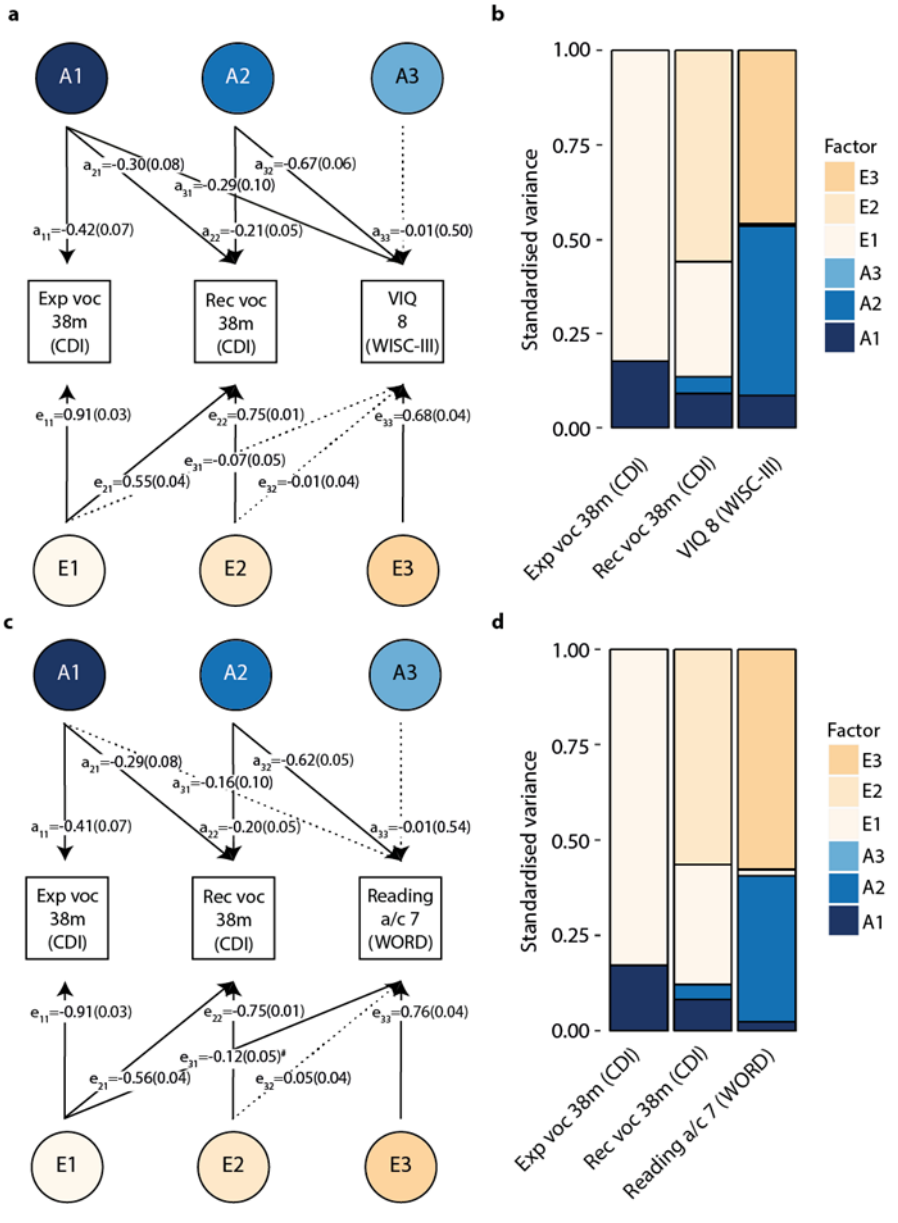


Figure 2: Path models and variance plots for early vocabulary and mid-childhood verbal intelligence or reading accuracy/comprehension. Trivariate Cholesky decomposition models (forward GSEM, GSEM software) were fitted based on all available observations for children across development (N≤6,092). **(a,c)** Path models (standardised path coefficients and standard errors) for expressive and receptive vocabulary at 38 months (CDI) and **(a)** VIQ assessed at 8 years (WISC-III) or **(c)** reading accuracy/comprehension at 7 years (WORD). Solid lines indicate path coefficients with $P \leq 0.05$, dashed lines indicate path coefficients with $P > 0.05$. **(b,d)** Standardised variance for models including **(b)** VIQ assessed at 8 years (WISC-III) and **(d)** reading accuracy/comprehension assessed at 7 years (WORD). # Path coefficient passing nominal significance ($P \leq 0.05$), but not the experiment-wide significance threshold ($P \leq 0.005$). Trait abbreviations are described in Table 1.

To evaluate overarching association patterns across mid-childhood/early-adolescent language and literacy skills (Figure S1, S6), we meta-analysed absolute path coefficients across all 13 forward GSEM models (Table S1). This meta-analysis confirmed the amplification of genetic influences that are unique to receptive vocabulary at 38 months (meta-path-coefficient a_{32}=0.62(0.06), $P<1x10^{-10}$, Figure 3, Table S4). In addition, we observed nominal evidence for an amplification of genetic influences that capture the entirety of expressive vocabulary at 38 months (meta-path-coefficient a_{31}=0.20(SE=0.08), *P*=0.009, Figure 3, Table S4). Consistent with individual GSEM models, there was little meta-analytic evidence for novel genetic influences arising after early childhood (meta-path-coefficient a_{33}=0.34(SE=0.29), *P*=0.24, Figure 3, Table S4). Meta-analyses of reading measures (N=7) and spelling measures (N=2), showed that developmental genetic amplification patterns observed across all LRAs primarily, but not exclusively, involved reading-related abilities (Table S4).

Cholesky decompositions are sensitive to the order of modelled traits, although SNP-h^2 estimations remain unchanged. We therefore created 13 additional GSEM models, as part of sensitivity analyses, reversing the order of expressive and receptive vocabulary at 38 months (reverse GSEM models, with path coefficients as detailed in Figure S4a). Consistent with forward GSEM models, there was little evidence for novel LRA-related genetic factors emerging after early childhood (A3, Figure S4-S5). For reverse GSEM, the first genetic factor (A1), capturing the entire SNP-h^2 of receptive vocabulary, accounted also for 11.8%(SE=5.5%) of the phenotypic variance in expressive vocabulary (Figure S4-S5, shown for the GSEM model including VIQ). A further 5.9%(SE=3.0%) of the phenotypic variance in expressive vocabulary was explained by a second genetic factor (A2), capturing genetic influences that are independent of receptive and unique to expressive vocabulary. Early genetic factors accounted for phenotypic variation in VIQ, reading and spelling abilities, but also phonemic awareness and/or non-word repetition (Figure S4-S5).

To identify the most predictive early genetic factors observed using either forward or reverse GSEM models, we studied factorial co-heritabilities and bivariate heritabilities. The largest contribution to the genetic variance of later LRAs was observed for genetic influences uniquely related to receptive vocabulary (A2, forward GSEM, Figure S2a), explaining up to 95%(SE=20%) in LRA SNP-h^2, especially for reading and VIQ (Table S3). In comparison, shared receptive/expressive vocabulary-related genetic influences (A1, reverse GSEM, Figure S4a) explained only up to 73%(SE=20%) of LRA SNP-h^2 (Table S3), although derived 95% confidence intervals overlap. Consistently, genetic covariance between receptive vocabulary and later LRAs accounted for the majority of their phenotypic covariance, with bivariate heritability estimates of up to 1.00(SE=0.22)(Figure 4, Table S5). In contrast, there was little evidence that genetic factors underlying expressive vocabulary, irrespective of its variance decomposition, substantially predicted variation in LRAs (Figure 4, Table S5), except for VIQ

(0.69(SE=0.24)). Thus, the majority of genetic variation in later LRAs can be attributed to a small proportion of genetic variance in early language that uniquely captures receptive vocabulary, and that has been amplified during development.

Finally, we assessed whether the identified amplification patterns extend to non-verbal cognition by studying GSEM models including PIQ at 8 years (Figure S7), as part of sensitivity analyses. Findings were largely similar to reading-related measures (Figure 2c-d). Specifically, using forward GSEM, (i) genetic influences unique to receptive vocabulary (A2) explained 25.7%(SE=6.4%) of phenotypic variance in PIQ (path-coefficient a_{32}=-0.51(SE=0.06), $P<1x10^{-10}$, Figure S7a-b; factorial co-heritability: 0.99(SE=0.04), Table S3); and (ii) there was little support for PIQ specific genetic influences that arise during mid-childhood (A3, Figure S7a-b). However, evidence for the contribution of genetic factors to the phenotypic covariance between receptive vocabulary and PIQ did not pass the experiment-wide threshold (bivariate heritability: 0.75(0.27), Figure 4, Table S5).

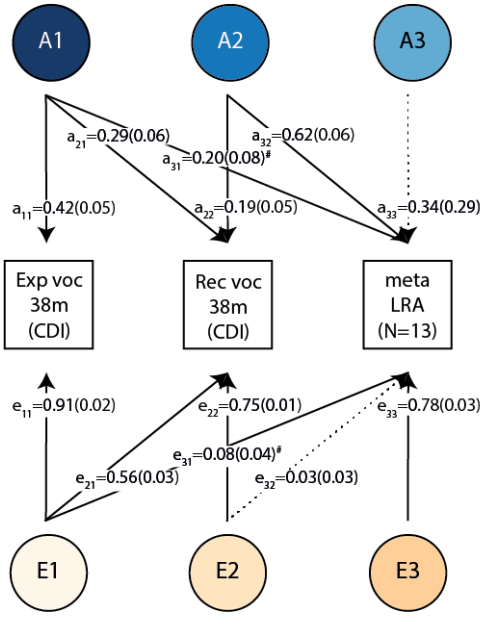


Figure 3: Meta-analyses of developmental structural models. Absolute path coefficients for 13 structural equation models (forward GSEM) corresponding to 13 LRAs in mid-childhood and early adolescence were meta-analysed, accounting for phenotypic interrelatedness. Detailed information, including estimates of effect heterogeneity, is shown in Table S4. # Path coefficient passing the nominal (P≤0.05), but not the experiment-wide significance threshold (P≤0.005). Trait abbreviations are described in Table 1.

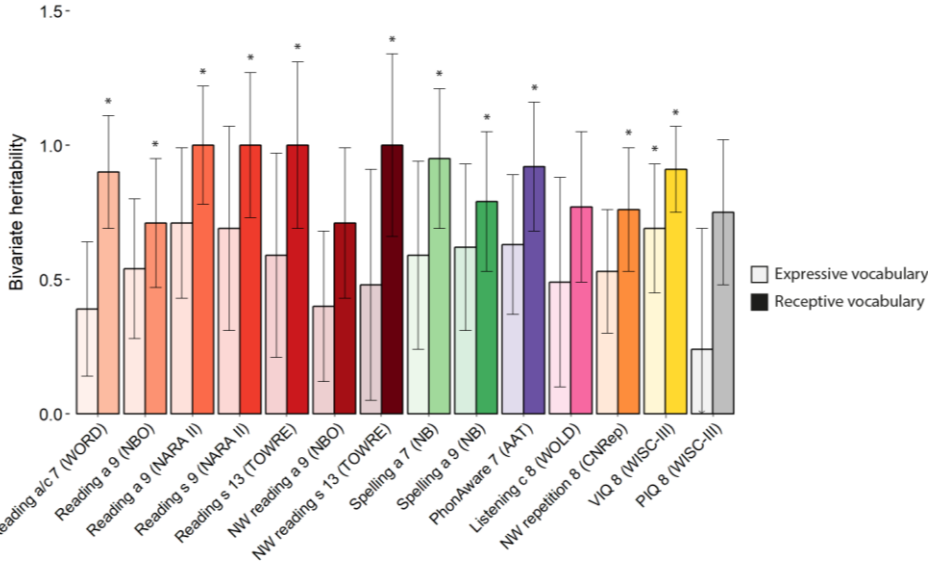


Figure 4: Bivariate heritability estimates. Bivariate heritability estimates (forward GSEM, GSEM software) reflect the proportion of the phenotypic covariance that is accounted for by the genetic covariance. Bivariate heritability estimates were truncated at one for reading a 9 (NARA II), reading s 9 (NARA II), reading s 13 (TOWRE) and NW reading s 13 (TOWRE). PIQ at 8 years was assessed for sensitivity analyses only. Bars represent standard errors. * Estimates passing the experiment-wide significance threshold ($P \leq 0.005$). Trait abbreviations are described in Table 1.

3.4. Discussion

Multivariate genetic variance analyses in this study showed that genetic factors contributing to mid-childhood/early-adolescent LRAs, including reading and spelling skills, but also phonological awareness, non-word repetition, verbal and non-verbal cognitive functioning, can already be captured by early-childhood language. Early genetic influences, especially those uniquely related to receptive vocabulary, are amplified during development and fully account for genetic variation in later reading, verbal and non-verbal cognitive skills. Independent of model specification, there was little evidence for novel genetic influences emerging during mid-childhood and early adolescence that would suggest specificity in the genetic LRA composition. Thus, developmental processes underlying language and literacy skills may not fully adhere to a paradigm that exclusively predicts genetic innovation during the transition from early to middle childhood[11,15].

The identification of amplification processes is consistent with twin research reporting moderate genetic correlations between latent factors for early language (including expressive vocabulary and syntax skills) and both mid-childhood and/or

adolescent latent language[11] and reading[10]. However, as genetic factors accounted only for about a third of the phenotypic correlations[10,11], findings have been interpreted as evidence for genetic innovation[11]. In the present study, early vocabulary-related genetic factors, especially those related to receptive vocabulary, explained the majority of genetic variance (≤99% SNP-h^2) for many later reading and cognitive skills. The difference in results, implicating amplification instead of innovation processes, might be due to two reasons. First, previous studies focused on early-childhood expressive language skills only. In the current study, however, the largest amplification was observed for a small proportion of genetic variance that is unique to early receptive and independent of early expressive vocabulary. Consistently, the majority of phenotypic covariance between early receptive vocabulary and later skills, especially literacy and cognition, was accounted for by shared genetic sources. In contrast, genetic influences in expressive vocabulary did not substantially contribute to the total genetic variance of later LRAs, despite some evidence for genetic interrelationships with VIQ. Thus, structural models omitting genetic factors influencing early receptive vocabulary may attribute developmental changes in the genetic architecture of mid-childhood/early-adolescent traits to genetic innovation processes. Note that VIQ findings were representative of many (less powerful) WISC-III subtests, including the WISC-III vocabulary subtest, showing similar association patterns (data not shown). Second, this study benefits from a direct estimation of genetic interrelationships between individuals, based on genotyping information[28], enabling the detection of small changes in SNP-h^2, compared to a more indirect assessment based on twin correlations.

The similarity in developmental genetic changes predicting the genetic composition of mid-childhood/early-adolescent reading and cognitive skills, as observed by factorial co-heritability estimates, is consistent with overarching developmental patterns. According to the 'generalist genes' hypothesis, cognitive abilities are presumed to share genetic variance components[30]. Our results suggest that an early generalist genetic component may manifest with the emergence of receptive vocabulary by the age of three years. This shared genetic component may imply developmentally stable biological mechanisms, but could also reflect different regulations of the same genes over time, although the current findings do not allow us to infer specific biological pathways.

At the same time, amplification processes predominate for genetic factors underlying early receptive vocabulary compared to genetic factors contributing to early expressive vocabulary, suggesting some degree of genetic specificity. Hence, the underlying genetic mechanisms may only partially adhere to the concept of 'generalist genes'[30], for the following reasons: Early language skills at the age of three, including vocabulary, comprehension and sentence construction have been linked to adolescent reading comprehension[31]. Notably, broadly defined early oral language, including receptive skills[32], has been shown to affect word recognition[33], while vocabulary

comprehension is also a precursor of listening comprehension[8]. Thus, receptive vocabulary skills might show wide-ranging links with both key predictors of reading comprehension, decoding and language comprehension, as proposed by the 'simple view of reading'[6]. Consistently, a delay in both expressive and receptive vocabulary at the age of two, is much more likely to lead to problems with later literacy, compared to delays in expressive vocabulary alone[34], and expressive and receptive vocabulary may be independently related to pre-reading skills[35]. Furthermore, variation in comprehension has been associated with non-linguistic cognitive measures, such as tool use and symbolic play, compared to expressive vocabulary[36]. Consequently, genetic variation for receptive vocabulary at 38 months may share genetic foundations with several key skills that are important for future reading, language and cognitive development, detectable as genetic amplification, and only partially overlap with cognitive mechanisms that are predicted by genetic factors influencing expressive vocabulary alone. Note that literacy abilities in this study primarily assessed accuracy and speed of reading and spelling, and that our findings may, thus, only partially apply to reading comprehension.

Increased SNP-h^2 estimates of language/literacy skills from mid-childhood to adolescence, compared to estimates of early-childhood language, may arise due to genotype-environment correlations, as children modify and select their environment in accordance with their genetic make-up[37]. Furthermore, the environmental variance may decrease with the start of schooling[38]. Finally, parent-reported vocabulary measures might be associated with higher random error rates (rendering them less reliable) than direct assessments of language and literacy skills using standardised psychological instruments, which consequently may affect the reliability of heritability estimations[39]. Thus, our findings do not preclude the emergence of novel genetic influences from mid-childhood onwards. Parent-reported vocabulary measures in ALSPAC have sufficient power (80%) to detect SNP-h^2 estimates of ≥0.15 (Appendix S8). However, compared to large-scale genome-wide studies of educational attainment[40] or direct assessments of language and literacy measures, their predictive power is low. This advocates a need for improvement of instruments assessing early language skills, especially as moderate to strong correlations between parental judgements and direct assessments of a child's vocabulary suggest sufficient instrument validity[23,41]. A further limitation of the current study is that the CDI Words & Sentences was developed for vocabulary assessment in children up to 30 months[21], whereas ALSPAC children were assessed at 38 months of age, potentially leading to ceiling effects. Finally, the lack of independent cohorts with data on both early expressive and receptive vocabulary prevents a direct replication of our findings.

The strength of this work lies in the identification of amplification processes using longitudinal models, suggesting that the developmental origins of many later complex skills, especially those related to literacy and cognition, lie in early childhood. Thus, cheaply and easily administered parent-reported CDI questionnaires, which are widely used to assess children's early language[42], might be useful instruments to capture genetic variation in language, literacy and cognitive skills many years later in life.

Key points

- It is known that individual differences in preschool vocabulary predict later language and literacy skills and that this relationship involves genetic mechanisms.
- We found evidence suggesting that genetic factors contributing to a wide range of mid-childhood/early-adolescent language- and literacy-related skills originate in early-childhood language
- Early genetic influences, especially those uniquely related to receptive vocabulary, are amplified during development and account for the majority of genetic variance in later reading, verbal and non-verbal skills, while processes of genetic innovation during mid-childhood and adolescence were negligible.
- Our findings highlight the predictive power of parental vocabulary reports to capture genetic variation in language, literacy and cognitive skills many years later in life.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. GWAS data was generated by Sample Logistics and Genotyping Facilities at the Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). EV, BSTP and SEF are supported by the Max Planck Society. BSTP is also supported by the Simons Foundation (Award ID: 514787). This publication is the work of the authors and EV and BSTP will serve as guarantors for the contents of this paper.

References

1. Bleses, D., Makransky, G., Dale, P. S., Højen, A. & Ari, B. A. Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics* **37**, 1461–1476 (2016).
2. Bornstein, M. H., Hahn, C.-S., Putnick, D. L. & Suwalsky, J. T. D. Stability of Core Language Skill from Early Childhood to Adolescence: A Latent Variable Approach. *Child Development* **85**, 1346–1356 (2014).
3. Duff, F. J., Reen, G., Plunkett, K. & Nation, K. Do infant vocabulary skills predict school-age language and literacy outcomes? *J Child Psychol Psychiatry* **56**, 848–856 (2015).
4. Lee, J. Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics* **32**, 69–92 (2011).
5. Durham, R. E., Farkas, G., Hammer, C. S., Bruce Tomblin, J. & Catts, H. W. Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility* **25**, 294–305 (2007).
6. Gough, P. B. & Tunmer, W. E. Decoding, Reading, and Reading Disability. *Remedial and Special Education* **7**, 6–10 (1986).
7. Jong, P. F. de & Leij, A. van der. Effects of Phonological Abilities and Linguistic Comprehension on the Development of Reading. *Scientific Studies of Reading* **6**, 51–77 (2002).
8. Dickinson, D. K. & Neuman, S. B. *Handbook of Early Literacy Research*. (Guilford Press, 2007).
9. Dich, N. & Cohn, A. C. A review of spelling acquisition: Spelling development as a source of evidence for the psychological reality of the phoneme. *Lingua* **133**, 213–229 (2013).
10. Harlaar, N., Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. Why Do Preschool Language Abilities Correlate With Later Reading? A Twin Study. *J Speech Lang Hear Res* **51**, 688–705 (2008).
11. Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev Sci* **15**, 233–49 (2012).
12. St Pourcain, B. *et al.* Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature communications* **5**, 4831 (2014).
13. Verhoef, E. *et al.* Disentangling polygenic associations between attention-deficit/hyperactivity disorder, educational attainment, literacy and language. *Translational Psychiatry* **9**, 35 (2019).
14. Plomin, R., DeFries, J. C., Knopik, V. S. & Neiderhiser, J. M. *Behavioral Genetics*. (Worth Publishers, 2016).
15. Briley, D. A. & Tucker-Drob, E. M. Explaining the increasing heritability of cognitive ability across development: a meta-analysis of longitudinal twin and adoption studies. *Psychol Sci* **24**, 1704–1713 (2013).
16. Haworth, C. *et al.* The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* **15**, 1112–1120 (2010).
17. Plomin, R. & DeFries, J. C. The origins of individual differences in infancy; the Colorado adoption project. *Science* **230**, 1369+ (1985).
18. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* **42**, 97–110 (2013).
19. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111–127 (2013).
20. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* **81**, 559–575 (2007).
21. Fenson, L. *et al. User's Guide and Technical Manual for the MacArthur Communicative Development Inventories*. (Singular Publishing, 1993).
22. Dale, P. S. The validity of a parent report measure of vocabulary and syntax at 24 months. *J Speech Hear Res* **34**, 565–571 (1991).
23. Ring, E. D. & Fenson, L. The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language* **20**, 141–159 (2000).
24. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904–909 (2006).
25. Brooks, R. & Meltzoff, A. N. Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: a longitudinal, growth curve modeling study. *J Child Lang* **35**, 207–220 (2008).

26. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
27. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540–2542 (2012).
28. St Pourcain, B. *et al.* Developmental changes within the genetic architecture of social communication behaviour: A multivariate study of genetic variance in unrelated individuals. *Biological Psychiatry* **83**, 598–606 (2017).
29. Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)* **95**, 221–227 (2005).
30. Plomin, R. & Kovas, Y. Generalist genes and learning disabilities. *Psychol Bull* **131**, 592–617 (2005).
31. Frost, J., Madsbjerg, S., Niedersøe, J., Olofsson, Å. & Sørensen, P. M. Semantic and phonological skills in predicting reading development: from 3–16 years of age. *Dyslexia* **11**, 79–92 (2005).
32. Bryant, P., Maclean, M. & Bradley, L. Rhyme, language, and children's reading. *Applied Psycholinguistics* **11**, 237–252 (1990).
33. NICHD Early Child Care Research Network. Pathways to reading: the role of oral language in the transition to reading. *Dev Psychol* **41**, 428–442 (2005).
34. Psyridou, M., Eklund, K., Poikkeus, A.-M. & Torppa, M. Reading outcomes of children with delayed early vocabulary: A follow-up from age 2–16. *Research in Developmental Disabilities* **78**, 114–124 (2018).
35. Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W. & Wolf, M. The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *J. Speech Lang. Hear. Res.* **50**, 1093–1109 (2007).
36. Bates, E., Dale, P. S. & Thal, D. Individual Differences and their Implications for Theories of Language Development. in *The Handbook of Child Language* 95–151 (John Wiley & Sons, Ltd, 2019). doi:10.1111/b.9780631203124.1996.00005.x.
37. Plomin, R., DeFries, J. C. & Loehlin, J. C. Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin* **84**, 309–322 (1977).
38. Samuelsson, S. *et al.* Response to early literacy instruction in the United States, Australia, and Scandinavia: A behavioral-genetic analysis. *Learn Individ Differ* **18**, 289–295 (2008).
39. Walters, R., Churchhouse, C. & Hosking, R. Heritability 101: What is 'heritability'? http://www.nealelab.is/blog/2017/9/13/heritability-101-what-is-heritability (2019).
40. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* **50**, 1112–1121 (2018).
41. Sachse, S. & Von Suchodoletz, W. Early identification of language delay by direct language assessment or parent report? *J Dev Behav Pediatr* **29**, 34–41 (2008).
42. Frank, M. C., Braginsky, M., Yurovsky, D. & Marchman, V. A. Wordbank: an open repository for developmental vocabulary data. *J Child Lang* **44**, 677–694 (2017).

Supplementary Materials

Supplementary Methods

Appendix S1: ALSPAC description

14,541 pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992 were recruited by ALSPAC. Initially, 14,541 pregnancies were enrolled for which the mother enrolled in the ALSPAC study and had either returned at least one questionnaire or attended a "Children in Focus" clinic by 19/07/99. This comprised a total of 14,676 fetuses, resulting in 14,062 live births and 13,988 children who were alive at the age of one year.

An attempt was made to bolster the initial sample, when the oldest children were approximately 7 years of age, with eligible cases who had failed to join the study originally. Consequently, there are data available for more than 14,541 pregnancies (see above) when considering variables collected from the age of seven onwards (and potentially abstracted from obstetric notes).

There are 913 pregnancies not in the initial sample (known as Phase I enrolment) that are currently represented on the built files and reflecting enrolment status at the age of 24. Of these new pregnancies, 452, 262 and 195 were recruited during Phases II, III and IV respectively. As a result, an additional 913 children are enrolled. The cohort profile paper[1] describes the phases of enrolment in more detail.

Thus, the total sample size for analyses using any data collected after the age of seven is 15,454 pregnancies, resulting in 15,589 fetuses. Of this total sample, 14,901 children were alive at the age of one year.

The Children in Focus (CiF) group, a 10% sample of the ALSPAC cohort, attended clinics at the University of Bristol at various time intervals between 4 to 61 months of age. The CiF group was randomly chosen from the last 6 months of ALSPAC births (1,432 families attended at least one clinic). Mothers that had moved out of the area, were lost to follow-up, or those partaking in another study of infant development in Avon were excluded.

Please note that the study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool (http://www.bristol.ac.uk/alspac/researchers/our-data).

Appendix S2: Genetic quality control

ALSPAC participants were genotyped using the Illumina HumanHap550 quad chip genotyping platforms, and genotypes were called using the Illumina GenomeStudio software. Standard genomic quality control[2] was performed at both the SNP and individual level using PLINK (v1.07)[3]. Individuals with a gender mismatch (comparing

reported gender with genetic gender), a large number of missing SNPs (>3%), non-European ancestry and a genetic relationship >0.05 were excluded from the analyses. SNPs that had a low call rate (<99%), were rare (<1%) or deviated from Hardy-Weinberg equilibrium ($P<5x10^{-7}$) were also excluded from the analysis. After quality control, 8,226 children and 465,740 SNPs remained. Among those, 6,092 children had also early-childhood vocabulary and/or mid-childhood/early-adolescent language and literacy information available.

Appendix S3: Mid-childhood/early-adolescent ALSPAC measures

Children were voluntarily brought to the clinic to be tested during a half-day visit. Tests were administered individually, by trained assessors, during 20-minute sessions. An extensive description of each assessment is available in the online ALSPAC documentation: http://www.bristol.ac.uk/alspac/researchers/access/.

Reading accuracy and comprehension age 7 (WORD)

To assess decoding and word reading both pictures and words were used, as part of the basic reading subtest of the Wechsler Objective Reading Dimensions (WORD)[4]. The WORD has high internal consistency and reliability, with coefficients above 0.9. Inter-correlation with the word-reading test from the Differential Ability Scale[5] was 0.82. In short, a series of four pictures, with for each picture four short, simple words underneath it, was shown to the child. Then, the child was asked to indicate, by pointing, the word that had the same beginning or ending sound as the picture. Next, a series of three pictures, again each with four words beneath it, starting with the same letter as the picture, were shown to the child. The child was asked to point to the word underneath each picture that correctly named the picture. Finally, the child was asked to read aloud a series of 48 unconnected words which increased in difficulty. If the child made six consecutive errors, the task was stopped. The reading accuracy and comprehension score was computed as the sum of items that the child read/responded correctly with a maximum score of 50.

Reading accuracy age 9 (NBO)

The child was asked to read aloud ten real words, selected from a larger selection of words as described by Nunes, Bryant and Olsen (NBO)[6]. The test-retest reliability of this assessment for word reading was 0.80 and a 0.85 correlation with the Schonell Word Reading Task[7] was observed. A score indicating reading accuracy was computed as the sum of the number of items that the child read correctly.

Reading speed and reading accuracy age 9 (NARA II)

The child was asked to read a passage from a booklet, following the revised Neale Analysis of Reading Ability (NARA II)[8]. The tester recorded both the time it took the child to read the passage, and also noted any errors made by the child. Alternate form reliability ranged from 0.84 to 0.92, depending on age at assessment, for accuracy and, between 0.50 and 0.83 for speed[9]. Measures of accuracy and speed correlated 0.95 and 0.76 with the Schonell Graded Word Reading Test respectively[7]. The maximum test score was 400 and 100 for speed and accuracy, respectively. Both raw scores were standardised by age.

Reading speed age 13 (TOWRE)

The child had 45 seconds to read as many words as possible from a list of 104 words, as included in the Test of Word Reading Efficiency (TOWRE)[10], to assess sight word efficiency. The alternate form reliability for this test ranged between 0.86 and 0.97, depending on age and sub-task. Correlations with the Woodcock Reading Mastery Tests ranged from 0.89 to 0.94, depending on age and sub-task[9]. The tester marked words that a child skipped, or got wrong. A reading speed score was computed as the sum of the number of correct words a child finished on.

Non-word reading accuracy age 9 (NBO)

The child was asked to read aloud ten non-words, selected from a larger selection of non-words taken from research conducted by Nunes and colleagues[6]. The test-retest reliability of the non-word reading task was 0.73 and correlation with the Schonell Word Reading Task[7] of 0.73 was observed. The tester emphasised to the child that the words ware made-up, and asked the child to read all the non-words in the way that they thought they should be read. A non-word reading accuracy score was computed as the sum of the number of items the child read correctly.

Non-word reading speed age 13 (TOWRE)

The child was asked to read as many non-words as possible out of a list of 63 words within 45 seconds. Word lists were derived from the non-word part of the Test of Word Reading Efficiency (TOWRE)[10] to assess decoding efficiency. The alternate form reliability was 0.94 and the test-retest reliability ranged between 0.82 and 0.97[9]. Correlations with the Woodcock Reading Mastery Tests were high and ranged from 0.89 to 0.91[9]. The tester marked words that a child skipped, or got wrong. A non-word reading speed score is computed as the sum of the number of correct non-words a child finished on.

Spelling accuracy age 7 (NB)

The child was asked to write down the spelling for a series of 15 words, chosen specifically for this age group after piloting on several hundred children (Nunes and Bryant, ALSPAC-specific measure). The words were of different frequencies, included regular and irregular words, and increased in difficulty. For each word, the tester first read only the word to the child, then a specific sentence incorporating the word, and finally alone again. A spelling accuracy score is computed as the number of words spelt correctly.

Spelling accuracy age 9 (NB)

Spelling accuracy at age 9 was assessed in a similar manner to that at age 7 (see above). However, the series of 15 words that a child was asked to spell were adjusted to match the age group of 9. A spelling accuracy score is computed as the number of words spelt correctly.

Phonemic awareness age 7 (AAT)

The task consisted of two practice and 40 test items of increasing difficulty, according to the Auditory Analysis Test (AAT)[11]. This test had correlations from 0.53 to 0.84 with the language arts subtests of Stanford Achievement Test[11]. Reliability of the AAT was not assessed in the initial task report[11]. However, test criteria of the commercial version of the AAT, TAAS[12] have been assessed and revealed high internal consistency (0.78). Interrelations with other tests of phonemic awareness ranged between 0.11 and 0.82 and suggest some construct validity[13]. For each item, the child was asked to first repeat the word and then produce it again but without part of the word (a phoneme or a number of phonemes). There were seven omission categories: 1) omission of a first syllable, 2) omission of a medial syllable, 3) omission of a final syllable, 4) omission of the initial consonant of a one-syllable word, 5) omission of the final consonant of a one-syllable word, 6) omission of the first consonant of a medial consonant, and 7) omission of the consonant blend of a medial consonant. Words from similar categories were not clustered. A phonemic awareness score is computed as the sum of correct responses.

Listening comprehension age 8 (WOLD)

A picture was shown to the child and the tester read aloud a paragraph about the picture, following a subset of the Wechsler Objective Language Dimensions (WOLD)[14] test. Next, the child was asked to answer fifteen questions on what they heard. A listening comprehension score is calculated as the sum of the items that the child got correct. The listening comprehension subtest has test-retest reliabilities between 0.83 and 0.88 in children aged six to eleven years[15] and correlation with the Peabody Picture Vocabulary Test-III[16] was 0.44[17].

Non-word repetition age 8 (CNRep)

The child was asked to listen to a series of 12 nonsense words, according to an adaptation of the Children's Test of Nonword Repetition (CNRep)[18]. The test-retest reliability was 0.80 and correlations with the digit span test ranged between 0.45 and 0.67[18]. The 12 nonsense words consisted of four nonsense words of three syllables, four nonsense words of four syllables, and four nonsense words of five syllables. All nonsense words were conforming to English rules for sound combinations. For each word, the child was asked to repeat the word after listening to it. The repetition attempt was scored as correct if there was no phonological deviation from the target form. A non-word repetition score was computed as the sum of the number of correct non-words.

Verbal intelligence age 8 (WISC-III)

A short form of the Wechsler Intelligence Scale for Children (WISC-III)[19], including alternate items for all subtests except for the coding subtest, was administered. The WISC-III comprises ten subtests five of which are verbal subtests: information, similarities, arithmetic, vocabulary, comprehension, and can be used to construct a verbal intelligence score. Based on the items used in the alternate item form of the WISC-III raw scores were calculated and the total age-scaled scores for the verbal scale were calculated using the look-up tables provided in the WISC-III manual, with a maximum VIQ score of 160. All scores were pro-rated. Test-retest correlations of the WISC-III verbal intelligence are high and ranged between 0.90 and 0.94, depending on the age at assessment and the duration of the test-retest interval[20]. The VIQ is also highly correlated with the Kaufman Brief Intelligence Test (0.79) and with the Standford-Binet IV (0.69)[20].

Performance intelligence age 8 (WISC-III)

A short form of the Wechsler Intelligence Scale for Children (WISC-III)[19], including alternate items for all subtests except for the coding subtest, was administered. The WISC-III contains five performance subtests: picture completion, coding, picture arrangement, block design and object assembly. Based on the items used in the alternate item form of the WISC-III raw scores were calculated and the total age-scaled scores for the performance scale were calculated using the look-up tables provided in the WISC-III manual, with a maximum PIQ score of 160. All scores were pro-rated. Test re-test correlations for the WISC-III performance intelligence are 0.89[21]. The correlation between PIQ assessed using WISC-III and the non-verbal score measured using the Otis-Lennon School Ability Test was 0.59[22].

Appendix S4: Structural equation modelling

A Cholesky decomposition can be described as follows[23]: for a multivariate trait P with phenotypic measurements t, a latent genetic factor (A_1) influences the first measure P_1, but may also explain variance in the remaining measures (P_2,...,P_t). Additionally, a second latent genetic factor (A_2) influences the second measure (P_2) and may explain variance, not yet captured by A_1, in all other measures (P_3,...,P_t). The final measure (P_t) is influenced by latent genetic factors (A_1,...,A_{t-1}), but also a genetic factor A_t. This latter genetic factor does not explain variance within any of the previous measures (P_1,...,P_{t-1})[24]. Genetic path coefficients were annotated with *a.* Here, the first number indicates the direction of the effect (the variable to which the arrow points) and the second the origin of the effect[24].

The expected phenotypic covariance matrix $\boldsymbol{\Sigma}$ for Z-standardised traits, based on the factor model is

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Gamma}\boldsymbol{\Theta}\boldsymbol{\Gamma}' \quad (1)$$

with a lower triangular matrix of genetic path coefficients $\boldsymbol{\Lambda}$, a diagonal matrix of latent genetic factor variances $\boldsymbol{\Phi}$ (standardised to unit variance), such that $\boldsymbol{\Phi}$ is an identity matrix *I*[25]. The residual variance can be decomposed into latent residual factors, with a lower triangular matrix of residual path coefficients $\boldsymbol{\Gamma}$ and a diagonal matrix of latent residual factor variances $\boldsymbol{\Theta}$ (standardised to unit variance), such that $\boldsymbol{\Theta}$ is an identity matrix *I*. For example, a trivariate model consisting of three measures (P_1 , P_2 and P_3), assuming three genetic factors (A_1, A_2 and A_3) and three residual factors (E_1, E_2 and E_3). The expected phenotypic covariance matrix for this trivariate model can be expressed as follows:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{p1}^2 & \sigma_{p12} & \sigma_{p13} \\ \sigma_{p12} & \sigma_{p2}^2 & \sigma_{p23} \\ \sigma_{p13} & \sigma_{p23} & \sigma_{p3}^2 \end{bmatrix} \quad (2)$$

with the relevant matrices

$$\boldsymbol{\Lambda} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \boldsymbol{\Phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{\Gamma} = \begin{bmatrix} e_{11} & 0 & 0 \\ e_{21} & e_{22} & 0 \\ e_{31} & e_{32} & e_{33} \end{bmatrix}, \boldsymbol{\Theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

with phenotypic variances σ_{p1}^2, σ_{p2}^2 and σ_{p3}^2 and phenotypic covariances σ_{p12} , σ_{p13} and σ_{p23} .

The trivariate AE Cholesky decomposition of three standardised measures (see above), can be visualised using a path diagram (Figure S3a). The expected phenotypic variances and covariances can be expressed as follows:

$$\sigma^2_{p1} = a^2_{11} + e^2_{11} = 1 \quad (4)$$

$$\sigma^2_{p2} = (a^2_{21} + a^2_{22}) + (e^2_{21} + e^2_{22}) = 1 \quad (5)$$

$$\sigma^2_{p3} = (a^2_{31} + a^2_{32} + a^2_{33}) + (e^2_{31} + e^2_{32} + e^2_{33}) = 1 \quad (6)$$

$$\sigma_{p12} = a_{11}a_{21} + e_{11}e_{21} \quad (7)$$

$$\sigma_{p13} = a_{11}a_{31} + e_{11}e_{31} \quad (8)$$

$$\sigma_{p23} = a_{31}a_{21} + a_{32}a_{22} + e_{31}e_{21} + e_{32}e_{22} \quad (9)$$

The variance of the latent genetic and residual factors has been standardised to unit variance and is not shown.

Bivariate genetic correlations between phenotypes, measuring the extent to which two phenotypes 1 and 2 share genetic factors (ranging from -1 to 1), can be derived using estimated genetic variances and covariances[26] according to:

$$r_g = \frac{\sigma_{g12}}{\sqrt{\sigma^2_{g1}\sigma^2_{g2}}} \quad (10)$$

with genetic covariance σ_{g12} between phenotypes 1 and 2 and the genetic variances σ^2_{g1} and σ^2_{g2}.

Appendix S5: Factorial co-heritability

A measure of factorial co-heritability was estimated to quantify the relative contribution of a genetic factor to the genetic variance of a phenotype, using the gsem package (R:gsem library, version 0.1.5). The factorial co-heritability f_g^2 was derived as:

$$f_g^2 = \frac{\sigma^2_{g\_it}}{\sum \sigma^2_{g\_it}} = \frac{\sigma^2_{g\_it}}{\sigma^2_{g\_t}} \quad (11)$$

where $\sigma^2_{g\_it}$ is the genetic variance of the genetic factor i contributing to trait t and $\sigma^2_{g\_t}$ the total genetic variance of trait t, all estimated from standardised path coefficients. Corresponding standard errors (SEs) were derived using the Delta method, and *P*-values approximated with a Wald test.

Appendix S6: Bivariate heritability

Bivariate heritability[27] details the proportion of phenotypic covariance between two traits that is accounted for by the genetic covariance and was estimated using the gsem package (R:gsem library, version 0.1.5). The genetic covariance was estimated based on unstandardised path coefficients and the phenotypic covariance on observed rank-transformed measures. SEs were approximated by the SE of the genetic covariance divided by the phenotypic covariance (as the SE of the phenotypic covariance is small)

and *P*-values based on a Wald-test, assuming normality. Reported bivariate heritability estimates are based on forward GSEM models, and reverse GSEM models provided nearly identical results.

Appendix S7: Meta-analysis across mid-childhood/early-adolescent language- and literacy-related abilities

A meta-analysis of absolute GSEM path coefficients across pre-defined domains including (i) reading-related measures, (ii) spelling-related measures, and (iii) all LRA outcomes (Table S1) was carried out across forward GSEM models. Estimates were combined using random-effects meta-regression intercepts, accounting for interrelatedness between LRAs (R:metafor library, Rv3.2.0, http://www.metafor-project.org/doku.php)[28]. For this, a variance/covariance matrix across measures was approximated by including the observed phenotypic correlation matrix, weighted by the standard errors of the path coefficients as estimated by GSEM, analogous to models accounting for correlated phylogenetic histories[29].

Appendix S8: Power analyses

For this study, we selected early vocabulary measures (15-38 months) that had a sample size >6,000. This corresponds to at least 80% power to detect a SNP-h^2 of 0.15 (P>0.05)[30]. For LRAs assessed later in life (7-13 years), sample sizes were slightly lower and measures that had at least 80% power to detect a SNP-h^2 of >0.20 (P>0.05)[30] were selected, corresponding to a sample size of ~4,000.

Supplementary Tables

Table S1: Meta-analysis domains of mid-childhood/early-adolescent language- and literacy-related abilities

LRA	Meta-analysis domains		
	All LRAs	Reading	Spelling
Reading a/c 7 (WORD)	✓	✓	×
Reading a 9 (NBO)	✓	✓	×
Reading a 9 (NARA II)	✓	✓	×
Reading s 9 (NARA II)	✓	✓	×
Reading s 13 (TOWRE)	✓	✓	×
NW reading a 9 (NBO)	✓	✓	×
NW reading s 13 (TOWRE)	✓	✓	×
Spelling a 7 (NB)	✓	×	✓
Spelling a 9 (NB)	✓	×	✓
PhonAware 7 (AAT)	✓	×	×
Listening c 8 (WOLD)	✓	×	×
NW repetition (CNRep)	✓	×	×
VIQ 8 (WISC-III)	✓	×	×

Absolute path coefficients estimated using genetic-relationship-matrix structural equation (GSEM) models were meta-analysed, accounting for trait-interrelationships. Meta-analysis were carried out across all LRAs, and for reading-related abilities as well as spelling-related abilities. ✓ indicates that a specific trait was included in a meta-analysis. × indicates that a specific trait was not included. Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CNRep, Children's Test of Nonword Repetition; LRAs, language- and literacy-related abilities; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

Table S2: SNP-heritability estimates

Measure	N	GCTA-h^2 (SE)	GSEM-h^2 (SE)
Expressive voc 38m (CDI)	6,092	0.18(0.06)	0.18(0.06)*
Receptive voc 38m (CDI)	6,092	0.12(0.06)	0.13(0.04)*
Reading a/c 7 (WORD)	5,723	0.42(0.06)	0.41(0.06)
Reading a 9 (NBO)	5,574	0.46(0.06)	0.46(0.06)
Reading a 9 (NARA II)	5,048	0.50(0.07)	0.49(0.07)
Reading s 9 (NARA II)	5,037	0.45(0.07)	0.43(0.07)
Reading s 13 (TOWRE)	4,131	0.40(0.09)	0.41(0.09)
NW reading a 9 (NBO)	5,569	0.32(0.06)	0.33(0.06)
NW reading s 13 (TOWRE)	4,121	0.38(0.06)	0.38(0.09)
Spelling a 7 (NB)	5,637	0.32(0.06)	0.33(0.06)
Spelling a 9 (NB)	5,564	0.38(0.06)	0.38(0.07)
PhonAware 7 (AAT)	5,749	0.39(0.06)	0.38(0.06)
Listening c 8 (WOLD)	5,324	0.32(0.07)	0.30(0.07)
NW repetition (CNRep)	5,315	0.32(0.07)	0.31(0.06)
VIQ 8 (WISC-III)	5,305	0.54(0.07)	0.54(0.06)
PIQ 8 (WISC-III)	5,296	0.26(0.07)	0.26(0.06)

SNP-heritability estimates were estimated based on rank-transformed scores, directly genotyped SNPs and individuals with a genetic relationship of <0.05, using Restricted Maximum Likelihood (REML) analyses as implemented in genome-wide complex trait analysis (GCTA) software. SNP-heritability estimates based on forward Genetic-relationship-matrix Structural Equation modelling (GSEM) were extracted for comparison. PIQ at 8 years was assessed for sensitivity analyses only. * GSEM-h^2 estimates as observed in the GSEM model with VIQ at 8 years. Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CDI, Communicative Development Inventory; CNRep, Children's Test of Nonword Repetition; GCTA, genome-wide complex trait analysis; GSEM, Genetic-relationship-matrix Structural Equation modelling; h^2, heritability; m, months; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; PIQ; performance intelligence quotient; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

Table S3: Factorial co-heritabilities

Measure	Forward GSEM				Reverse GSEM			
	Expressive vocabulary 38 months#		Receptive vocabulary 38 months##		Expressive vocabulary 38 months###		Receptive vocabulary 38 months####	
	Factorial co-heritability (SE)	*P*	Factorial co-heritability (SE)	*P*	Factorial co-heritability (SE)	*P*	Factorial co-heritability (SE)	*P*
Reading a/c 7 (WORD)	0.06(0.07)	0.41	0.94(0.08)	$<1x10^{-10}$	0.43(0.23)	0.06	0.57(0.23)	0.01
Reading a 9 (NBO)	0.10(0.10)	0.29	0.33(0.45)	0.47	0.10(0.30)	0.73	0.33(0.24)	0.17
Reading a 9 (NARA II)	0.15(0.12)	0.19	0.85(0.12)	$<1x10^{-10}$	0.29(0.23)	0.22	0.71(0.23)	0.002
Reading s 9 (NARA II)	0.09(0.09)	0.36	0.91(0.09)	$<1x10^{-10}$	0.40(0.24)	0.10	0.60(0.23)	0.01
Reading s 13 (TOWRE)	0.09(0.11)	0.43	0.91(0.38)	0.01	0.42(0.40)	0.29	0.57(0.30)	0.05
NW reading a 9 (NBO)	0.07(0.09)	0.46	0.50(0.73)	0.49	0.23(0.54)	0.68	0.35(0.29)	0.23
NW reading s 13 (TOWRE)	0.05(0.08)	0.56	0.95(0.20)	$2x10^{-6}$	0.49(0.27)	0.07	0.51(0.27)	0.05
Spelling a 7 (NB)	0.09(0.10)	0.38	0.90(0.37)	0.01	0.45(0.27)	0.10	0.55(0.27)	0.04
Spelling a 9 (NB)	0.12(0.12)	0.29	0.43(0.60)	0.47	0.14(0.38)	0.71	0.40(0.29)	0.16
PhonAware 7 (AAT)	0.16(0.13)	0.21	0.79(1.03)	0.44	0.35(0.94)	0.71	0.63(0.42)	0.14
Listening c 8 (WOLD)	0.06(0.09)	0.52	0.74(0.99)	0.45	0.40(0.82)	0.63	0.40(0.32)	0.20
NW repetition 8 (CNRep)	0.20(0.16)	0.21	0.54(0.81)	0.50	0.17(0.53)	0.75	0.58(0.40)	0.15
VIQ 8 (WISC-III)	0.16(0.11)	0.14	0.84(0.11)	$<1x10^{-10}$	0.27(0.20)	0.17	0.73(0.20)	$2x10^{-4}$
PIQ 8 (WISC-III)	0.01(0.05)	0.78	0.99(0.04)	$<1x10^{-10}$	0.59(0.26)	0.02	0.41(0.26)	0.11

Factorial co-heritabilities reflect the proportion of total genetic variance explained by a specific genetic factor. SEs were derived using the Delta method and P-values based on a Wald-test assuming normality. PIQ at 8 years was assessed for sensitivity analyses only. # Proportion of genetic influences for expressive vocabulary including those shared with receptive vocabulary (A1, forward GSEM model, Figure S3) with respect to the total LRA SNP-h2: a31 * a31/ (a31 * a31 + a32 * a32 + a33 * a33). ## Proportion of genetic influences for receptive vocabulary independent of expressive vocabulary (A2, forward GSEM model, Figure S3) with respect to the total LRA SNP-h2: a32 * a32/ (a31 * a31 + a32 * a32 + a33 * a33). ### Proportion of genetic influences for expressive vocabulary independent of receptive vocabulary (A2, reverse GSEM model, Figure S5) with respect to the total LRA SNP-h2: a32 * a32/ (a31 * a31 + a32 * a32 + a33 * a33). #### Proportion of genetic influences for receptive vocabulary including those shared with expressive vocabulary (A1, reverse GSEM model, Figure S5) with respect to the total LRA SNP-h2: a32 * a32/ (a31 * a31 + a32 * a32 + a33 * a33). The experiment-wide threshold is P≤0.005. Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CNRep, Children's Test of Nonword Repetition; GSEM, Genetic-relationship-matrix Structural Equation modelling; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; PIQ; performance intelligence quotient; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

Table S4: Meta-analysis across pre-defined language- and literacy-related ability combinations

Path		All LRAs (N=13)			Reading (N=7)			Spelling (N=2)		
		Coefficient (SE)	*P*	P_{het}	Coefficient (SE)	*P*	P_{het}	Coefficient (SE)	*P*	P_{het}
Genetic influences	a_{11}	0.42(0.05)	$<1x10^{-10}$	1.00	0.42(0.06)	$<1x10^{-10}$	1.00	0.42(0.06)	$<1x10^{-10}$	0.97
	a_{21}	0.29(0.06)	$1x10^{-6}$	1.00	0.29(0.07)	$2x10^{-5}$	1.00	0.29(0.08)	$2x10^{-4}$	0.98
	a_{31}	0.20(0.08)	$9x10^{-3}$	0.70	0.19(0.09)	0.03	0.70	0.19(0.10)	0.04	0.54
	a_{22}	0.19(0.05)	$3x10^{-5}$	1.00	0.19(0.05)	$3x10^{-5}$	1.00	0.17(0.07)	$8x10^{-3}$	0.99
	a_{32}	0.62(0.06)	$<1x10^{-10}$	0.97	0.62(0.06)	$<1x10^{-10}$	0.94	0.52(0.13)	$1x10^{-4}$	0.50
	a_{33}	0.34(0.29)	0.24	1.00	0.37(0.29)	0.20	0.95	0.38(0.32)	0.23	0.65
Residual influences	e_{11}	0.91(0.02)	$<1x10^{-10}$	1.00	0.91(0.03)	$<1x10^{-10}$	1.00	0.91(0.03)	$<1x10^{-10}$	0.97
	e_{21}	0.56(0.03)	$<1x10^{-10}$	1.00	0.56(0.03)	$<1x10^{-10}$	1.00	0.56(0.04)	$<1x10^{-10}$	0.98
	e_{31}	0.08(0.04)	0.02	0.62	0.09(0.04)	0.03	0.25	0.07(0.04)	0.11	0.98
	e_{22}	0.75(0.01)	$<1x10^{-10}$	1.00	0.75(0.01)	$<1x10^{-10}$	1.00	0.76(0.02)	$<1x10^{-10}$	0.99
	e_{32}	0.03(0.03)	0.36	0.91	0.04(0.04)	0.23	0.83	0.02(0.04)	0.67	0.58
	e_{33}	0.78(0.03)	$<1x10^{-10}$	$8x10^{-4}$	0.76(0.04)	$<1x10^{-10}$	0.09	0.80(0.04)	$<1x10^{-10}$	0.25

Absolute path coefficients for 13 structural equation models (forward GSEM, Figure S3) were meta-analysed across pre-defined domains (Table S1) accounting for interrelatedness between traits. Path coefficients were considered significant if they passed the experiment-wide *P*-value threshold of 0.005. Heterogeneity among effect estimates was assessed using Cochran's Q-test. Genetic and residual path coefficients follow the schematic Cholesky decomposition model depicted in Figure S3a. Abbreviations: LRAs, literacy- and language-related abilities; P_{het}, P-value for the test of heterogeneity

Table S5: Bivariate heritability estimates

Measure	Expressive vocabulary 38 months		Receptive vocabulary 38 months	
	Bivariate heritability (SE)	*P*	Bivariate heritability (SE)	*P*
Reading a/c 7 (WORD)	0.39(0.25)	0.12	0.90(0.21)	$2x10^{-5}$
Reading a 9 (NBO)	0.54(0.26)	0.04	0.71(0.24)	0.003
Reading a 9 (NARA II)	0.71(0.28)	0.01	1.00(0.22)#	$9x10^{-7}$
Reading s 9 (NARA II)	0.69(0.38)	0.07	1.00(0.27)#	$8x10^{-10}$
Reading s 13 (TOWRE)	0.59(0.38)	0.12	1.00(0.31)#	$3x10^{-4}$
NW reading a 9 (NBO)	0.40(0.28)	0.16	0.71(0.28)	0.01
NW reading s 13 (TOWRE)	0.48(0.43)	0.26	1.00(0.34)*	$8x10^{-4}$
Spelling a 7 (NB)	0.59(0.35)	0.09	0.95(0.26)	$3x10^{-4}$
Spelling a 9 (NB)	0.62(0.31)	0.04	0.79(0.26)	0.002
PhonAware 7 (AAT)	0.63(0.26)	0.02	0.92(0.24)	$1x10^{-4}$
Listening c 8 (WOLD)	0.49(0.39)	0.20	0.77(0.28)	0.006
Non-word repetition 8 (CNRep)	0.53(0.23)	0.02	0.76(0.23)	0.001
VIQ 8 (WISC-III)	0.69(0.24)	0.005	0.91(0.16)	$3x10^{-8}$
PIQ 8 (WISC-III)	0.24(0.45)	0.59	0.75(0.27)	0.006

Bivariate heritability estimates, reflecting the proportion of the phenotypic covariance that is accounted for by the genetic covariance. SEs were approximated by the SE of the genetic covariance divided by the phenotypic covariance (as the SE of the phenotypic covariance is small) and *P*-values are based on a Wald-test, assuming normality. Estimates are based on forward GSEM models, and reverse GSEM models provided nearly identical results (data not shown). PIQ at 8 years was assessed for sensitivity analyses only. The experiment-wide threshold is $P \leq 0.005$. # Estimates were truncated at one. Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CNRep, Children's Test of Nonword Repetition; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; PIQ; performance intelligence quotient; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

Supplementary Figures

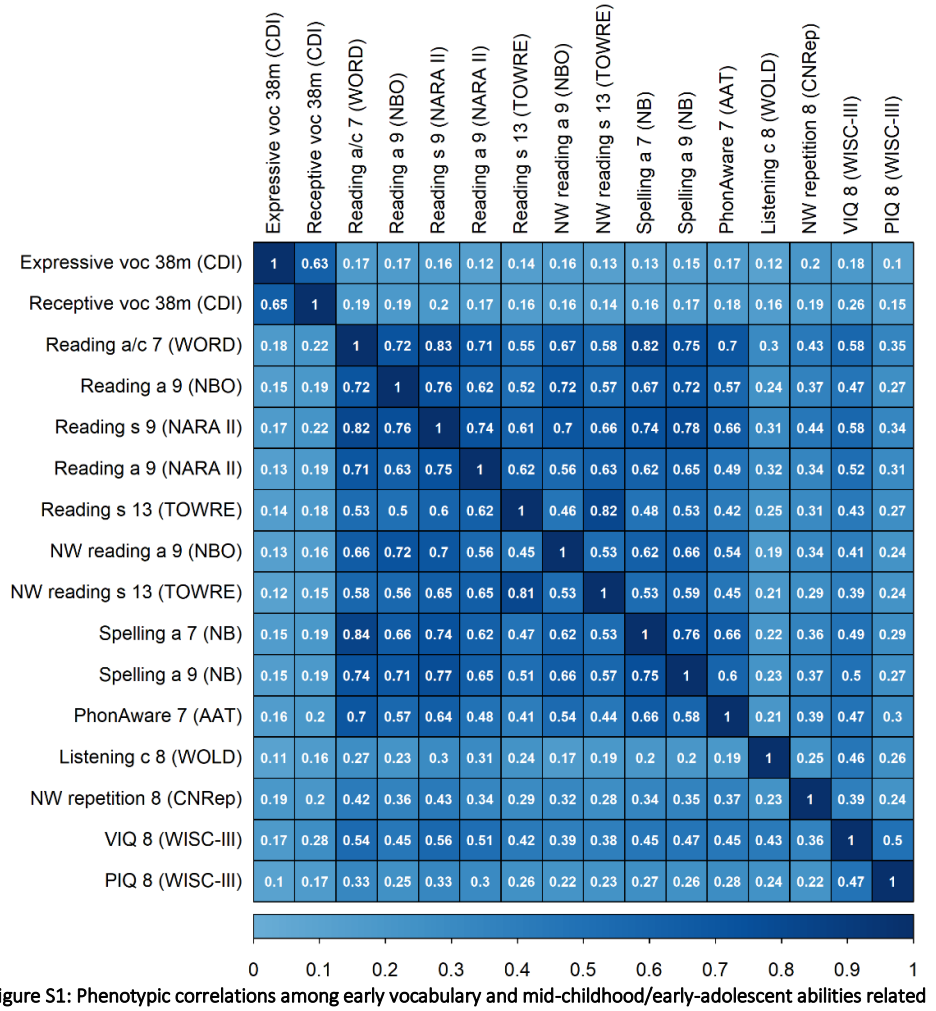
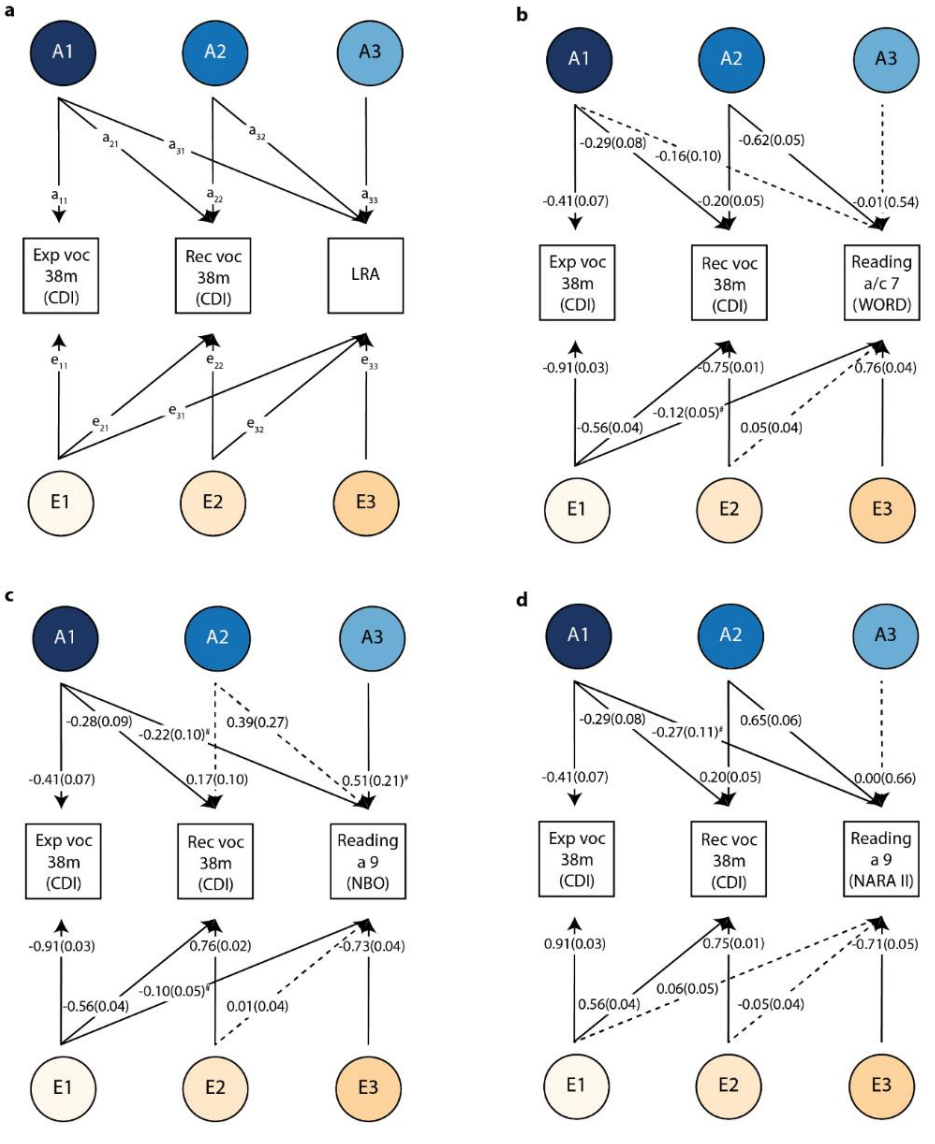
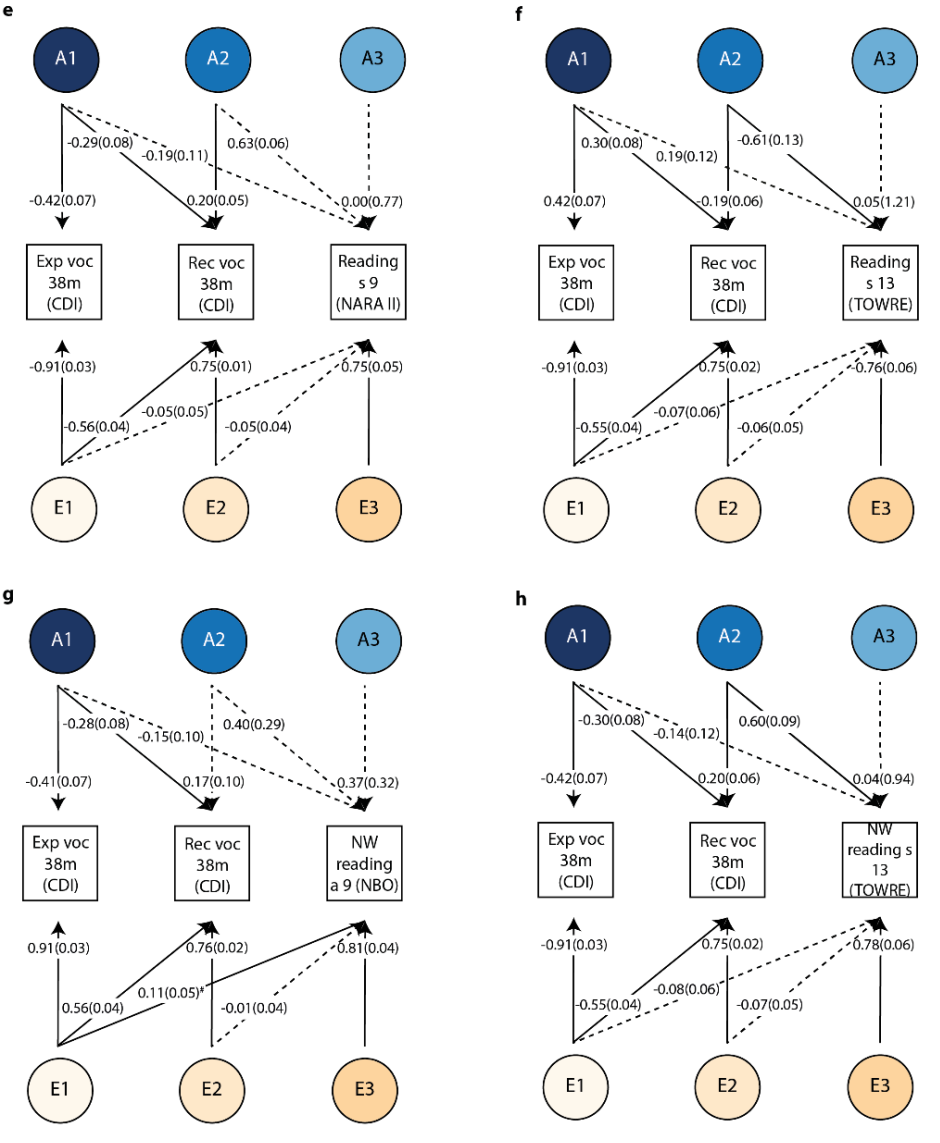
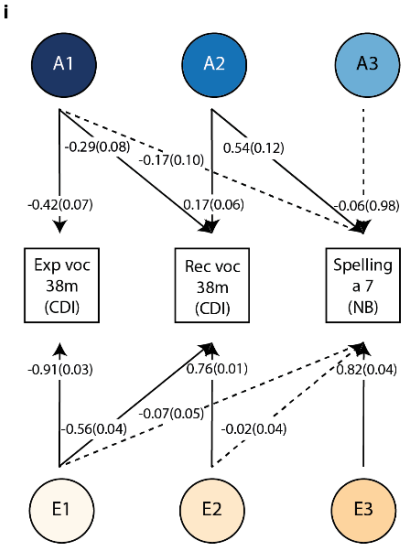
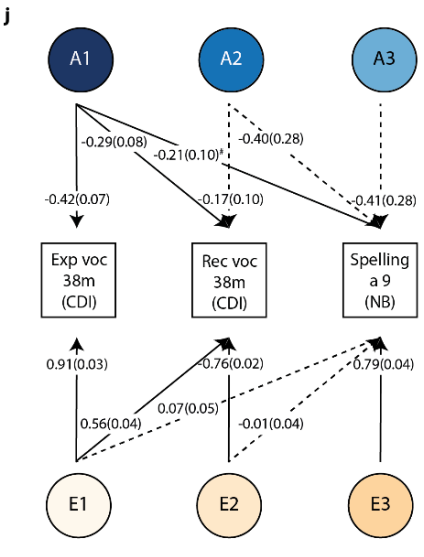


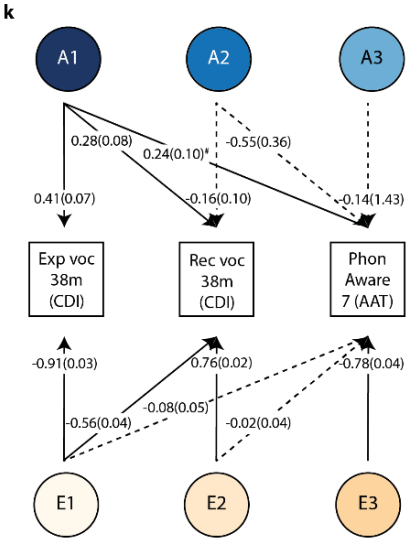
Figure S1: Phenotypic correlations among early vocabulary and mid-childhood/early-adolescent abilities related to literacy, language and cognition. Phenotypic correlations among untransformed measures are represented in the lower triangle and were estimated with Spearman's rank correlation coefficients. Phenotypic correlations among transformed measures are represented in the upper triangle and were estimated with Pearson correlation coefficients. Only correlation coefficients passing the experiment-wide significance threshold ($P \leq 0.005$) are shown. PIQ at 8 years was assessed for sensitivity analyses only. Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CDI, Communicative Development Inventory; CNRep, Children's Test of Nonword Repetition; m, months; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; PIQ, performance intelligence quotient; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

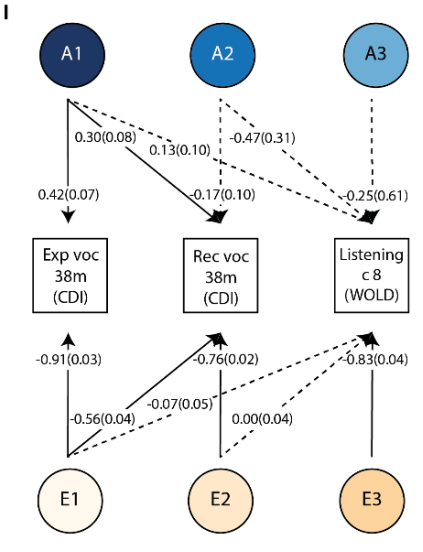

a
A1
A2
A3
a21
a31
a32
a11
a22
a33
Exp voc 38m (CDI)
Rec voc 38m (CDI)
LRA
e11
e22
e33
e21
e31
e32
E1
E2
E3
b
A1
A2
A3
-0.29(0.08)
-0.16(0.10)
-0.62(0.05)
-0.41(0.07)
-0.20(0.05)
-0.01(0.54)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Reading a/c 7 (WORD)
-0.91(0.03)
-0.75(0.01)
0.76(0.04)
-0.56(0.04)
-0.12(0.05)#
0.05(0.04)
E1
E2
E3
c
A1
A2
A3
-0.28(0.09)
-0.22(0.10)#
0.39(0.27)
-0.41(0.07)
0.17(0.10)
0.51(0.21)#
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Reading a 9 (NBO)
-0.91(0.03)
0.76(0.02)
-0.73(0.04)
-0.56(0.04)
-0.10(0.05)#
0.01(0.04)
E1
E2
E3
d
A1
A2
A3
-0.29(0.08)
-0.27(0.11)#
0.65(0.06)
-0.41(0.07)
0.20(0.05)
0.00(0.66)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Reading a 9 (NARA II)
0.91(0.03)
0.75(0.01)
-0.71(0.05)
0.56(0.04)
0.06(0.05)
-0.05(0.04)
E1
E2
E3


e
A1
A2
A3
-0.29(0.08)
-0.19(0.11)
0.63(0.06)
-0.42(0.07)
0.20(0.05)
0.00(0.77)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Reading s 9 (NARA II)
-0.91(0.03)
0.75(0.01)
0.75(0.05)
-0.56(0.04)
-0.05(0.05)
-0.05(0.04)
E1
E2
E3
f
A1
A2
A3
0.30(0.08)
0.19(0.12)
-0.61(0.13)
0.42(0.07)
-0.19(0.06)
0.05(1.21)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Reading s 13 (TOWRE)
-0.91(0.03)
0.75(0.02)
0.76(0.06)
-0.55(0.04)
-0.07(0.06)
-0.06(0.05)
E1
E2
E3
g
A1
A2
A3
-0.28(0.08)
-0.15(0.10)
0.40(0.29)
-0.41(0.07)
0.17(0.10)
0.37(0.32)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
NW reading a 9 (NBO)
0.91(0.03)
0.76(0.02)
0.81(0.04)
0.56(0.04)
0.11(0.05)*
-0.01(0.04)
E1
E2
E3
h
A1
A2
A3
-0.30(0.08)
-0.14(0.12)
0.60(0.09)
-0.42(0.07)
0.20(0.06)
0.04(0.94)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
NW reading s 13 (TOWRE)
-0.91(0.03)
0.75(0.02)
0.78(0.06)
-0.55(0.04)
-0.08(0.06)
-0.07(0.05)
E1
E2
E3


i
A1
A2
A3
-0.29(0.08)
-0.17(0.10)
0.54(0.12)
-0.42(0.07)
0.17(0.06)
-0.06(0.98)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Spelling a 7 (NB)
-0.91(0.03)
0.76(0.01)
0.82(0.04)
-0.56(0.04)
-0.07(0.05)
-0.02(0.04)
E1
E2
E3


j
A1
A2
A3
-0.29(0.08)
-0.21(0.10)*
-0.40(0.28)
-0.42(0.07)
-0.17(0.10)
-0.41(0.28)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Spelling a 9 (NB)
0.91(0.03)
-0.76(0.02)
0.79(0.04)
0.56(0.04)
0.07(0.05)
-0.01(0.04)
E1
E2
E3


k
A1
A2
A3
0.28(0.08)
0.24(0.10)*
-0.55(0.36)
0.41(0.07)
-0.16(0.10)
-0.14(1.43)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Phon Aware 7 (AAT)
-0.91(0.03)
0.76(0.02)
0.78(0.04)
-0.56(0.04)
-0.08(0.05)
-0.02(0.04)
E1
E2
E3


l
A1
A2
A3
0.30(0.08)
0.13(0.10)
-0.47(0.31)
0.42(0.07)
-0.17(0.10)
-0.25(0.61)
Exp voc 38m (CDI)
Rec voc 38m (CDI)
Listening c 8 (WOLD)
-0.91(0.03)
-0.76(0.02)
0.83(0.04)
-0.56(0.04)
-0.07(0.05)
0.00(0.04)
E1
E2
E3

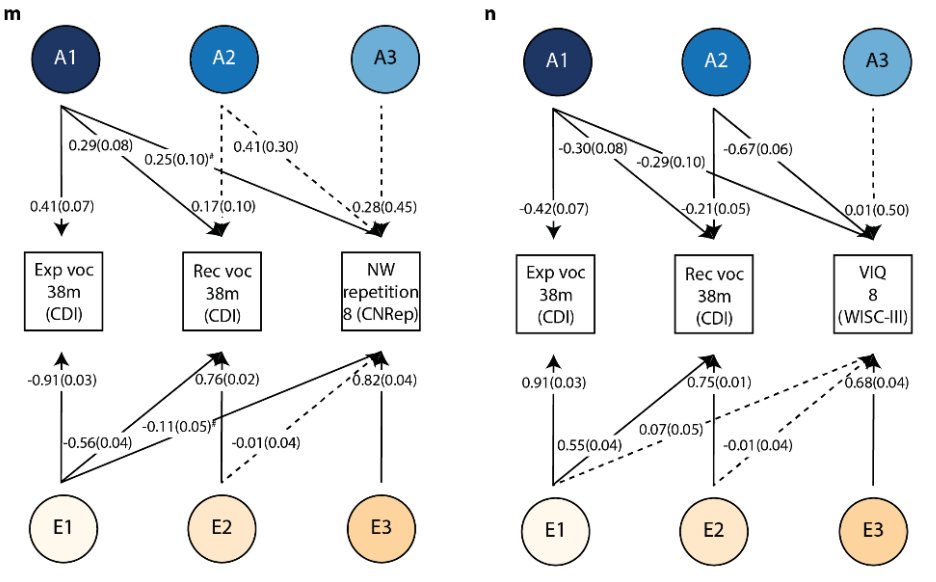


Figure S2: Path models of early vocabulary and mid-childhood/early-adolescent literacy- and language-related abilities (forward GSEM). Cholesky decompositions were fitted using GSEM, according to forward GSEMs and based on all available observations for children across development (N≤6,092). **(a)** Schematic path model with path coefficient labels for a Cholesky decomposition model of vocabulary at 38 months, including expressive and receptive vocabulary (in that order), and one later LRA. **(b-n)** Path models of standardised path coefficients and corresponding standard errors for 13 forward GSEM models, one for each fitted LRA. Solid lines indicate path coefficients passing a *P*-value threshold of $P \leq 0.05$, dashed lines indicate non-significant path coefficients $P > 0.05$. # Path coefficient passing nominal significance ($P \leq 0.05$), but not the experiment-wide significance threshold ($P \leq 0.005$). Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CDI, Communicative Development Inventory; CNRep, Children's Test of Nonword Repetition; Exp, expressive; LRA, language- and literacy-related ability; m, months; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; Rec, receptive; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

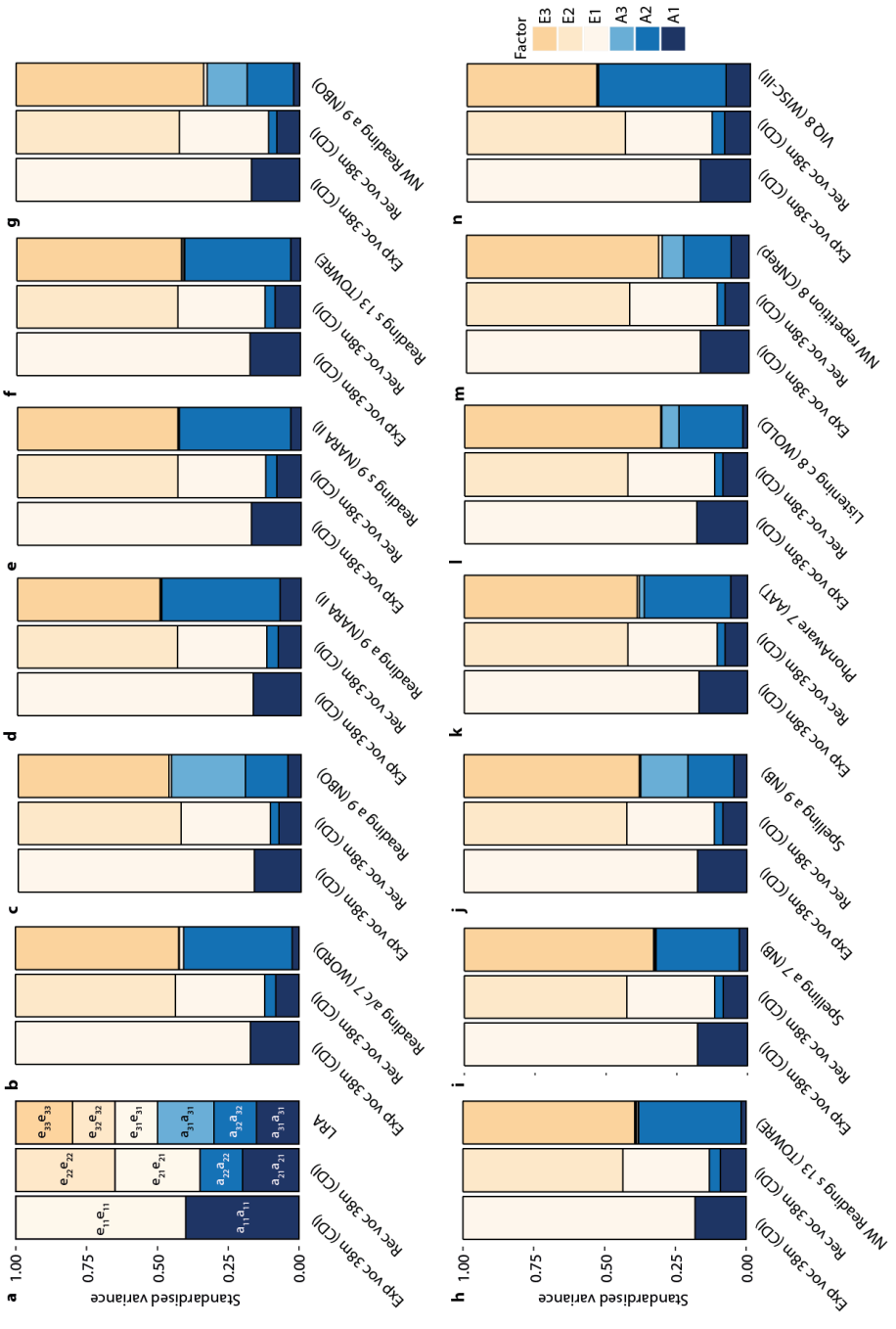
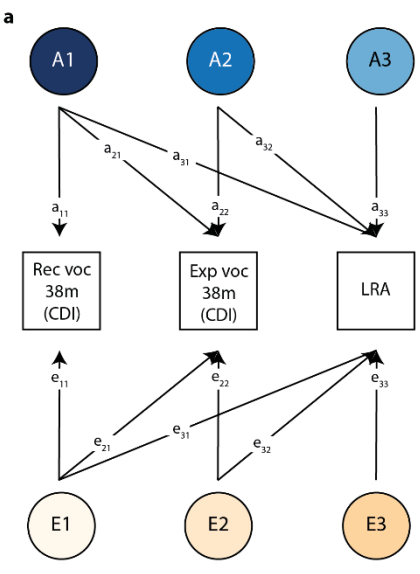
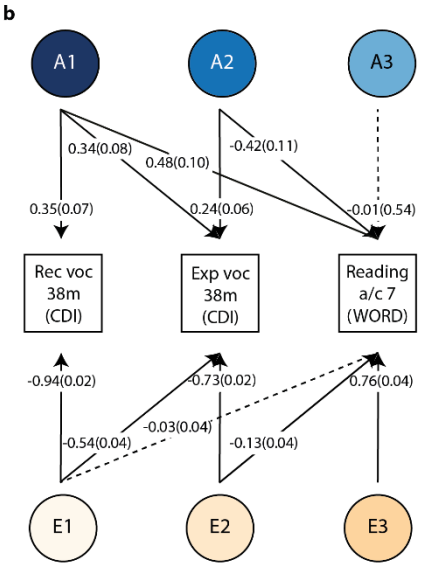
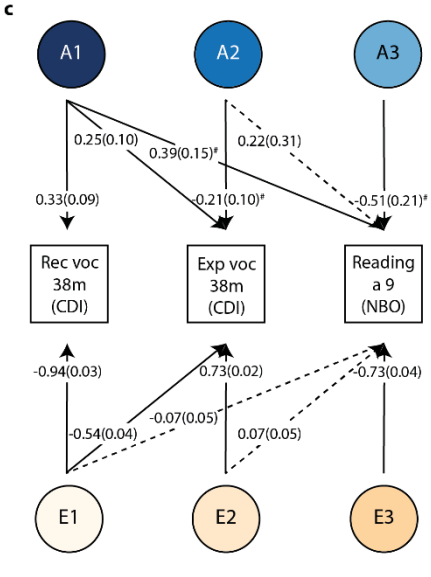
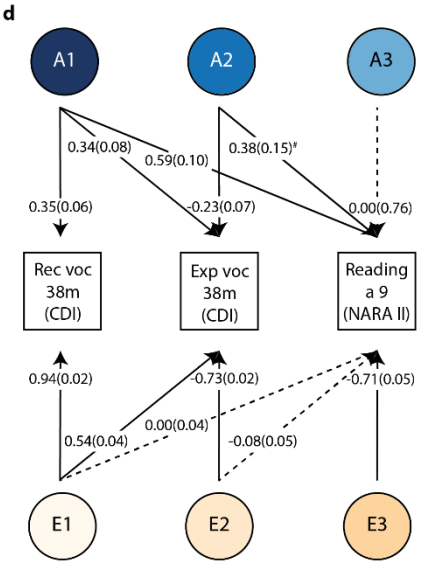

Factor
E3
E2
E1
A3
A2
A1
Standardised variance
1.00
0.75
0.50
0.25
0.00
a
b
c
d
e
f
g
h
i
j
k
l
m
n
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
LRA
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Reading a/c 7 (WORD)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Reading a 9 (NBO)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Reading a 9 (NARA II)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Reading s 9 (NARA II)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Reading s 13 (TOWRE)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
NW Reading a 9 (NBO)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
NW Reading s 13 (TOWRE)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Spelling a 7 (NB)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Spelling a 9 (NB)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
PhonAware 7 (AAT)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Listening c 8 (WOLD)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
NW repetition 8 (CNRep)
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
VIQ 8 (WISC-III)

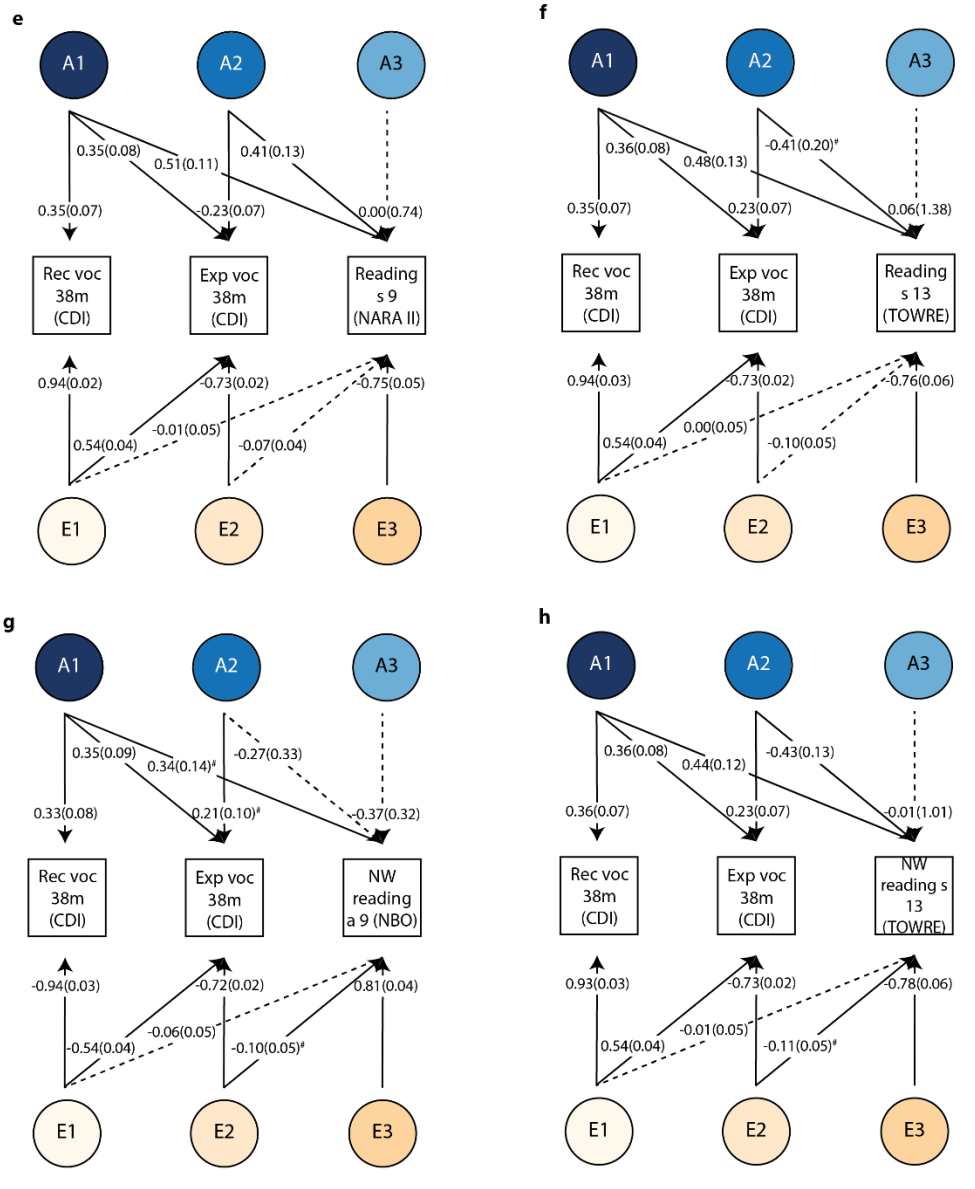
Figure S3: Variance plots for path models of early vocabulary and mid-childhood/early-adolescent literacy- and language-related abilities (forward GSEM). Standardised variance explained by genetic and residual factors as derived by Cholesky decompositions using forward GSEM (Figure S3), based on all available observations for children across development (N≤6,092). **(a)** Variance plot with path coefficient labels for a Cholesky decomposition model of vocabulary at 38 months, including expressive and receptive vocabulary (in that order), and one later LRA. **(b-n)** Standardised variance explained by genetic and residual factors as modelled in 13 forward GSEM models, one for each fitted LRA. Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CDI, Communicative Development Inventory; CNRep, Children's Test of Nonword Repetition; Exp, expressive; LRA, language- and literacy-related ability; m, months; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; Rec, receptive; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

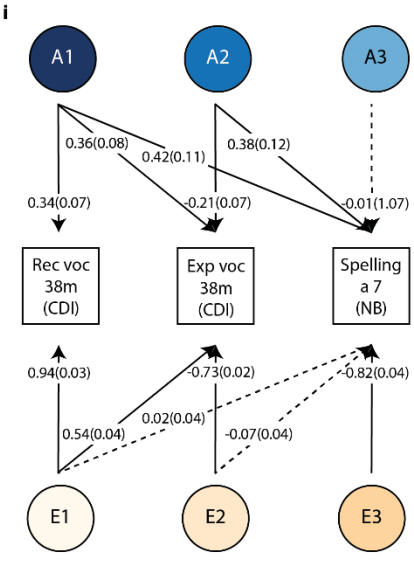

a
A1
A2
A3
a21
a31
a32
a11
a22
a33
Rec voc 38m (CDI)
Exp voc 38m (CDI)
LRA
e11
e22
e33
e21
e31
e32
E1
E2
E3

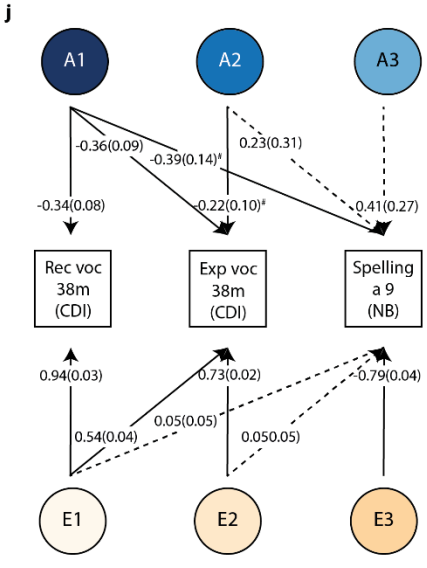

b
A1
A2
A3
0.34(0.08)
0.48(0.10)
-0.42(0.11)
0.35(0.07)
0.24(0.06)
-0.01(0.54)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Reading a/c 7 (WORD)
-0.94(0.02)
-0.73(0.02)
0.76(0.04)
-0.03(0.04)
-0.54(0.04)
-0.13(0.04)
E1
E2
E3

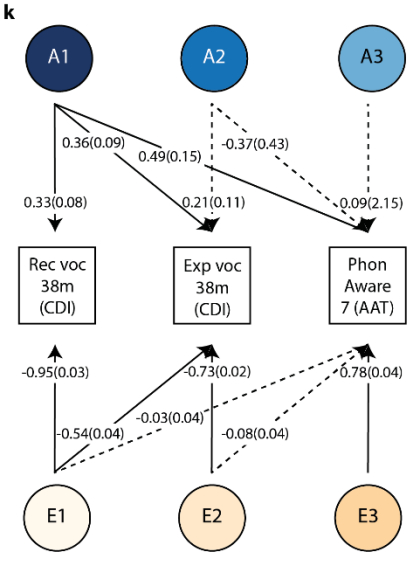

c
A1
A2
A3
0.25(0.10)
0.39(0.15)#
0.22(0.31)
0.33(0.09)
-0.21(0.10)#
-0.51(0.21)#
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Reading a 9 (NBO)
-0.94(0.03)
0.73(0.02)
-0.73(0.04)
-0.07(0.05)
-0.54(0.04)
0.07(0.05)
E1
E2
E3

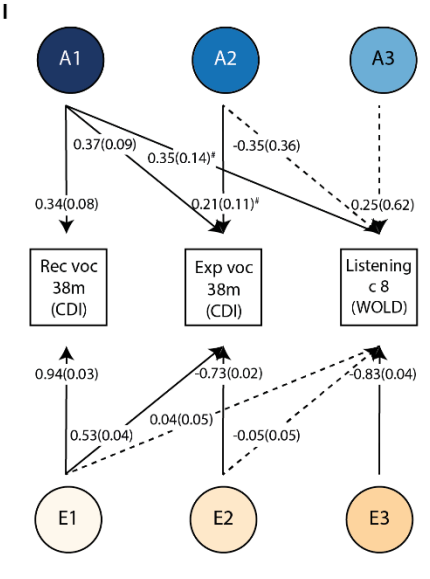

d
A1
A2
A3
0.34(0.08)
0.59(0.10)
0.38(0.15)#
0.35(0.06)
-0.23(0.07)
0.00(0.76)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Reading a 9 (NARA II)
0.94(0.02)
-0.73(0.02)
-0.71(0.05)
0.00(0.04)
0.54(0.04)
-0.08(0.05)
E1
E2
E3


e
A1
A2
A3
0.35(0.08)
0.51(0.11)
0.41(0.13)
0.35(0.07)
-0.23(0.07)
0.00(0.74)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Reading s 9 (NARA II)
0.94(0.02)
-0.73(0.02)
-0.75(0.05)
0.54(0.04)
-0.01(0.05)
-0.07(0.04)
E1
E2
E3
f
A1
A2
A3
0.36(0.08)
0.48(0.13)
-0.41(0.20)#
0.35(0.07)
0.23(0.07)
0.06(1.38)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Reading s 13 (TOWRE)
0.94(0.03)
-0.73(0.02)
-0.76(0.06)
0.54(0.04)
0.00(0.05)
-0.10(0.05)
E1
E2
E3
g
A1
A2
A3
0.35(0.09)
0.34(0.14)#
-0.27(0.33)
0.33(0.08)
0.21(0.10)#
-0.37(0.32)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
NW reading a 9 (NBO)
-0.94(0.03)
-0.72(0.02)
0.81(0.04)
-0.54(0.04)
-0.06(0.05)
-0.10(0.05)#
E1
E2
E3
h
A1
A2
A3
0.36(0.08)
0.44(0.12)
-0.43(0.13)
0.36(0.07)
0.23(0.07)
-0.01(1.01)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
NW reading s 13 (TOWRE)
0.93(0.03)
-0.73(0.02)
-0.78(0.06)
0.54(0.04)
-0.01(0.05)
-0.11(0.05)#
E1
E2
E3


i
A1
A2
A3
0.36(0.08)
0.42(0.11)
0.38(0.12)
0.34(0.07)
-0.21(0.07)
-0.01(1.07)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Spelling a 7 (NB)
0.94(0.03)
-0.73(0.02)
-0.82(0.04)
0.54(0.04)
0.02(0.04)
-0.07(0.04)
E1
E2
E3


j
A1
A2
A3
-0.36(0.09)
-0.39(0.14)#
0.23(0.31)
-0.34(0.08)
-0.22(0.10)#
0.41(0.27)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Spelling a 9 (NB)
0.94(0.03)
0.73(0.02)
-0.79(0.04)
0.54(0.04)
0.05(0.05)
0.050.05)
E1
E2
E3


k
A1
A2
A3
0.36(0.09)
0.49(0.15)
-0.37(0.43)
0.33(0.08)
0.21(0.11)
0.09(2.15)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Phon Aware 7 (AAT)
-0.95(0.03)
-0.73(0.02)
0.78(0.04)
-0.54(0.04)
-0.03(0.04)
-0.08(0.04)
E1
E2
E3


l
A1
A2
A3
0.37(0.09)
0.35(0.14)#
-0.35(0.36)
0.34(0.08)
0.21(0.11)#
0.25(0.62)
Rec voc 38m (CDI)
Exp voc 38m (CDI)
Listening c 8 (WOLD)
0.94(0.03)
-0.73(0.02)
-0.83(0.04)
0.53(0.04)
0.04(0.05)
-0.05(0.05)
E1
E2
E3

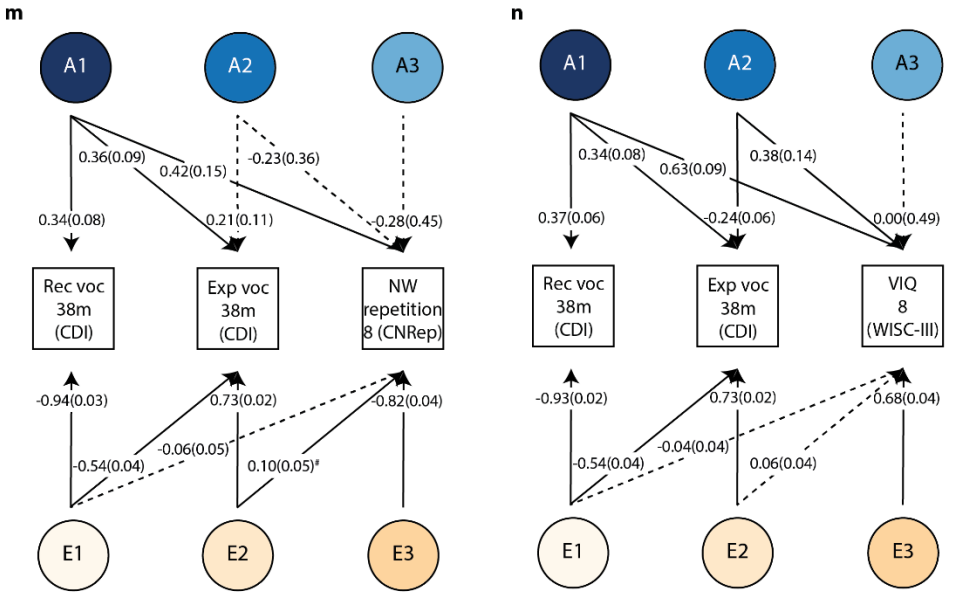


Figure S4: Path models of early vocabulary and mid-childhood/early-adolescent literacy- and language-related abilities (reverse GSEM). Cholesky decompositions were fitted using GSEM, according to reverse GSEMs and based on all available observations for children across development (N≤6,092). **(a)** Schematic path model with path coefficient labels for a Cholesky decomposition model of vocabulary at 38 months, including receptive and expressive vocabulary (in that order), and one later LRA. **(b-n)** Path models of standardised path coefficients and corresponding standard errors for 13 reverse GSEM models, one for each fitted LRA. Solid lines indicate path coefficients passing a *P*-value threshold of $P \leq 0.05$, dashed lines indicate non-significant path coefficients $P > 0.05$. # Path coefficient passing nominal significance ($P \leq 0.05$), but not the experiment-wide significance threshold ($P \leq 0.005$). Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CDI, Communicative Development Inventory; CNRep, Children's Test of Nonword Repetition; Exp, expressive; LRA, language- and literacy-related ability; m, months; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; Rec, receptive; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

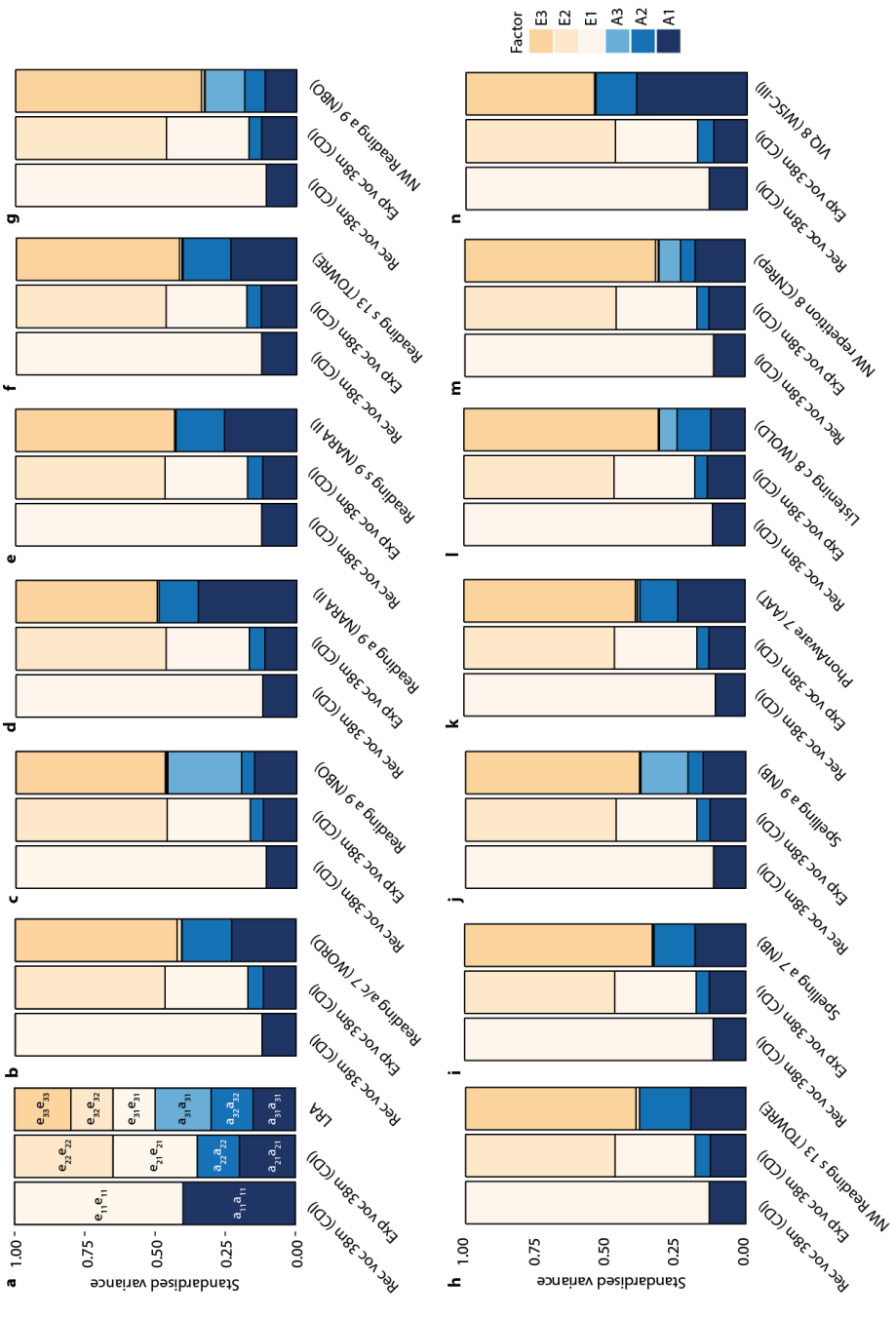

Factor
E3
E2
E1
A3
A2
A1
Standardised variance
1.00
0.75
0.50
0.25
0.00
a
b
c
d
e
f
g
h
i
j
k
l
m
n
LRA
Exp voc-3bm (CDI)
Rec voc-3bm (CDI)
Reading a/c 7 (WORD)
Reading a 9 (NB0)
Reading a 9 (NARA II)
Reading s 13 (TOWRE)
Reading s 9 (NARA II)
NW Reading a 9 (NB0)
Speiling a 7 (NB)
Spelling a 9 (NB)
NW Reading s 13 (TOWRE)
Phonaware 7 (AAT)
Listening c 8 (WOLD)
NW repetition 8 (CNRep)
VIQ 8 (WISC-III)

Figure S5: Variance plots for path models of early vocabulary and mid-childhood/early-adolescent literacy- and language-related abilities (reverse GSEM). Standardised variance explained by genetic and residual factors as derived by Cholesky decompositions using reverse GSEM (Figure S5), based on all available observations for children across development (N≤6,092). **(a)** Variance plot with path coefficient labels for a Cholesky decomposition model of vocabulary at 38 months, including receptive and expressive vocabulary (in that order), and one later LRA. **(b-n)** Standardised variance explained by genetic and residual factors as modelled in 13 reverse GSEM models, one for each fitted LRA. Abbreviations: a, accuracy; AAT, Auditory Analysis Test; c, comprehension; CDI, Communicative Development Inventory; CNRep, Children's Test of Nonword Repetition; Exp, expressive; LRA, language- and literacy-related ability; m, months; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; NB, ALSPAC-specific assessment developed by Nunes and Bryant; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NW, nonword; PhonAware, phonemic awareness; Rec, receptive; s, speed; TOWRE, Test Of Word Reading Efficiency; VIQ, verbal intelligence quotient; voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WOLD, Wechsler Objective Language Dimensions; WORD, Wechsler Objective Reading Dimension

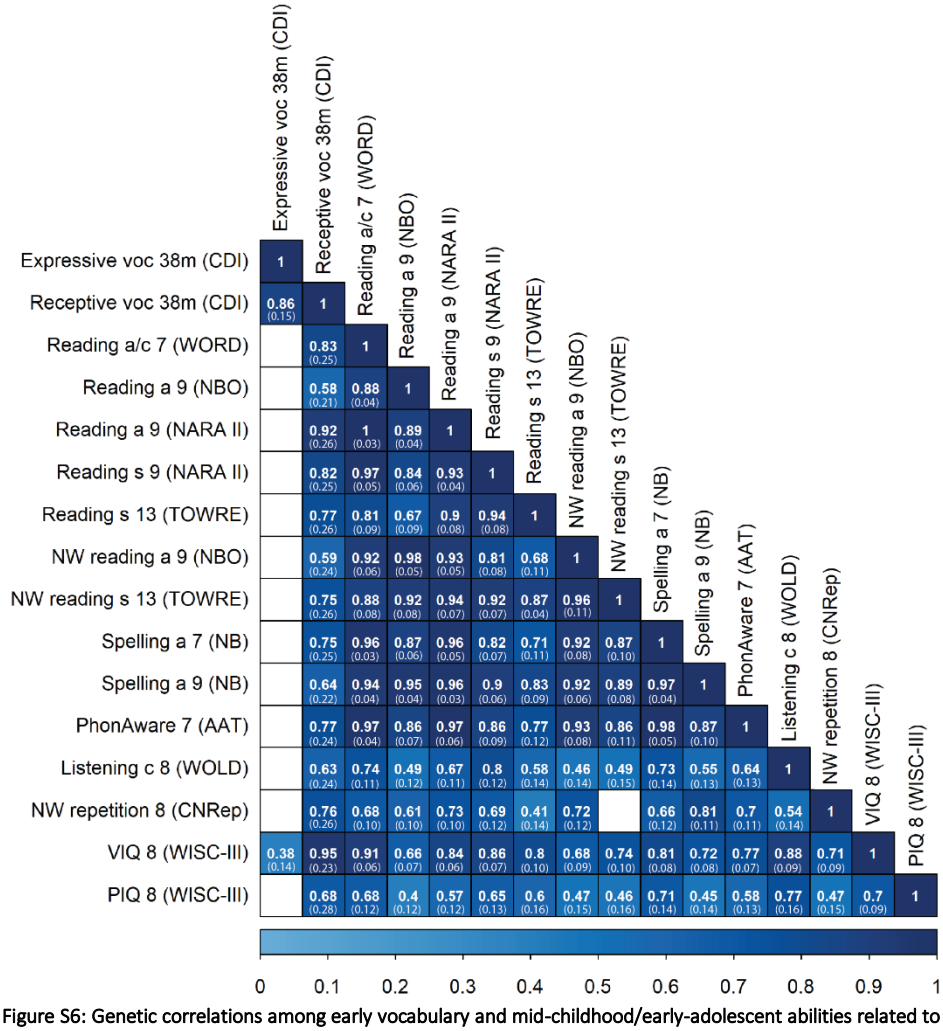


Figure S6: Genetic correlations among early vocabulary and mid-childhood/early-adolescent abilities related to literacy, language and cognition. Genetic correlations were calculated based on rank-transformed scores using Restricted Maximum Likelihood (REML) analyses as implemented in genome-wide complex trait analysis (GCTA) software, based on directly genotyped SNPs and unrelated individuals (genetic relationship of <0.05). Only genetic correlations that passed the experiment-wide significance threshold ($P \leq 0.005$) are shown. Corresponding standard errors are provided between brackets. PIQ at 8 years was assessed for sensitivity analyses only. Abbreviations are shown in Figure S1.

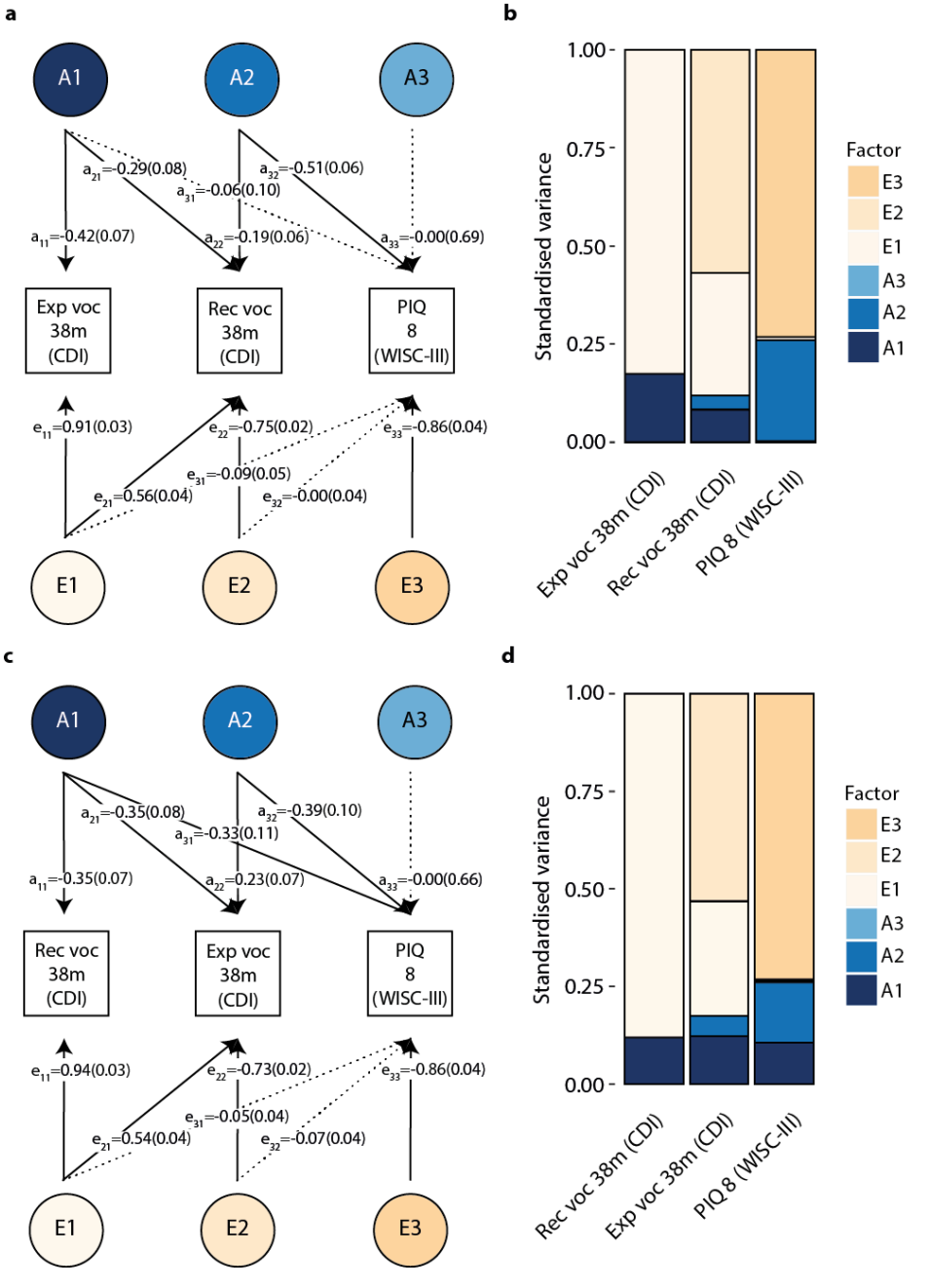


Figure S7: Path model and variance plot for early vocabulary and mid-childhood performance intelligence. A Cholesky decomposition was fitted using GSEM, according to **(a,b)** forward GSEM and **(c,d)** reverse GSEM, based on all available observations for children across development (N≤6,092). **(a)** Path model of standardised path coefficients and corresponding standard errors for a Cholesky decomposition of vocabulary at 38 months, including expressive and receptive vocabulary (in that order), and performance intelligence scores at 8 years. Solid lines indicate path coefficients passing a *P*-value threshold of $P \leq 0.05$, dashed lines indicate non-significant

path coefficients (*P*>0.05). **(b)** Standardised variance explained by genetic and residual factors modelled in a. **(c)** Path model of standardised path coefficients and corresponding standard errors for a Cholesky decomposition of vocabulary at 38 months, including receptive and expressive vocabulary (in that order), and performance intelligence scores at 8 years. Solid lines indicate path coefficients passing a *P*-value threshold of *P*≤0.05, dashed lines indicate non-significant path coefficients (*P*>0.05). **(d)** Standardised variance explained by genetic and residual factors modelled in c. Abbreviations: CDI, Communicative Development Inventory; Exp, expressive; m, months; Rec, receptive; voc, vocabulary; PIQ, performance intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III

Supplementary References

1. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111–127 (2013).
2. Marees, A. T. *et al.* A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International Journal of Methods in Psychiatric Research* **27**, e1608 (2018).
3. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* **81**, 559–575 (2007).
4. *WORD, Wechsler Objective Reading Dimensions Manual*. (Psychological Corporation, 1993).
5. Elliott, C. D., Salerno, J. D., Dumont, R. & Willis, J. O. The Differential Ability Scales—Second Edition. in *Contemporary intellectual assessment: Theories, tests, and issues, 4th ed* 360–382 (The Guilford Press, 2018).
6. Nunes, T., Bryant, P. & Olsson, J. Learning Morphological and Phonological Spelling Rules: An Intervention Study. *Scientific Studies of Reading* **7**, 289–307 (2003).
7. Schonell, F. J. & Goodacre, E. The Psychology and Teaching of Reading. Fifth Edition. (1974).
8. Neale, M. D. *Neale Analysis of Reading Ability: Second Revised British Edition*. (NFER-Nelson, 1997).
9. Boyle, J. & Fisher, S. *Educational Testing: A Competence-Based Approach*. (John Wiley & Sons, 2008).
10. *TOWRE, Test of Word Reading Efficiency: Examiner's Manual*. (PRO-ED, 1999).
11. Rosner, J. & Simon, D. P. The Auditory Analysis Test: An Initial Report. *Journal of Learning Disabilities* **4**, 384–392 (1971).
12. Rosner, J. *TAAS: Test of Auditory Analysis Skills*. (Academic Therapy Publications, 1975).
13. Yopp, H. K. The validity and reliability of phonemic awareness tests. *Reading Research Quarterly* **23**, 159–177 (1988).
14. Rust, J. & Wechsler, D. *WOLD: Wechsler objective language dimensions.* (Psychological Corp. ; Artsberg Enterprises Ltd. [distributor, 1996).
15. Gathercole, S. E., Alloway, T. P., Willis, C. & Adams, A.-M. Working memory in children with reading disabilities. *J Exp Child Psychol* **93**, 265–281 (2006).
16. Dunn, L. M., Dunn, L. M., Bulheller, S. & Häcker, H. *Peabody picture vocabulary test*. (American Guidance Service Circle Pines, MN, 1965).
17. Rathvon, N. *Early Reading Assessment: A Practitioner's Handbook*. (Guilford Press, 2004).
18. Gathercole, S. E., Willis, C. S., Baddeley, A. D. & Emslie, H. The Children's Test of Nonword Repetition: a test of phonological working memory. *Memory* **2**, 103–127 (1994).
19. Wechsler, D., Golombok, S. & Rust, J. *WISC-III UK Wechsler Intelligence Scale for Children – Third Edition UK Manual*. (The Psychological Corporation, 1992).
20. Sattler, J. M. *Assessment of children: Cognitive applications, 4th ed*. (Jerome M Sattler Publisher, 2001).
21. Canivez, G. L. & Watkins, M. W. Long-term stability of the Wechsler Intelligence Scale for Children—Third Edition. *Psychological Assessment* **10**, 285–291 (1998).
22. Guilmette, T. J., Kennedy, M. L. & Queally, P. T. A Comparison of the WISC-III and the Otis-Lennon School Ability Test with Students Referred for Learning Disabilities. *Journal of Psychoeducational Assessment* **19**, 239–244 (2001).
23. Cherny, S. S. Cholesky Decomposition. in *Encyclopedia of Statistics in Behavioral Science* (John Wiley & Sons, Ltd, 2005).
24. Neale, M. & Maes, H. H. M. *Methodology for genetic studies of twins and families.* (Kluwer Academic Publishers, 2004).
25. Martin, N. G. & Eaves, L. J. The genetical analysis of covariance structure. *Heredity (Edinb)* **38**, 79–95 (1977).
26. Falconer, P. D. S. & Mackay, P. T. F. C. *Introduction to Quantitative Genetics*. (Longman, 1995).
27. Janssens, M. J. J. Co-heritability: Its relation to correlated response, linkage, and pleiotropy in cases of polygenic inheritance. *Euphytica* **28**, 601–608 (1979).
28. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* 1–48 (2010).
29. Lajeunesse, M. J. Meta-analysis and the comparative phylogenetic method. *Am. Nat.* **174**, 369–381 (2009).

30. Visscher, P. M. *et al.* Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples. *PLOS Genetics* **10**, e1004269 (2014).

Chapter 4

The developmental genetic architecture of vocabulary skills during the first three years of life: capturing emerging associations with later-life reading and cognition

This chapter has been accepted for publication and is available as preprint:
Verhoef, E., Shapland, C.Y., Fisher, S.E., Dale, P.S., St Pourcain, B. The developmental genetic architecture of vocabulary skills during the first three years of life: capturing emerging associations with later-life reading and cognition (accepted). *PLOS Genetics*. doi: 10.1371/journal.pgen.1009144.

Abstract

Individual differences in early-life vocabulary measures are heritable and associated with subsequent reading and cognitive abilities, although the underlying mechanisms are little understood. Here, we (i) investigate the developmental genetic architecture of expressive and receptive vocabulary in toddlerhood and (ii) assess origin and developmental stage of emerging genetic associations with mid-childhood verbal and non-verbal skills.

Studying up to 6,524 unrelated children from the population-based Avon Longitudinal Study of Parents and Children (ALSPAC) cohort, we dissected the phenotypic variance of longitudinally assessed early-life vocabulary measures (15-38 months) and later-life reading and cognitive skills (7-8 years) into genetic and residual components, by fitting multivariate structural equation models to genome-wide genetic-relationship matrices.

Our findings show that the genetic architecture of early-life vocabulary is dynamic, involving multiple distinct genetic factors. Two of them are developmentally stable and contribute to genetic variation in mid-childhood skills: Genetic links with later-life verbal abilities (reading, verbal intelligence) emerged with expressive vocabulary at 24 months. The underlying genetic factor explained 10.1% variation (path coefficient: 0.32(SE=0.06)) in early language, but also 6.4% (path coefficient: 0.25(SE=0.12)) and 17.9% (path coefficient: 0.42(SE=0.13)) variation in mid-childhood reading and verbal intelligence, respectively. An independent stable genetic factor was identified for receptive vocabulary at 38 months, explaining 2.1% (path coefficient: 0.15(SE=0.07)) phenotypic variation. This genetic factor was also linked to both verbal and non-verbal cognitive abilities in mid-childhood, accounting for 24.7% of the variation in non-verbal intelligence (path coefficient: 0.50(SE=0.08)), 33.0% in reading (path coefficient: 0.57(SE=0.07)) and 36.1% in verbal intelligence (path coefficient: 0.60(0.10)), corresponding to the majority of genetic variance (≥66.4%).

Thus, the genetic foundations of mid-childhood reading and cognition are diverse. They involve at least two independent genetic factors that emerge at different stages during early language development and may implicate differences in cognitive processes that are already detectable during toddlerhood.

4.1. Introduction

The number of words produced and understood by children during the first few years of life is a rapidly changing developmental phenotype that is often used to assess the level of language acquisition[1]. One of the first precursors of expressive vocabulary (i.e. word production) in typically developing children is canonical babbling, which emerges around the age of four to six months[2], followed by the spontaneous production of first words between 10 to 15 months of age[3]. With progressing development, the number of produced words increases, reaching a median of 40 words at 16 months[1], often trailed by a period of rapid growth till the age of about 22 months[4] and a steady increase after that. This results in the production of approximately 500 words at 30 months[5] and about 2,600 words at six years of age[6]. The development of receptive vocabulary (i.e. word comprehension) typically precedes expressive vocabulary in developing children[7], with the understanding of the first few words emerging between 6 to 9 months of age[8]. Thus, receptive vocabulary is often larger than expressive vocabulary in size[7]. For example, the number of words understood by infants at 16 months of age has a median of 169 words, and is, thus, approximately 129 words larger compared to their expressive vocabulary at the same time[1]. This discrepancy increases during development, with a receptive vocabulary size of about 20,000 to 24,000 words at the age of six years, which is about six times larger than its expressive counterpart[6].

The rate of language acquisition, and thus vocabulary size, varies between children during early language development[9,10]. These large interindividual differences can partially be explained by genetic variation. Twin studies estimated that genetic influences could account for 17% to 25% of variation in expressive vocabulary at 24 months[11,12], 10% to 14% of variation in expressive vocabulary at 36 months[11] and 28% of variation in receptive vocabulary at 14 months[13]. Studies using genotype data from unrelated children provided similar estimates, with single-nucleotide polymorphism heritability (SNP-h^2) estimates of 13% to 14% for expressive vocabulary at 15 to 30 months of age[14] and 12% for receptive vocabulary at 38 months of age[15].

Despite some stable genetic contributions during early development, there is evidence for age-specific genetic influences on vocabulary skills. For example, measures of expressive vocabulary size assessed between 15 and 36 months were genetically only moderately correlated, with estimates ranging from 0.48 to 0.69[11,14]. Additionally, a considerable proportion (3% to 28%) of the total variation in early expressive language assessed at 24, 36 and 48 months could be explained by measure-specific additive genetic variance and not by a shared latent factor[16]. However, the field is still missing an in-depth characterisation of the genetic architecture underlying early-life vocabulary development that characterises age-specific genetic influences across infancy and toddlerhood starting from the first-word stage as well as differences between receptive and expressive language skills.

Genetic links between early language processes (assessed from 24 to 48 months of age) and subsequent language- and literacy-related abilities (assessed from mid-childhood to adolescence) have been reported by studies of both twins and unrelated individuals[15–17]. This research suggested that genetic variance in mid-childhood/adolescent language, literacy and cognitive development can already be captured by genetic factors contributing to language skills in toddlerhood, i.e. before the age of four years. More specifically, genetic influences underlying receptive vocabulary at 38 months could capture, through amplification, the majority of genetic variation contributing to a wide spectrum of mid-childhood/early-adolescent literacy and (verbal) cognitive skills in a sample of unrelated individuals[15]. So far, however, our understanding of the developmental origin of these factors is incomplete.

Here, we (i) examine stability and change in the developmental genetic architecture of language during the first three years of life and (ii) assess origin and developmental stage of emerging genetic associations with verbal and non-verbal abilities during mid-childhood. We model multivariate genetic architectures underlying these traits as directly captured by genome-wide information (based on genetic-relationship-matrices, GRMs) for up to 6,524 unrelated youth from the UK Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort[18,19]. We apply GRM structural equation modelling (GSEM)[20], analogous to twin research-modelling techniques, and dissect the phenotypic variation into additive genetic and residual variance structures.

4.2. Results

Analysis strategy

A two-stage analysis strategy was followed: During the first stage of the analysis (Stage 1), we examine the multivariate genetic variance structure of expressive and receptive vocabulary from 15 to 38 months of age (Table 1). A structural equation model (SEM) only was fitted to vocabulary measures with at least nominal evidence for SNP-h^2 (P<0.05). During the second stage (Stage 2), we extend these models, and assess the emerging genetic links between early-life vocabulary (15 to 38 months) and reading, verbal intelligence quotient (VIQ) scores and performance (non-verbal) intelligence scores (PIQ) during mid-childhood (7 to 8 years of age, S1 Table). For all SEMs studied, we report path coefficients (the square root of individual factor variance contributions) and the corresponding percentage of explained phenotypic variance, in addition to total SNP-h^2, genetic and residual correlations, factorial co-heritability (the proportion of total SNP-h^2 explained by a specific genetic factor) and bivariate heritability (the contribution of genetic factors to the observed phenotypic correlation between two measures) (S3, S4, S5 Appendix).

Table 1. Early-life expressive and receptive vocabulary in ALSPAC

Measure	Psychological instrument	Mean Score (SE)	Mean Age (SE)	N (%male)	GCTA-SNP-h^2 (SE)
Expressive vocabulary	MacArthur CDI[a]	14.29(17.76)	1.28(0.08)	6,524(51.1)	0.11(0.05)
	MacArthur CDI[b]	64.21(35.11)	2.03(0.09)	6,014(51.7)	0.16(0.06)
	MacArthur CDI[b]	113.33(17.44)	3.21(0.10)	6,092(51.4)	0.18(0.06)
Receptive vocabulary	MacArthur CDI[a]	75.85(31.78)	1.28(0.08)	6,524(51.1)	0.08(0.05)
	MacArthur CDI[b]	109.75(23.75)	3.21(0.10)	6,092(51.4)	0.12(0.06)

Expressive vocabulary and receptive vocabulary were assessed between 15-38 months of age in independent children (genetic relationship<0.05). [a] Adapted form of the MacArthur CDI:Words & Gestures, consisting of 134 words. [b] Adapted from of the MacArthur CDI:Words & Sentences, consisting of 123 words. Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; CDI, Communicative Development Inventory; GCTA, Genome-wide Complex Trait Analysis; h^2, heritability; SNP, single-nucleotide polymorphism.

Stage 1: The developmental genetic architecture of early-life vocabulary skills

Univariate SNP-heritability estimates for early-life vocabulary measures: Measures of early-life language included expressive vocabulary at 15, 24 and 38 months and receptive vocabulary at 15 and 38 months (Table 1). They were assessed with parent-reported questionnaires and analysed as rank-transformed scores (see Methods). For comparison with multivariate models, we first estimated SNP-h^2 using Genome-based Restricted Maximum Likelihood as implemented in Genome-wide Complex Trait Analysis (GCTA) software[21]. Common genetic variation accounted for a modest proportion of phenotypic variation in early-life vocabulary throughout, except for receptive vocabulary at 15 months, where SNP-h^2 was consistent with zero (Table 1). GCTA-SNP-h^2 estimates for expressive vocabulary at 15, 24 and 38 months were 11%(SE=5%), 16%(SE=6%) and 18%(SE=6%), respectively. For receptive vocabulary at 15 and 38 months, SNP-h^2 was estimated at 8%(SE=5%) and 12%(SE=6%), respectively. Given little evidence for SNP-h^2 for receptive vocabulary at 15 months (P>0.05; Table 1), we excluded this measure from further correlation and GSEM analyses to facilitate the convergence of the models. Note that it was not possible to include the receptive vocabulary score at 24 months due to discrepancies in the questionnaire coding scheme (see Methods).

Bivariate phenotypic and genetic correlations among early-life vocabulary measures: Early-life vocabulary measures were phenotypically interrelated, although correlations decreased with increasing age windows (Fig 1a). The largest phenotypic correlation (r_p) was estimated between expressive and receptive vocabulary at 38

months (r_p=0.63). Bivariate genetic correlations (r_g) among early-life vocabulary measures emerged from 24 months of age onwards (Figs 1b). Mirroring phenotypic relationships, the largest genetic correlation was observed between expressive and receptive vocabulary assessed at 38 months (GCTA-r_g=0.86(SE=0.15), *P*=0.004).

Multivariate genetic variance structures between early-life vocabulary measures: Using GSEM, we studied the multivariate genetic architecture underlying early vocabulary development, while allowing for both shared (i.e. across age and/or ability) and unique (i.e. age- and ability-specific) genetic influences. A multivariate SEM was fitted to expressive vocabulary at 15, 24 and 38 months as well as receptive vocabulary at 38 months (in this order), following a Cholesky decomposition. SNP-h^2 estimates were nearly identical for all early-life vocabulary measures using univariate GCTA and multivariate GSEM approaches (Table S2). Estimated bivariate genetic correlations using GSEM were also highly consistent with GCTA findings (Figs 1b and 1c), with overlapping 95%-confidence intervals (95%-CIs). GSEM-estimated residual correlations among vocabulary measures were modest to moderate (Fig 1d), suggesting further shared aetiological mechanisms not captured by common variation.

Structural models of vocabulary measures assessed during the first three years of life revealed that the underlying genetic architecture is dynamic, with evidence for age-specific genetic influences (Fig 2). The first genetic factor (A1) accounted for 10.6%(SE=5.0%) of the phenotypic variation in expressive vocabulary at 15 months (Fig 2, S3 Table), which can be estimated by squaring the corresponding estimated path coefficient, here a_{11} (path coefficient a_{11}:0.33(SE=0.08), *P*=$2x10^{-5}$). By structural model design, the phenotypic variance explained by a_{11} corresponds to the SNP-h^2 of expressive vocabulary at 15 months (S2 and S3 Tables). Genetic factor A1 was also related to expressive vocabulary at 24 months (path coefficient a_{21}:0.21(SE=0.10), *P*=0.04), explaining 4.6%(SE=4.4%) of the phenotypic variation and accounting for almost a third of the SNP-h^2 (factorial co-heritability: 31.2%(SE=23.4%), S4 Table). However, there was little evidence for shared genetic influences between expressive vocabulary at 15 months and either expressive or receptive vocabulary scores at 38 months (Fig 2, S3 Table). This pattern of findings suggests that genetic influences underlying expressive vocabulary at 15 months play a decreasing role during the course of later vocabulary development, consistent with data from genetic correlation and bivariate heritability analyses (Figs 1b and 1c, S5 Table).

Expressive vocabulary at 24 months loaded on a second genetic factor (A2), explaining an additional 10.1%(SE=4.0%) of the phenotypic variation (path-coefficient a_{22}: 0.32(SE=0.06), *P*=$4x10^{-7}$; Fig 2, S3 Table) and the majority of the SNP-h^2 (factorial co-heritability: 68.8%(SE=23.4%), S4 Table). This genetic factor was also shared with both expressive (path coefficient a_{32}:0.27(SE=0.09), *P*=0.005) and receptive (path coefficient

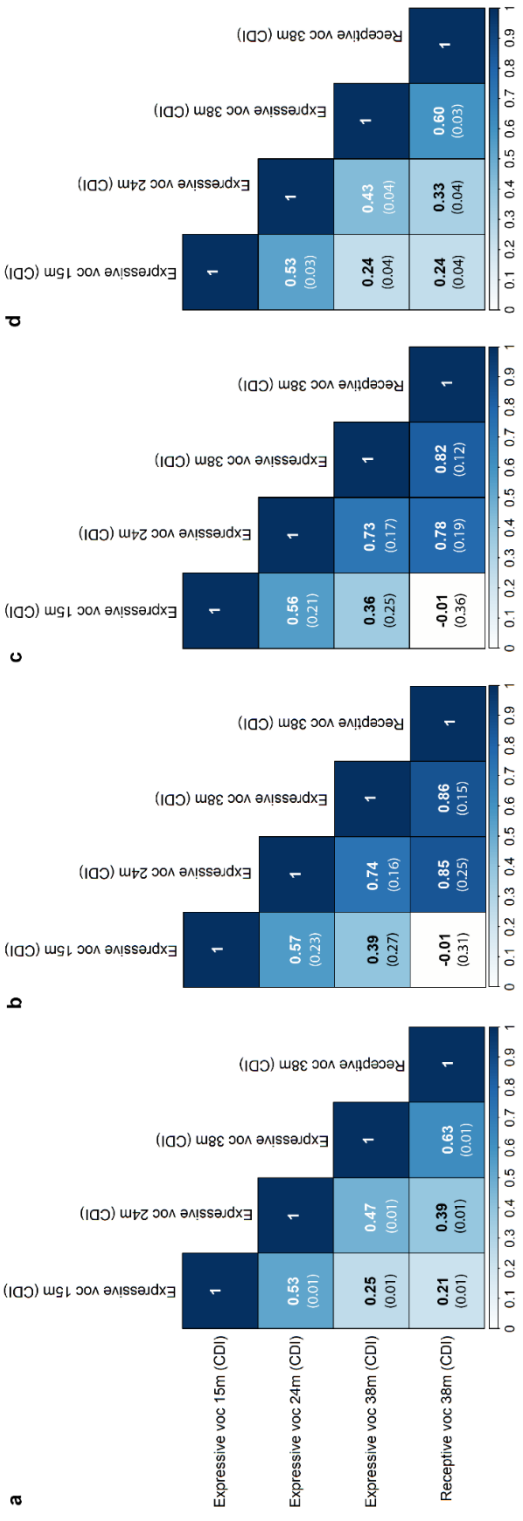


Fig 1. Phenotypic, genetic and residual correlations among early-life vocabulary scores (15 to 38 months). Correlation patterns are shown for rank-transformed measures with sufficient evidence for SNP-h^2 ($P<0.05$) Standard errors are shown in brackets. Genetic analyses were conducted using genetic relationship matrices based on directly genotyped SNPs and individuals with a genetic relationship of <0.05. **(a)** Phenotypic correlations were estimated with Pearson correlation coefficients. **(b)** GCTA genetic correlations based on GREML. **(c)** GSEM genetic correlations. **(d)** GSEM residual correlations. Abbreviations: CDI, Communicative Development Inventory; GCTA, Genome-based Restricted Maximum Likelihood as implemented in genome-wide complex trait analysis (GCTA) software; GREML, Genome-based restricted maximum likelihood; GSEM, genetic-relationship-matrix structural equation modelling; m, months; voc, vocabulary.

a_{42}:0.33(SE=0.08), P=4x10^{-5}) vocabulary at 38 months, accounting for 7.1%(SE=5.0%) and 11.0%(5.3%) of the phenotypic variation, respectively (Fig 2, S3 Table). For receptive vocabulary at 38 months, this genetic factor captured the majority of the SNP-h^2 (factorial co-heritability: 88.9%(SE=23.1%), Table S4), suggesting a largely shared genetic aetiology with expressive vocabulary at 24 months, as confirmed by their high genetic correlation (GSEM-r_g=0.78(SE=0.19), Fig 1c).

For receptive vocabulary at 38 months, this genetic factor captured the majority of the SNP-h^2 (factorial co-heritability: 88.9%(SE=23.1%), Table S4), suggesting a largely shared genetic aetiology with expressive vocabulary at 24 months, as confirmed by their high genetic correlation (GSEM-r_g=0.78(SE=0.19), Fig 1c).

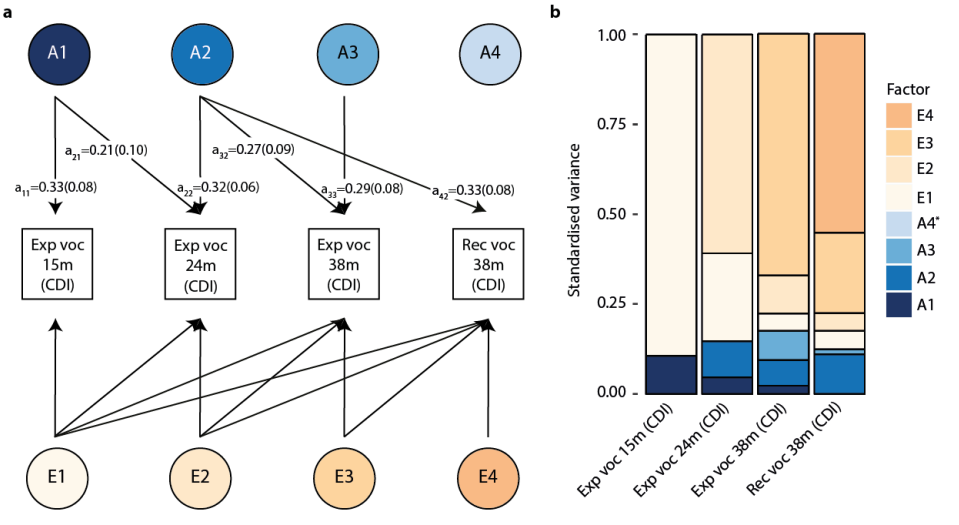


Fig 2. Structural model of early-life vocabulary scores (15 to 38 months). Genetic-relationship matrix structural equation modelling (GSEM) of early-life vocabulary scores (15, 24 and 38 months of age) based on all available observations for children across development (N≤6,524; Cholesky decomposition model). **(a)** Path diagram with standardised path coefficients and corresponding standard errors. Only paths with a path coefficient passing a P-value threshold of 0.05 are shown. Full information on path coefficients and their standard errors can be found in S3 Table. **(b)** Standardised variance explained by genetic and residual factors modelled in (a). * The proportion of phenotypic variance explained by genetic factor A4 in receptive vocabulary at 38 months is negligible. Abbreviations: CDI, Communicative Development Inventory; Exp, expressive; m, months of age; Rec, receptive; voc, vocabulary

The third genetic factor (A3) was only related to expressive vocabulary at 38 months (path coefficient a_{33}:0.29(SE=0.08), P=0.001) and explained 8.2%(SE=4.9%) of the phenotypic variation (Fig 2, S3 Table), corresponding to nearly half of the SNP-h^2 (factorial co-heritability: 47.0%(SE=25.1%), S4 Table). This genetic factor was unrelated to receptive vocabulary at 38 months (path coefficient a_{43}:0.12(SE=0.12), P=0.35). Thus, it is likely that the genetic correlation between expressive and receptive vocabulary at

38 months (GSEM-r_g=0.82(SE=0.12, Fig 1c) is primarily driven by genetic variance shared with expressive vocabulary at 24 months.

Finally, there was little support for the presence of a fourth genetic factor (A4) that would be exclusively related to receptive vocabulary at 38 months (Fig 2, S3 Table). However, according to findings from our previous work, such a factor is likely to account only for very little phenotypic variance in receptive vocabulary at 38 months[15]. Therefore, it may only become detectable once modelled together with other heritable traits sharing underlying genetic influences.

Stage 2: Multivariate genetic variance structures between early-life vocabulary and mid-childhood reading, verbal and performance intelligence

In a second step, we assessed the emergence of genetic links with mid-childhood reading accuracy/comprehension at 7 years, verbal intelligence quotient scores (VIQ) at 8 years and performance intelligence quotient scores (PIQ) at 8 years (S1 Table) across the studied vocabulary measures during the first three years of life, using rank-transformed measures. The selected measures of reading and verbal intelligence are representative of previously reported genetic association patterns between vocabulary at 38 months and a wide spectrum of language, literacy and cognitive abilities in ALSPAC[15]. We contrast these verbal abilities with a measure of non-verbal intelligence (PIQ) to evaluate differences in developmental association patterns with respect to early-life vocabulary. Thus, the model from the first step (Fig 2) was extended to include, in turn, each of the three mid-childhood skills, resulting in three further SEMs (with measures included in chronological order).

At the phenotypic level, all early-life vocabulary measures showed low to modest correlations with both mid-childhood verbal and non-verbal skills (Fig 3a), with the largest phenotypic correlation between receptive vocabulary at 38 months and VIQ at 8 years (r_p=0.26). The selected mid-childhood skills, reading, VIQ and PIQ, were all moderately heritable, with GCTA-SNP-h^2 estimates of 42%(SE=6%), 54%(SE=7%) and 26%(SE=7%), respectively. These estimates largely corresponded to GSEM-SNP-h^2 estimates (S2 Table). Using GCTA, bivariate genetic correlations of mid-childhood skills with early-life vocabulary measures (Fig 3b) revealed moderate genetic correlations of VIQ with expressive vocabulary at 24 (GCTA-r_g=0.41(SE=0.14),*P*=0.003) and 38 months (GCTA-r_g=0.38(SE=0.14),*P*=0.003), but high genetic correlations of both verbal and non-verbal skills with receptive vocabulary at 38 months (reading GCTA-r_g=0.83(SE=0.25), *P*=9x10^{-6}; VIQ GCTA-r_g=0.95(SE=0.23), *P*=1x10^{-8}; PIQ GCTA-r_g=0.68(SE=0.28), *P*=0.004). GCTA and GSEM genetic correlation estimates were highly consistent, with overlapping

95%-CIs (Figs 3b and 3c). GSEM-estimated residual correlations between early-life and mid-childhood measures were low (Fig 3d).

Using multivariate structural models, our results showed, first, that there is little evidence for genetic links between expressive vocabulary at 15 months (A1) and vocabulary, reading or cognition abilities after the age of 24 months (Fig 4, S6, S7, S8 Tables). Second, the developmentally novel genetic factor emerging for expressive vocabulary at 24 months (A2), explained further genetic variance in receptive and expressive vocabulary at 38 months (as outlined above) and, importantly, mid-childhood verbal skills. Specifically, it was related to both reading accuracy/comprehension (path coefficient a_{52}:0.25(SE=0.12), *P*=0.04) and VIQ (path coefficient a_{52}=0.42(SE=0.13), *P*=0.001), and accounted for 6.4%(6.2%) and 17.9%(11.1%) of their phenotypic variation, respectively (Fig 4, S6 and S7 Tables). However, this genetic factor was not linked to PIQ at 8 years (path coefficient a_{52}:-0.03(SE=0.12), *P*=0.78)(Fig 4e and 4f, S8 Table). These findings may reflect some genetic specificity for verbal skills (reading and VIQ), compared to non-verbal cognition, though the 95%-CIs for the identified path coefficients overlap (path coefficients a_{52}-reading accuracy/comprehension: 95%-CI=0.01-0.49, a_{52}-VIQ: 95%-CI=0.17-0.68, a_{52}-PIQ: 95%-CI=-0.26-0.20, derived assuming normality). Third, genetic influences identified for expressive vocabulary at 38 months (A3) were unrelated to receptive vocabulary assessed at the same age (as outlined above) and later mid-childhood abilities (Fig 4, S6, S7, S8 Tables). Thus, the genetic correlation observed between expressive vocabulary at 38 months and mid-childhood VIQ (GSEM-r_g=0.35(SE=0.13), Fig 3c) is primarily driven by genetic variance shared with expressive vocabulary at 24 months.

Fourth, joint modelling of early-life vocabulary measures with mid-childhood abilities enabled the identification of a genetic factor that affects receptive vocabulary at 38 months (A4) and that is independent of early-life expressive vocabulary genetic factors (path coefficient a_{44}:0.15(SE=0.07), *P*=0.04, Fig 4c). Although this genetic factor accounted for only a tiny proportion of the phenotypic variation in receptive vocabulary at 38 months 2.1%(SE=1.9%)), it explained 33.0%(SE=8.2%), 36.1%(SE=11.5%) and 24.7%(SE=7.5%) of the phenotypic variation in reading accuracy/comprehension, VIQ and PIQ, respectively (path coefficients a_{54}-reading accuracy/comprehension:0.57(SE=0.07), $P<1x10^{-10}$; a_{54}-VIQ: 0.60(0.10), $P=3x10^{-10}$; a_{54}-PIQ: 0.50(0.08), $P<1x10^{-10}$). The genetic variance explained by genetic factor A4 corresponds to the majority of the estimated SNP-h^2 for mid-childhood abilities, as indicated by factorial co-heritabilities (reading: 82.3%(SE=16.1%), VIQ: 66.4%(SE=19.9%), PIQ: 91.8%(SE=15.1%), S9 Table). Finally, there was little evidence for novel genetic factors emerging during mid-childhood (A5, Fig 4), consistent with previous findings[15]. Thus, the fitted multivariate models for early-life vocabulary and mid-childhood skills were consistent with both the identified multivariate genetic

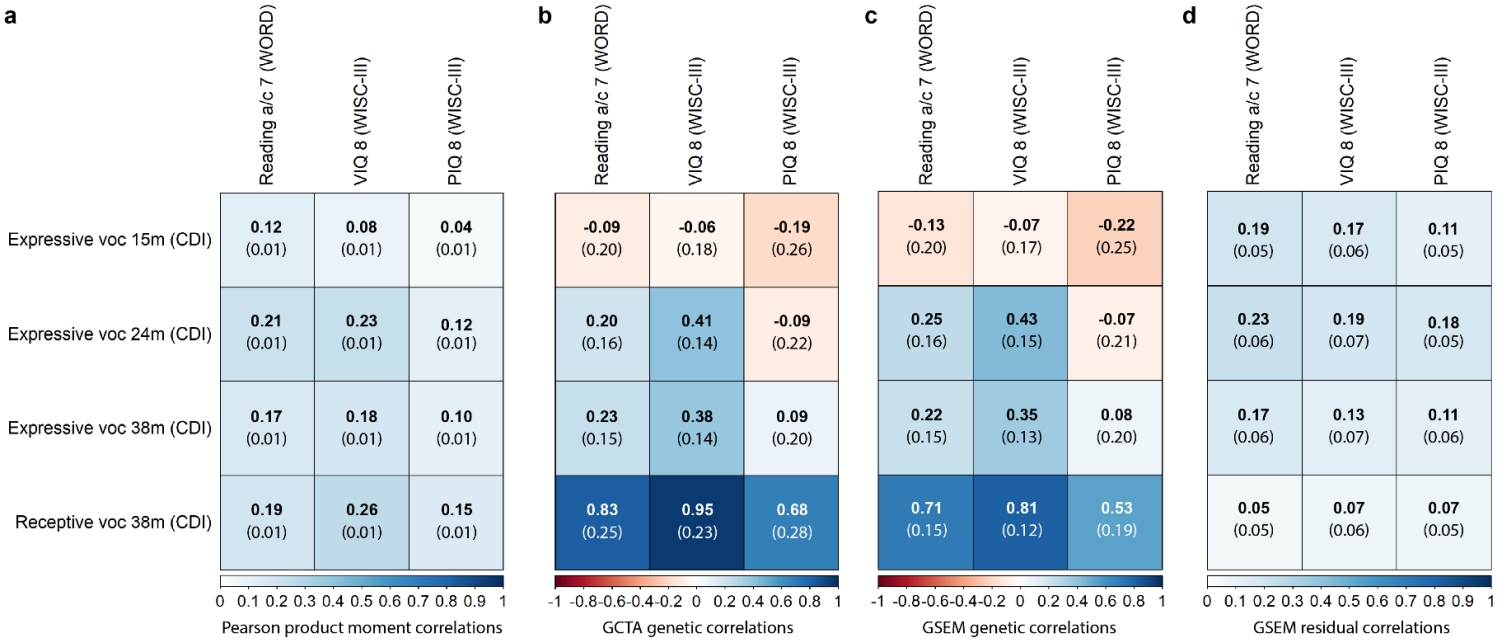
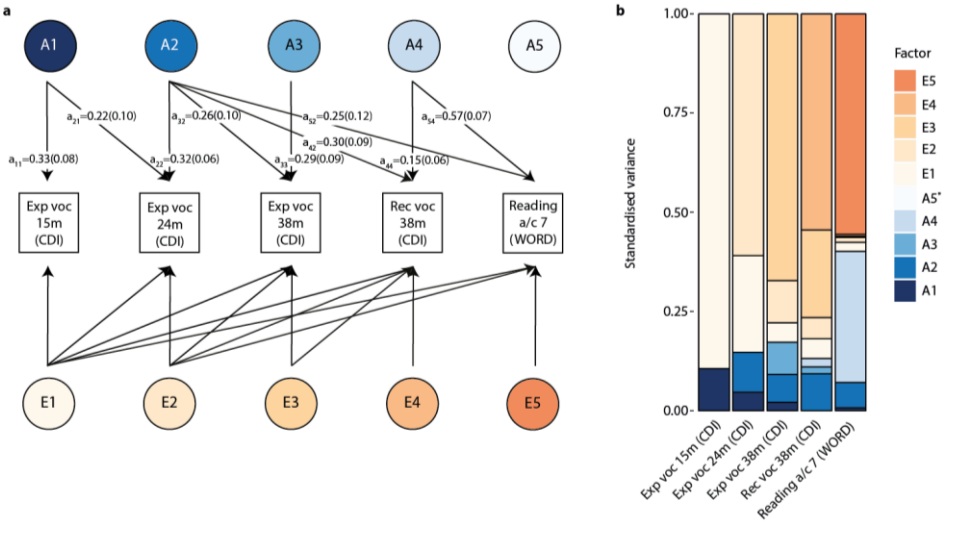


Fig 3. Phenotypic, genetic and residual correlations between early-life vocabulary scores and mid-childhood reading, verbal intelligence and performance intelligence. Correlation patterns are shown for rank-transformed measures with sufficient evidence for SNP-h^2 (P<0.05) Standard errors are shown in brackets. **(a)** Phenotypic correlations were estimated with Pearson correlation coefficients. **(b)** GCTA genetic correlations based on GREML. **(c)** GSEM genetic correlations. **(d)** GSEM residual correlations. Abbreviations: a, accuracy; c, comprehension; CDI, Communicative Development Inventory; GCTA, Genome-based Restricted Maximum Likelihood as implemented in genome-wide complex trait analysis (GCTA) software; GREML, Genome-based restricted maximum likelihood; GSEM, genetic-relationship-matrix structural equation modelling; m, months; PIQ; performance intelligence quotient; VIQ; verbal intelligence quotient, voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WORD, Wechsler Objective Reading Dimension. Genetic analyses were conducted using genetic relationship matrices based on directly genotyped SNPs and individuals with a genetic relationship of <0.05

architecture of early-life vocabulary (Fig 2) and the previously reported amplification of genetic factors for vocabulary at 38 months[15].

The phenotypic covariance of mid-childhood reading, VIQ and PIQ with receptive vocabulary at 38 months (Fig 3a) was primarily due to genetic covariance, with bivariate heritability estimates of 0.87(SE=0.21), 0.88(SE=0.16) and 0.68(SE=0.27), respectively (S10 Table). This is consistent with little evidence for residual correlation between receptive vocabulary at 38 months and mid-childhood measures (Fig 3d). For verbal mid-childhood skills, such as VIQ, evidence for bivariate heritability with early-life expressive vocabulary was already detectable at 24 months of age (bivariate heritability: 0.54(SE=0.19)), as well as at 38 months of age (bivariate heritability: 0.60(SE=0.24)).



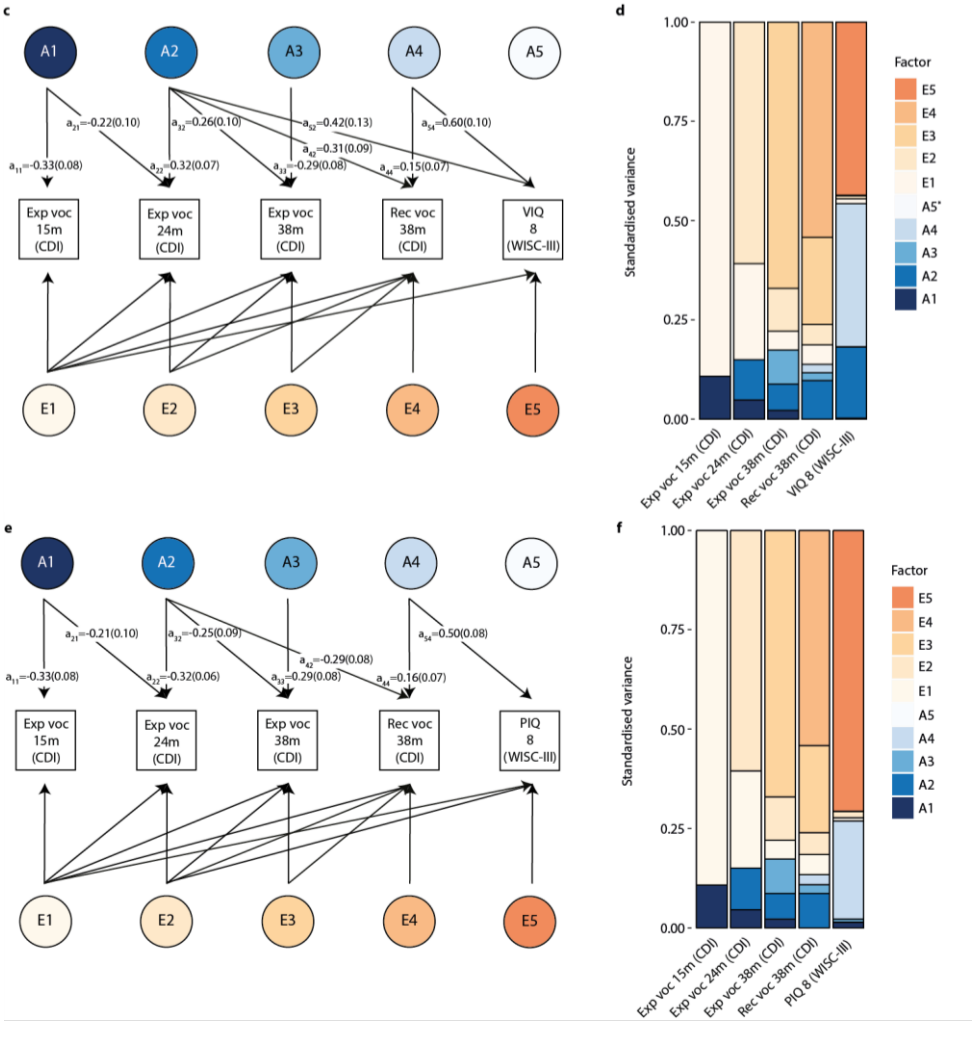


Fig 4. Structural models of early-life vocabulary and mid-childhood reading and cognition. Genetic-relationship matrix structural equation modelling (GSEM) of early-life vocabulary scores (15, 24 and 38 months of age) in combination with mid-childhood **(a,b)** reading accuracy/comprehension at 7 years, **(c,d)** VIQ scores at 8 years or **(e,f)** PIQ scores at 8 years, based on all available observations for children across development (N≤6,524). **(a,c,e)** Path diagrams with standardised path coefficients and corresponding standard errors including mid-childhood **(a)** reading accuracy/comprehension, **(c)** VIQ and **(e)** PIQ outcomes**.** Only paths with a path coefficient passing a *P*-value threshold of 0.05 are shown. Full information on all path coefficients and their standard errors can be found in S6, S7, S8 Tables. **(b,d,f)** Standardised variance explained by genetic and residual factors as modelled in a,c,e for models including **(b)** reading accuracy/comprehension, **(d)** VIQ, and **(f)** PIQ. * The proportion of phenotypic variance explained by genetic factor A5 is negligible. Abbreviations: a, accuracy; c, comprehension; CDI, Communicative Development Inventory; Exp, expressive; m, months of age; Rec, receptive; PIQ, performance intelligence quotient; VIQ, verbal intelligence quotient; voc, vocabulary; WISC-III, Wechsler Intelligence Scale for Children III; WORD, Wechsler Objective Reading Dimension

4.3. Discussion

This genome-wide longitudinal analysis of vocabulary size during the first three years of life assessed in unrelated children demonstrates that the genetic architecture underlying expressive and receptive vocabulary is dynamic, with evidence for both age- and ability-specific genetic influences. Genetic continuity was found for two independent early-life genetic factors, which contribute to the genetic variance of reading and cognitive skills in mid-childhood. One stable early-life genetic source of variation was related to expressive vocabulary and emerged at 24 months of age, accounting for between 6.4% and 17.9% of the phenotypic variation in mid-childhood abilities, especially verbal skills such as reading and VIQ. A second, independent and stable early-life genetic factor was identified for receptive vocabulary at 38 months and explained between 24.7% and 36.1% of the phenotypic variance in both mid-childhood verbal and non-verbal cognitive abilities, including PIQ, corresponding to the majority of SNP-h^2 (≥66%). Given the modest SNP-h^2 of early-life vocabulary scores, ranging from 11% to 18%, this suggests not only genetic stability, but also an amplification of early genetic variance during the life-course that contributes to the markedly increased SNP-h^2 of later-life reading and cognition (27% to 54%).

The identification of multiple independent genetic factors related to vocabulary during the first three years of life may reflect rapid changes in mastering behavioral and language skills. Genetic influences identified for expressive vocabulary at 15 months (A1) were also related to expressive vocabulary at 24 months, but were not linked to vocabulary, reading or cognition measures beyond this age. Thus, these early genetic influences might primarily affect the very first stages of language development that, once achieved, have little impact on subsequent verbal and cognitive development. A plausible candidate process for this is the acquisition of phonological skills to identify phonemes and sequences from speech and their storage for future production[23].

The stable independent genetic factor emerging for expressive vocabulary at 24 months (A2) contributes to the genetic architectures underlying verbal processes throughout childhood, in contrast to genetic factor A1. Specifically, the genetic influences captured by A2 were related to both expressive and receptive vocabulary at 38 months, as well as mid-childhood verbal abilities such as reading and VIQ, but not PIQ (although 95%-CIs of estimated genetic path coefficients overlap with those for PIQ). This genetic factor may reflect stages of language learning that take place after the production of words in isolation at the age of 10 to 15 months[3]. This includes, for example, an increasing vocabulary size as well as the use of more complex grammatical structures, marked by the emergence of two-word combinations around the age of 18 to 24 months[1,24]. It has been shown that lexical and grammatical development share underlying acquisition mechanisms[25] and measures of expressive vocabulary and

grammatical development at two and three years of age are both phenotypically and genetically correlated[11].

Expressive vocabulary at 38 months loaded on an additional independent genetic factor (A3) that was not related to receptive vocabulary at the same age, nor to any of the studied mid-childhood reading and IQ measures. This genetic factor may, thus, involve genetic associations with processes that affect expressive vocabulary at an early age, but do not play a role in later cognition. They may, for example, reflect social abilities, which are known to impact on vocabulary development and vice versa[26]. Note that expressive vocabulary at 38 months is nonetheless genetically related to mid-childhood verbal processes due to shared genetic influences that were already detectable at 24 months (A2).

The majority of SNP-h^2 for mid-childhood reading, VIQ and also PIQ was accounted for by a genetic factor that emerged at 38 months of age for receptive vocabulary (A4), consistent with previous findings[15]. Although this stable genetic factor explained only a very small part of the phenotypic variance in receptive vocabulary (2.1%), it accounted from 66% to 92% of the phenotypic variation in later reading performance, verbal and non-verbal cognition, with very little residual contributions. Due to the wide spectrum of associated mid-childhood phenotypes that are linked with this genetic factor, including both later verbal and non-verbal cognitive abilities, it is possible that the genetically encoded biological processes are important for cognitive development in general. It merits noting that the genetic factor A4 was only detectable once modelled together with a mid-childhood skill sharing underlying genetic variance, probably due to the low proportion of phenotypic variation that it explained in early-life receptive vocabulary.

Previous twin studies demonstrating genetic links between language use in early childhood and later language/literacy skills have been based on a latent factor approach jointly capturing genetic variance of expressive language skills between the ages of 2 and 4 years[16,17]. Here, we used a sample of unrelated children with genome-wide genotyping data and distinguish language measures during the first three years of life based on both modality and age at assessment. We extend and refine the previous twin findings by showing that (i) early-life expressive vocabulary at 15 months of age is influenced by a genetic factor that is only shared across expressive vocabulary scores during infancy, and (ii) that there are at least two independent genetic factors during early life that are associated with mid-childhood reading and cognition. Genetic associations with mid-childhood verbal cognitive processes arise as early as 24 months of age, whereas genetic influences that are relevant for mid-childhood general cognitive development emerge as early as 38 months of age for receptive vocabulary, and are independent of expressive vocabulary. This latter distinction is important as receptive vocabulary at 38 months also shares a genetic factor with expressive vocabulary at 24 months and subsequent reading and verbal intelligence. The diversity in genetic factors may implicate differences in

overarching cognitive processes that are already detectable during toddlerhood. This is important as genetic influences associated with early-life vocabulary could fully account for the SNP-h^2 of mid-childhood reading, verbal and non-verbal intelligence. The presence of such genetic stability implicating verbal processes and general cognition from toddlerhood to, at least, mid-childhood may, furthermore, suggests shared biological underpinnings. Thus, joint genome-wide association study analyses across developmental stages may facilitate an increase in study power.

In addition to the strengths of this study described above, this study benefits from modelling multivariate genetic variance structures in unrelated individuals directly, based on genome-wide information, using novel structural equation modelling techniques. It is, however, not possible to infer biological mechanisms underlying the identified genetic factor structures with the current methodology. We furthermore exploit the phenotypic richness of the ALSPAC cohort, including longitudinally assessed vocabulary measures during early development as well as reading and cognitive outcomes in mid-childhood. This study has also several limitations. Given the rapidly changing nature of early vocabulary size, increasingly larger and complex word lists are required to reliably assess vocabulary size at 24 and 38 months compared to 15 months of age. Thus, the observed differences in genetic factor structures during early life may reflect differences in CDI instruments, although this is unlikely to fully explain our findings, given substantial phenotypic correlations between expressive vocabulary scores at 15 and 24 months of age (r_p=0.53). Furthermore, vocabulary assessments at 38 months of age might be affected by ceiling effects, as the MacArthur CDI:Words & Sentences was developed for children up to 30 months[5]. This may have reduced phenotypic variation and, thus, power to detect genetic variance components at 38 months. In addition, it has recently been shown that heritability and genetic relationships estimated in samples of unrelated individuals, especially for cognition-related traits[27,28], might be inflated by indirect genetic effects, reflecting a type of gene-environment correlation[29]. The observed association patterns between early-life vocabulary and mid-childhood reading and cognitive skills may therefore represent both shared genetic variance and indirect genetic effects. Future research using family-based data is warranted to assess the impact of indirect genetic effects on the reported association patterns. Finally, the sparcity of large data sets with longitudinal information on expressive and receptive vocabulary during infancy and toddlerhood, in addition to genome-wide data, currently prevents a direct replication of our findings in independent cohorts.

Taken together, our findings reveal a dynamic genetic landscape underlying vocabulary during the first three years of life. We found evidence for genetic continuity of two independent early-life genetic factors that contribute to both verbal and general cognitive abilities in mid-childhood and manifest at different developmental stages during early-life language development. Thus, the genetic foundations for both mid-

childhood reading and cognition lie in toddlerhood, but are diverse, and may implicate aetiological differences in overarching cognitive processes that are detectable long before the age of schooling.

4.4. Methods

Sample description and trait selection

Cohort information: Participants were born in 1991 or 1992 and included in ALSPAC, a UK population-based birth cohort (S1 Appendix)[18,19]. Ethical approval was provided by the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Informed consent for questionnaire and clinical data was obtained from participants following recommendations of the ALSPAC Ethics and Law Committee at the time. Consent for biological samples was collected in accordance with the Human Tissue Act (2004).

Genetic analyses: Genotyping and genotype calling was performed using the Illumina HumanHap550 quad chip and Illumina GenomeStudio software. Quality control of genetic data was applied using PLINK (v1.07)[30] at both the SNP and individual level following standard procedures. Individuals were excluded in case of gender mismatch between reported and genetic sex information, >3% missing SNP information, non-European ancestry, or interindividual relatedness (genomic relatedness>0.05). SNPs were excluded if they had a low call rate (<99%), were rare (<1%) and/or deviated from Hardy-Weinberg equilibrium ($P<5x10^{-7}$). After quality control, 7,924 children and 465,740 SNPs with high-quality genetic data remained.

Early-life vocabulary measures: Expressive and receptive vocabulary was assessed at 15, 24 and 38 months of age using parental-reports (predominantly mother) of age-specific defined word lists adapted from the MacArthur Communicative Development Inventory (CDI). At 15 months, expressive and receptive vocabulary were assessed with an abbreviated version of the MacArthur CDI:Words & Gestures (133 words, 8 to 16 months of age)[31]. Scores were recorded as the number of words a child could "say and understand" (expressive vocabulary), and "understand" plus "say and understand" (receptive vocabulary), respectively. At 24 and 38 months of age, an abbreviated vocabulary list from the MacArthur CDI:Words & Sentences (123 words, 16-30 months of age)[5] was used. At both ages, expressive vocabulary was ascertained as the total number of words a child could "say" plus "say and understand". Receptive vocabulary at 38 months was measured as the total number of words a child could "understand" plus "say and understand". The receptive vocabulary score at 24 months was excluded due to discrepancies in the applied coding scheme (reflecting the total number of words a child could "understand" only, excluding words a child could "say and understand"), and recoded scores have not yet been released by ALSPAC.

CDI expressive vocabulary scores have high reliability and validity, showing correlations with direct assessments of over 0.70[32,33]. Receptive vocabulary assessed using parental report correlated 0.55 with direct assessment[33]. In total, N≤6,524 children (Table 1) had vocabulary scores and genome-wide genetic data available for analyses.

Mid-childhood measures: For the selection of mid-childhood measures, we build on our previous work identifying genetic links between vocabulary at 38 months and thirteen mid-childhood/adolescent literacy and cognitive measures[15]. As it is not possible, due to computational constrains, to study longitudinal genetic architectures of early-life vocabulary measures in combination with a wide spectrum of mid-childhood language, literacy and cognitive abilities, we selected three mid-childhood measures that are representative of previously observed developmental association patterns[15] (N≤5,296; S1 Table). The studied mid-childhood measures included reading accuracy/comprehension at 7 years, assessed using the Wechsler Objective Reading Dimensions (WORD)[34], as well as both VIQ and PIQ assessed at 8 years using the Wechsler Intelligence Scale for Children (WISC-III)[35]. Detailed descriptions, including validity and reliability, of each measure are available in the Supporting Information (S2 Appendix).

Phenotype transformation: All early-life vocabulary and mid-childhood measures were adjusted for sex, age (except for VIQ and PIQ as they were derived using age-specific norms), and the first two principal components (adjusting for subtle differences in ancestry[36]), and subsequently rank-transformed. In addition, early-life vocabulary measures were adjusted for age squared, as vocabulary develops rapidly during early childhood[37]. Phenotypic correlations between early-life vocabulary measures were estimated using untransformed (Spearman rank-correlation) and rank-transformed (Pearson correlation) scores respectively, and patterns were largely unaffected by trait transformation (S1 Fig). Phenotypic correlations between early-life vocabulary and mid-childhood reading, VIQ and PIQ measures were estimated using rank-transformed (Pearson correlation) scores only.

Genome-wide Complex Trait Analysis

Total SNP-h^2 was estimated using Genome-based restricted maximum likelihood (GREML) analyses[38,39], as implemented in GCTA software[21], based on a GRM including directly genotyped SNPs only (GCTA-SNP-h^2). Measures with little evidence for GCTA-SNP-h^2 (P>0.05) were excluded from further analyses.

Bivariate GREML[39] was applied to estimate bivariate genetic correlations among early-life vocabulary measures and between early-life vocabulary and mid-childhood reading, VIQ and PIQ measures.

Multivariate genetic analyses

To study the genetic architecture of vocabulary in a developmental context, we used Genetic-relationship-matrix Structural Equation Models (GSEMs)[20]. This is a multivariate structural equation modelling technique, which combines multivariate analysis methodologies established in twin research[40,41] with estimates of genetic relationships between unrelated individuals, as captured by genome-wide genetic markers[20] (S3 Appendix). Specifically, GSEMs dissect the phenotypic covariance structure into one or more additive genetic factors (A), capturing genetic variance tagged by common genotyped SNPs, as well as one or more residual factors (E) that resemble the residual variance, containing both untagged genetic variation and unique environmental influences (including measurement error). Here, multivariate GSEMs were fitted to the data through a Cholesky decomposition model, with the phenotypic variance decomposed into as many latent genetic and residuals factors as there are observed variables, without any restrictions on the structure[42] (S3 Appendix). Structural models were based on all available observations across individuals and thus allow for missing data (saturated model; R:gsem library, version 0.1.5). Genetic relationships between individuals were assessed with GRMs, including directly genotyped SNPs only, as implemented in GCTA software[21].

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. This publication is the work of the authors and EV and BSTP will serve as guarantors for the contents of this paper. The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institue and LabCorp (Laboratory Corporation of America) using support from 23andMe. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). EV, BSTP and SEF are supported by the Max Planck Society. BSTP is also supported by the Simons Foundation (Award ID: 514787). CYS works in a unit that receives support from the University of Bristol and the UK Medical Research Council.

References

1. Fenson, L. *et al.* Variability in early communicative development. *Monogr Soc Res Child Dev* **59**, 1–173; discussion 174-85 (1994).
2. Kennison, S. M. *Introduction to language development*. (SAGA, 2014).
3. Clark, E. V. *First language acquisition*. (Cambridge University Press, 2016).
4. Goldfield, B. A. & Reznick, J. S. Early lexical acquisition: rate, content, and the vocabulary spurt. *J Child Lang* **17**, 171–183 (1990).
5. Fenson, L. *et al. User's Guide and Technical Manual for the MacArthur Communicative Development Inventories*. (Singular Publishing, 1993).
6. Burger, A. & Chong, I. Receptive Vocabulary. in *Encyclopedia of Child Behavior and Development* (eds. Goldstein, S. & Naglieri, J. A.) 1231–1231 (Springer US, 2011). doi:10.1007/978-0-387-79061-9_2359.
7. Owens Jr., R., E. *Language Development: An Introduction*. (Pearson, 2015).
8. Bergelson, E. & Swingley, D. At 6-9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3253–3258 (2012).
9. Fenson, L. & Marchman, V. A. *MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual*. (Paul H. Brookes Publishing Company, 2007).
10. Frank, M. C., Braginsky, M., Yurovsky, D. & Marchman, V. A. Wordbank: an open repository for developmental vocabulary data. *J Child Lang* **44**, 677–694 (2017).
11. Dionne, G., Dale, P. S., Boivin, M. & Plomin, R. Genetic Evidence for Bidirectional Effects of Early Lexical and Grammatical Development. *Child Development* **74**, 394–412 (2003).
12. Dale, P. S., Dionne, G., Eley, T. C. & Plomin, R. Lexical and grammatical development: a behavioural genetic perspective. *Journal of Child Language* **27**, 619–642 (2000).
13. Reznick, J. S., Corley, R. & Robinson, J. A longitudinal twin study of intelligence in the second year. *Monogr Soc Res Child Dev* **62**, i–vi, 1–154; discussion 155-60 (1997).
14. St Pourcain, B. *et al.* Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature communications* **5**, 4831 (2014).
15. Verhoef, E., Shapland, C. Y., Fisher, S. E., Dale, P. S. & Pourcain, B. S. The developmental origins of genetic factors influencing language and literacy: Associations with early-childhood vocabulary. *Journal of Child Psychology and Psychiatry* (2020) doi:10.1111/jcpp.13327.
16. Harlaar, N., Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. Why Do Preschool Language Abilities Correlate With Later Reading? A Twin Study. *J Speech Lang Hear Res* **51**, 688–705 (2008).
17. Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev Sci* **15**, 233–49 (2012).
18. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* **42**, 97–110 (2013).
19. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111–127 (2013).
20. St Pourcain, B. *et al.* Developmental changes within the genetic architecture of social communication behaviour: A multivariate study of genetic variance in unrelated individuals. *Biological Psychiatry* **83**, 598–606 (2017).
21. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
22. Fenson, L., Dale, P. S. & Reznic, S. Technical Manual for the MacArthur Communicative Development Inventories. (1991).
23. Curtin, S. & Archer, S. L. Speech perception. in *The Cambridge Handbook of Child Language* (eds. Bavin, E. L. & Naigles, L. R.) 137–158 (Cambridge University Press, 2015). doi:10.1017/CBO9781316095829.007.
24. Hoff, E. *Language Development*. (Cengage Learning, 2013).
25. Bates, E. & Goodman, J. C. On the inseparability of grammar and the lexicon: Evidence from acquisition. in *Language development: The essential readings* 134–162 (Blackwell Publishing, 2001).
26. Silverman, R. D. & Hartranft, A. M. *Developing Vocabulary and Oral Language in Young Children*. (Guilford Publications, 2014).

27. Selzam, S. *et al.* Comparing Within- and Between-Family Polygenic Score Prediction. *The American Journal of Human Genetics* **105**, 351–363 (2019).
28. Cheesman, R. *et al.* Comparison of Adopted and Nonadopted Individuals Reveals Gene–Environment Interplay for Education in the UK Biobank: *Psychological Science* (2020) doi:10.1177/0956797620904450.
29. Kong, A. *et al.* The nature of nurture: Effects of parental genotypes. *Science* **359**, 424–428 (2018).
30. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* **81**, 559–575 (2007).
31. Fenson, L. *et al*. Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics* **21**, 95–116 (2000).
32. Dale, P. S. The validity of a parent report measure of vocabulary and syntax at 24 months. *J Speech Hear Res* **34**, 565–571 (1991).
33. Ring, E. D. & Fenson, L. The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language* **20**, 141–159 (2000).
34. *WORD, Wechsler Objective Reading Dimensions Manual*. (Psychological Corporation, 1993).
35. Wechsler, D., Golombok, S. & Rust, J. *WISC-III UK Wechsler Intelligence Scale for Children – Third Edition UK Manual*. (The Psychological Corporation, 1992).
36. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* **38**, 904–909 (2006).
37. Brooks, R. & Meltzoff, A. N. Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: a longitudinal, growth curve modeling study. *J Child Lang* **35**, 207–220 (2008).
38. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565–569 (2010).
39. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540–2542 (2012).
40. Martin, N. G. & Eaves, L. J. The genetical analysis of covariance structure. *Heredity (Edinb)* **38**, 79–95 (1977).
41. Neale, M. & Maes, H. H. M. *Methodology for genetic studies of twins and families*. (Kluwer Academic Publisher, 2004).
42. Neale, M., Boker, S., Xie, G. & Maes, H. H. M. *Mx: Statistical modeling*. (Department of Psychiatry, 2006).

Supplementary Materials

Supporting Methods

S1 Appendix. ALSPAC description

Pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992 were invited to take part in the study. The initial number of pregnancies enrolled is 14,541 (for these at least one questionnaire has been returned or a "Children in Focus" clinic had been attended by 19/07/99). Of these initial pregnancies, there was a total of 14,676 foetuses, resulting in 14,062 live births and 13,988 children who were alive at one year of age.

When the oldest children were approximately seven years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join the study originally. As a result, when considering variables collected from the age of seven onwards (and potentially abstracted from obstetric notes) there are data available for more than the 14,541 pregnancies mentioned above. The number of new pregnancies not in the initial sample (known as Phase I enrolment) that are currently represented on the built files and reflecting enrolment status at the age of 24 is 913 (456, 262 and 195 recruited during Phases II, III and IV respectively), resulting in an additional 913 children being enrolled. The phases of enrolment are described in more detail in the cohort profile paper and its update[1,2]. The total sample size for analyses using any data collected after the age of seven is therefore 14,545 pregnancies, resulting in 15,589 foetuses. Of these 14,901 were alive at one year of age.

A 10% sample of the ALSPAC cohort, known as the Children in Focus (CiF) group, attended clinics at the University of Bristol at various time intervals between 4 to 61 months of age. The CiF group were chosen at random from the last six months of ALSPAC births (1,432 families attended at least one clinic). Excluded were those mothers who had moved out of the area or were lost to follow-up, and those partaking in another study of infant development in Avon.

Details of all available data is fully searchable through a data dictionary on the study website (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/).

S2 Appendix. Mid-childhood ALSPAC measures

Reading accuracy and comprehension age 7 (WORD)

Decoding and word reading were assessed using the basic reading subtest of the Wechsler Objective Reading Dimensions (WORD)[3], including both pictures and words. This test has reliability and internal consistency coefficients ≥0.9. Its validity is also high, with an inter-correlation of 0.82 with the word-reading test from the Differential Ability Scale[4]. In short, the child was shown a series of four pictures, and each picture had four short words underneath it. For each picture, the child was asked to point to the word underneath that had the same beginning or ending sound as the picture. This was followed by a series of three pictures that all had four words beneath them starting with the same letter as the picture. This time the child was asked to point to the word that correctly named the picture. Finally, the child was presented with a series of 48 unconnected words, which increased in difficulty, and asked to read them aloud. The task was stopped if the child made six consecutive errors. A reading accuracy and comprehension score, that had a maximum score of 50, was computed as the sum of the number of items the child read/responded to correctly.

Verbal intelligence age 8 (WISC-III)

To assess verbal intelligence the child was assessed using a short form of the Wechsler Intelligence Scale for Children (WISC-III)[5]. The WISC-III comprises ten subtests and alternate items were administered for all subtests, but the coding subtest. The information, similarities, arithmetic, vocabulary, and comprehension subtests were used to create a score indicating verbal intelligence. The WISC-III verbal intelligence score has high test-retest correlations, ranging between 0.90 and 0.94, dependent on the age at assessment and the duration of the test-retest interval[6]. Correlations with the Kaufman Brief Intelligence Test and the Stanford-Binet IV suggest good construct validity, with estimates of 0.79 and 0.69 respectively[6]. After the calculation of raw scores based on the items used in the alternate item form of the WISC-III, total age-scaled scores for the verbal scale were calculated according to the look-up tables in the WISC-III manual. The maximum VIQ score is 160 and all scores were pro-rated.

Performance intelligence age 8 (WISC-III)

Performance intelligence was assessed using a short form of the Wechsler Intelligence Scale for Children (WISC-III)[5], including alternate items for all subtests, with the exception of the coding subtest. The five performance subtests of the WISC-III were used to create a score indicating performance intelligence: picture completion, coding, picture arrangement, block design and object assembly. The WISC-III performance intelligence quotient (PIQ) score has a correlation of 0.59 with the non-verbal score measured using the Otis-Lennon School Ability Test was 0.59[7]. It has high reliability, with

test-retest correlations of 0.89[8]. Raw PIQ scores were calculated based on the items used in the alternate item form of the WISC-III and had a maximum PIQ score of 160. Next, total age-scaled scores for the performance scale were calculated according to the look-up tables in the WISC-III manual. All scores were pro-rated.

S3 Appendix. Genetic-relatedness-matrix Structural equation modelling

A Cholesky decomposition[9] describes a multivariate trait P with phenotypic measurements t, resulting in a range of measures (P_1,P_2,...,P_t). The first measure (P_1) can be influenced by a latent genetic factor (A_1) that may also explain variance in the remaining measures (P_2,...,P_t). The second measure (P_2) can also be influenced by a second latent genetic factor (A_2) that is independent of A_1 and captures additional variation. As the first latent genetic factor, the second latent genetic factor may also capture variance in all other measures (P_3,...,P_t). The final measure (P_t) can thus be influenced by latent genetic factors (A_1,...,A_{t-1}), but also a latent genetic factor A_t. This latter genetic factor (A_t) is independent of previous latent genetic factors (A_1,...,A_{t-1}) and does not explain variance within any of the previous measures (P_1,...,P_{t-1})[10].

Based on the factor model, the expected phenotypic covariance matrix $\boldsymbol{\Sigma}$ for Z-standardised traits is:

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Gamma}\boldsymbol{\Theta}\boldsymbol{\Gamma}' \quad (1)$$

with a lower triangular matrix of genetic factor loadings $\boldsymbol{\Lambda}$, a diagonal matrix of latent genetic factor variances $\boldsymbol{\Phi}$ (standardised to unit variance), a lower triangular matrix of residual factor loadings $\boldsymbol{\Gamma}$ and a diagonal matrix of latent residual factor variances $\boldsymbol{\Theta}$. Both $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ were standardised to unit variance, such that they represent an identity matrix $\boldsymbol{I}$[11].

For example, for a Cholesky decomposition of three measures (P_1, P_2 and P_3), assuming three latent genetic factors (A_1, A_2 and A_3) and three residual factors (E_1, E_2 and E_3) this translates into the following expected phenotypic covariance matrix:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{p1}^2 & \sigma_{p12} & \sigma_{p13} \\ \sigma_{p12} & \sigma_{p2}^2 & \sigma_{p23} \\ \sigma_{p13} & \sigma_{p23} & \sigma_{p3}^2 \end{bmatrix} \quad (2)$$

with phenotypic variances $\sigma_{p1}^2, \sigma_{p2}^2$ and σ_{p3}^2, phenotypic covariances σ_{p12}, σ_{p13} and σ_{p23}, and the relevant matrices

$$\boldsymbol{\Lambda} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \boldsymbol{\Phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{\Gamma} = \begin{bmatrix} e_{11} & 0 & 0 \\ e_{21} & e_{22} & 0 \\ e_{31} & e_{32} & e_{33} \end{bmatrix}, \boldsymbol{\Theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

with genetic factor loadings a and residual factor loadings e.

Factor loadings, also known as path coefficients, were annotated with a or e followed by two numbers to indicate the specific path. In this notation, the first number indicates the direction of effect (the measure to which the path goes) and the second number indicates the origin of the effect[10]. For example, a_{21} indicates the genetic factor loading for the path originating from A_1 and affecting P_2.

A Cholesky decomposition of three standardised measures (see above), can be visualised using a path diagram (S2 Fig). The expected phenotypic variances and covariances can be expressed as follows:

$$\sigma^2_{p1} = a^2_{11} + e^2_{11} = 1 \quad (4)$$

$$\sigma^2_{p2} = (a^2_{21} + a^2_{22}) + (e^2_{21} + e^2_{22}) = 1 \quad (5)$$

$$\sigma^2_{p3} = (a^2_{31} + a^2_{32} + a^2_{33}) + (e^2_{31} + e^2_{32} + e^2_{33}) = 1 \quad (6)$$

$$\sigma_{p12} = a_{11}a_{21} + e_{11}e_{21} \quad (7)$$

$$\sigma_{p13} = a_{11}a_{31} + e_{11}e_{31} \quad (8)$$

$$\sigma_{p23} = a_{31}a_{21} + a_{32}a_{22} + e_{31}e_{21} + e_{32}e_{22} \quad (9)$$

The variance of the latent genetic and residual factors has been standardised to unit variance and is not shown.

Bivariate genetic correlation estimates (ranging from -1 to 1) reflect the extent to which two measures share genetic factors and can be derived using estimated genetic variances and covariances[12] according to:

$$r_g = \frac{\sigma_{g12}}{\sqrt{\sigma^2_{g1}\sigma^2_{g2}}} \quad (10)$$

with genetic covariance σ_{g12} between measures P_1 and P_2, and the genetic variances σ^2_{g1} and σ^2_{g2}.

S4 Appendix. Factorial co-heritability

Factorial co-heritability was estimated to quantify the relative contribution of a genetic factor to the total genetic variance of a phenotype, using the gsem package (R:gsem library, version 0.1.5). We derived the factorial co-heritability (f_g^2) according to:

$$f^2_g = \frac{\sigma^2_{g\_it}}{\sum \sigma^2_{g\_it}} = \frac{\sigma^2_{g\_it}}{\sigma^2_{g\_t}} \quad (11)$$

with the genetic variance of genetic factor i contributing to trait t $\sigma^2_{g\_it}$ and the total genetic variance of trait t $\sigma^2_{g\_t}$. Estimates were derived using standardised path

coefficients, corresponding standard errors (SEs) using the Delta method, and *P*-values approximated with a Wald test.

S5 Appendix. Bivariate heritability

The proportion of phenotypic covariance between two traits that is accounted for by the genetic covariance was expressed as bivariate heritability[13], as incorporated in the gsem package (R:gsem library, version 0.1.5). It was estimated based on unstandardised path coefficients and the phenotypic covariance estimated for rank-transformed measures. Corresponding SEs were approximated by dividing the SE of the genetic covariance by the phenotypic covariance, based on the assumption that the SE of the phenotypic covariance is small. *P*-values were calculated using a Wald-test assuming normality.

S6 Appendix. Websites

GCTA: https://cnsgenomics.com/software/gcta/

GSEM: https://gitlab.gwdg.de/beate.stpourcain/gsem

Supporting Tables

S1 Table. Mid-childhood measures in ALSPAC

Measure	Psychological instrument	Mean Score (SE)	Mean Age (SE)	N (%males)
Reading accuracy and comprehension	WORD	28.52 (9.25)	7.53 (0.31)	5,723 (50.9)
Verbal intelligence[a]	WISC-III	108.04 (16.74)	8.64 (0.31)	5,305 (49.9)
Performance intelligence[a]	WISC-III	100.24(16.95)	8.64(0.31)	5,296 (49.9)

To compare findings between verbal and non-verbal cognitive processes, mid-childhood performance intelligence was also studied. Skills were assessed using standardised instruments and studied in independent ALSPAC participants only (genetic relatedness<0.05). a. Scores were derived using age norms and adjusted for sex and principal components only before transformation (see Methods). Representative of previous findings[14] reading skills and verbal intelligence during mid-childhood (7-8 year) were selected from ALSPAC. Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; WISC-III, Wechsler Intelligence Scale for Children III; WORD, Wechsler Objective Reading Dimension

S2 Table. SNP heritability estimates

Measure	N	GCTA-h^2(SE)	GSEM-h^2(SE)
Expressive vocabulary 15m (CDI)	6,524	0.11(0.05)	0.11(0.05)[b]
Receptive vocabulary 15m (CDI)[a]	6,524	0.08(0.05)	NA
Expressive vocabulary 24m (CDI)	6,014	0.16(0.06)	0.15(0.06)[b]
Expressive vocabulary 38m (CDI)	6,092	0.18(0.06)	0.18(0.06)[b]
Receptive vocabulary 38m (CDI)	6,092	0.12(0.06)	0.12(0.05)[b]
Reading a/c 7 (WORD)	5,723	0.42(0.06)	0.41(0.06)
VIQ 8 (WISC-III)	5,305	0.54(0.07)	0.54(0.06)
PIQ 8 (WISC-III)	5,296	0.26(0.07)	0.27(0.06)

SNP-heritability estimates were estimated based on rank-transformed scores, directly genotyped SNPs and individuals with a genetic relationship of <0.05 using Restricted Maximum Likelihood (REML) analyses as implemented in genome-wide complex trait analysis (GCTA) software. SNP-heritability estimates based on Genetic-relationship-matrix Structural Equation modelling (GSEM) were extracted for comparison. a. Due to limited evidence (P>0.05) for SNP-h^2 of receptive vocabulary assessed at 15 months using GCTA, this trait was excluded from further analyses. b. GSEM-h^2 estimates as observed in the GSEM model with reading accuracy/comprehension at 7 years. Abbreviations: a, accuracy; c, comprehension; CDI, Communicative Development Inventory; GCTA, genome-wide complex trait analysis; GSEM, Genetic-relationship-matrix Structural Equation modelling; h^2, heritability; PIQ, verbal intelligence quotient; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; WORD, Wechsler Objective Reading Dimension

S3 Table. Standardised path coefficients and variance explained for early-life vocabulary measures

Path	Standardised path coefficient		Standardised variance explained (%)
	Estimate (SE)	*P*	Estimate (SE)
a_{11}	0.33(0.08)	$2x10^{-5}$	10.6(5.0)
a_{21}	0.21(0.10)	0.04	4.6(4.4)
a_{31}	0.15(0.11)	0.18	2.2(3.3)
a_{41}	$-3x10^{-3}$(0.11)	0.98	$7x10^{-4}$(0.1)
a_{22}	0.32(0.06)	$4x10^{-7}$	10.1(4.0)
a_{32}	0.27(0.09)	0.005	7.1(5.0)
a_{42}	0.33(0.08)	$4x10^{-5}$	11.0(5.3)
a_{33}	0.29(0.08)	0.001	8.2(4.9)
a_{43}	0.12(0.12)	0.35	1.4(2.9)
a_{44}	$1x10^{-4}$(0.24)	1.00	$1x10^{-6}$(0.005)
e_{11}	0.95(0.03)	$<1x10^{-10}$	89.4(5.0)
e_{21}	0.49(0.04)	$<1x10^{-10}$	24.4(3.7)
e_{31}	0.22(0.04)	$6x10^{-8}$	4.8(1.8)
e_{41}	0.23(0.04)	$3x10^{-9}$	5.2(1.8)
e_{22}	-0.78(0.03)	$<1x10^{-10}$	60.9(4.0)
e_{32}	-0.33(0.04)	$<1x10^{-10}$	10.6(2.7)
e_{34}	-0.22(0.04)	$3x10^{-8}$	4.9(1.8)
e_{33}	0.82(0.03)	$<1x10^{-10}$	67.0(4.4)
e_{43}	0.47(0.04)	$<1x10^{-10}$	22.3(3.3)
e_{44}	0.74(0.02)	$<1x10^{-10}$	55.3(2.9)

Genetic-relationship matrix structural equation modelling (GSEM) of early-life vocabulary scores (15, 24 and 38 months of age) based on all available observations for children across development (N≤6,524; Cholesky decomposition model). A visual representation is provided in Fig 2.

S4 Table. Factorial co-heritability for early-life vocabulary measures

Path	Factorial co-heritability (%)	
	Estimate (SE)	*P*
a_{11}	100.0 (0.0)	$<1x10^{-10}$
a_{21}	31.2 (23.4)	0.18
a_{31}	12.7 (18.1)	0.48
a_{41}	0.005 (0.5)	0.99
a_{22}	68.8 (23.4)	0.003
a_{32}	40.3 (25.6)	0.12
a_{42}	88.9 (23.1)	$1x10^{-4}$
a_{33}	47.0 (25.1)	0.06
a_{43}	11.1 (23.0)	0.63
a_{44}	$8x10^{-6}$(0.04)	1.00

Factorial co-heritability reflects the proportion of total SNP-h^2 estimated for a trait explained by a specific genetic factor. SEs were derived using the Delta method and *P*-values based on a Wald-test assuming normality (Supporting Methods). For example, the factorial co-heritability of a_{42} was estimated as $a_{42}*a_{42}$ / ($a_{41}*a_{41}$ + $a_{42}*a_{42}$ + $a_{43}*a_{43}$ + $a_{44}*a_{44}$) and implies that genetic factor A2 explains 88.9%(SE=23.1%) of the total SNP-h^2 estimated for receptive vocabulary at 38 months (see Fig 2).

S5 Table. Bivariate heritability for early-life vocabulary measures

	Expressive voc 15m (CDI)	Expressive voc 24m (CDI)	Expressive voc 38m (CDI)	Receptive voc 38m (CDI)
Expressive voc 15m (CDI)				
Expressive voc 24m (CDI)	0.13 (0.08)			
Expressive voc 38m (CDI)	0.20 (0.16)	0.25* (0.09)		
Receptive voc 38m (CDI)	-0.004 (0.17)	0.28* (0.11)	0.19* (0.07)	

Bivariate heritability reflects the proportion of the phenotypic covariance between two traits that is accounted for by the genetic covariance. Standard errors (SEs) are shown in brackets and were approximated by the SE of the genetic covariance divided by the phenotypic covariance (as the SE of the phenotypic covariance is small). *P*-values are based on a Wald-test, assuming normality. * Bivariate heritability estimates passing a significance threshold of $P<0.05$. Abbreviations: CDI, communicative development inventory; m, months; voc, vocabulary

S6 Table. Standardised path coefficients and variance explained for early-life vocabulary and mid-childhood reading accuracy/comprehension

Path	Standardised path coefficient		Standardised variance explained (%)
	Estimate (SE)	*P*	Estimate (SE)
a_{11}	0.33(0.08)	$3x10^{-5}$	10.6(5.1)
a_{21}	0.22(0.10)	0.04	4.7(4.5)
a_{31}	0.15(0.11)	0.19	2.1(3.3)
a_{41}	0.01(0.11)	0.93	0.01(0.2)
a_{51}	-0.08(0.12)	0.50	0.7(2.1)
a_{22}	0.32(0.06)	$1x10^{-6}$	10.0(4.1)
a_{32}	0.26(0.10)	0.01	7.0(5.1)
a_{42}	0.30(0.09)	$3x10^{-4}$	9.3(5.2)
a_{52}	0.25(0.12)	0.04	6.4(6.2)
a_{33}	0.29(0.09)	0.001	8.1(4.9)
a_{43}	0.13(0.11)	0.24	1.7(3.0)
a_{53}	0.02(0.16)	0.93	0.02(0.5)
a_{44}	0.15(0.06)	0.02	2.1(1.9)
a_{54}	0.57(0.07)	$<1x10^{-10}$	33.0(8.2)
a_{55}	$-3x10^{-4}$(0.58)	1.00	$9x10^{-6}$(0.03)
e_{11}	0.95(0.03)	$<1x10^{-10}$	89.4(5.1)
e_{21}	0.49(0.04)	$<1x10^{-10}$	24.4(3.8)
e_{31}	0.22(0.04)	$5x10^{-8}$	4.8(1.8)
e_{41}	0.22(0.04)	$2x10^{-9}$	5.0(1.7)
e_{51}	0.15(0.04)	$2x10^{-4}$	2.3(1.2)
e_{22}	-0.28(0.03)	$<1x10^{-10}$	61.0(4.1)
e_{32}	-0.33(0.04)	$<1x10^{-10}$	10.6(2.7)
e_{42}	-0.23(0.04)	$4x10^{-9}$	5.3(1.8)
e_{52}	-0.11(0.04)	0.01	1.2(1.0)
e_{33}	-0.82(0.03)	$<1x10^{-10}$	67.2(4.4)
e_{43}	-0.47(0.03	$<1x10^{-10}$	22.1(3.1)
e_{53}	-0.06(0.04)	0.18	0.3(0.5)
e_{44}	0.74(0.02)	$<1x10^{-10}$	54.4(2.3)
e_{54}	-0.07(0.04)	0.08	0.5(0.5)
e_{55}	0.75(0.04)	$<1x10^{-10}$	55.5(5.6)

Genetic-relationship matrix structural equation modelling (GSEM) of early-life vocabulary scores (15, 24 and 38 months of age) in combination with mid-childhood reading accuracy/comprehension at 7 years, based on all available observations for children across development (N≤6,524). A visual representation is provided in Figs 4a and 4b.

S7 Table. Standardised path coefficients and variance explained for early-life vocabulary and mid-childhood verbal intelligence

Path	Standardised path coefficient		Standardised variance explained (%)
	Estimate (SE)	*P*	Estimate (SE)
a_{11}	-0.33(0.08)	$2x10^{-5}$	10.8(5.0)
a_{21}	-0.22(0.10)	0.03	4.8(4.6)
a_{31}	-0.15(0.11)	0.18	2.2(3.3)
a_{41}	-0.01(0.10)	0.39	0.02(0.3)
a_{51}	0.05(0.13)	0.70	0.2(1.2)
a_{22}	0.32(0.07)	$1x10^{-6}$	10.1(4.1)
a_{32}	0.26(0.10)	0.008	6.6(5.0)
a_{42}	0.31(0.09)	$3x10^{-4}$	9.7(5.3)
a_{52}	0.42(0.13)	0.001	17.9(11.1)
a_{33}	-0.29(0.08)	$4x10^{-4}$	8.5(4.8)
a_{43}	-0.14(0.11)	0.22	1.9(3.1)
a_{53}	-0.02(0.19)	0.91	0.1(0.8)
a_{44}	0.15(0.07)	0.04	2.2(2.1)
a_{54}	0.60(0.10)	$3x10^{-10}$	36.1(11.5)
a_{55}	$4x10^{-4}$(0.56)	1.00	$2x10^{-5}$(0.1)
e_{11}	-0.94(0.03)	$<1x10^{-10}$	89.2(5.0)
e_{21}	-0.49(0.04)	$<1x10^{-10}$	24.2(3.8)
e_{31}	-0.22(0.04)	$6x10^{-8}$	4.8(1.8)
e_{41}	-0.22(0.04)	$2x10^{-9}$	4.9(1.6)
e_{51}	-0.11(0.04)	0.01	1.3(1.0)
e_{22}	0.78(0.03)	$<1x10^{-10}$	60.8(4.1)
e_{32}	0.33(0.04)	$<1x10^{-10}$	10.8(2.7)
e_{42}	0.23(0.04)	$6x10^{-9}$	5.1(1.8)
e_{52}	0.08(0.05)	0.08	0.7(0.8)
e_{33}	0.82(0.03)	$<1x10^{-10}$	67.1(4.4)
e_{43}	0.47(0.03)	$<1x10^{-10}$	21.9(3.1)
e_{53}	0.03(0.05)	0.45	0.1(0.3)
e_{44}	-0.74(0.02)	$<1x10^{-10}$	54.2(2.3)
e_{54}	0.02(0.04)	0.59	0.05(0.2)
e_{55}	-0.66(0.05)	$<1x10^{-10}$	43.6(6.0)

Genetic-relationship matrix structural equation modelling (GSEM) of early-life vocabulary scores (15, 24 and 38 months of age) in combination with mid-childhood verbal intelligence scores at 8 years, based on all available observations for children across development (N≤6,524). A visual representation is provided in Figs 4c and 4d.

S8 Table. Standardised path coefficients and variance explained for early-life vocabulary and mid-childhood performance intelligence

Path	Standardised path coefficient		Standardised variance explained (%)
	Estimate (SE)	*P*	Estimate (SE)
a_{11}	-0.33(0.08)	$2x10^{-5}$	10.8(5.1)
a_{21}	-0.21(0.10)	0.04	4.6(4.4)
a_{31}	-0.15(0.11)	0.18	2.2(3.3)
a_{41}	-0.004(0.11)	0.97	0.001(0.1)
a_{51}	0.12(0.13)	0.36	1.4(3.0)
a_{22}	-0.32(0.06)	$2x10^{-7}$	10.5(4.0)
a_{32}	-0.25(0.09)	0.01	6.5(4.7)
a_{42}	-0.29(0.08)	$4x10^{-4}$	8.6(4.8)
a_{52}	-0.03(0.12)	0.78	0.1(0.8)
a_{33}	0.29(0.08)	$3x10^{-4}$	8.6(4.7)
a_{43}	0.15(0.10)	0.14	2.2(3.1)
a_{53}	0.09(0.14)	0.55	0.7(2.4)
a_{44}	0.16(0.07)	0.02	2.5(2.1)
a_{54}	0.50(0.08)	$<1x10^{-10}$	24.7(7.5)
a_{55}	0.01(0.45)	0.99	0.003(0.4)
e_{11}	-0.94(0.03)	$<1x10^{-10}$	89.2(5.1)
e_{21}	-0.49(0.04)	$<1x10^{-10}$	24.4(3.8)
e_{31}	-0.22(0.04)	$7x10^{-8}$	4.8(1.8)
e_{41}	-0.23(0.04)	$2x10^{-9}$	5.1(1.7)
e_{51}	-0.09(0.04)	0.04	0.8(0.8)
e_{22}	0.78(0.03)	$<1x10^{-10}$	60.5(4.0)
e_{32}	0.33(0.04)	$<1x10^{-10}$	10.9(2.7)
e_{42}	0.23(0.04)	$3x10^{-9}$	5.5(1.8)
e_{52}	0.12(0.05)	0.01	1.5(1.2)
e_{33}	-0.82(0.03)	$<1x10^{-10}$	67.0(4.4)
e_{43}	-0.47(0.03)	$<1x10^{-10}$	21.9(3.1)
e_{53}	-0.03(0.05)	0.52	0.1(0.3)
e_{44}	0.74(0.02)	$<1x10^{-10}$	54.1(2.6)
e_{54}	-0.01(0.04)	0.82	0.01(0.1)
e_{55}	0.84(0.04)	$<1x10^{-10}$	70.6(6.4)

Genetic-relationship matrix structural equation modelling (GSEM) of early-life vocabulary scores (15, 24 and 38 months of age) in combination with mid-childhood performance intelligence scores at 8 years, based on all available observations for children across development (N≤6,524). A visual representation is provided in Figs 4e and 4f.

S9 Table. Factorial co-heritability for genetic factors contributing to mid-childhood reading, verbal intelligence and performance intelligence

Measure	Factorial co-heritability (%)									
	Expressive voc 15 months[a]		Expressive voc 24 months[b]		Expressive voc 38 months[c]		Receptive voc 38 months[d]		Mid-childhood LRA[e]	
	Estimate (SE)	*P*	Estimate (SE)	*P*	Estimate (SE)	*P*	Estimate (SE)	*P*	Estimate (SE)	*P*
Reading a/c 7 (WORD)	1.7(5.2)	0.74	15.9(15.1)	0.29	0.1(1.2)	0.96	82.3(16.1)	$3x10^{-7}$	$2x10^{-5}$(0.1)	1.00
VIQ 8 (WISC-III)	0.4(2.3)	0.84	33.0(20.0)	0.10	0.1(1.6)	0.95	66.4(19.9)	$8x10^{-4}$	$5x10^{-5}$(0.1)	1.00
PIQ 8 (WISC-III)	5.0(11.2)	0.65	0.4(2.9)	0.89	2.7(9.0)	0.76	91.8(15.1)	$1x10^{-9}$	0.01(1.7)	1.00

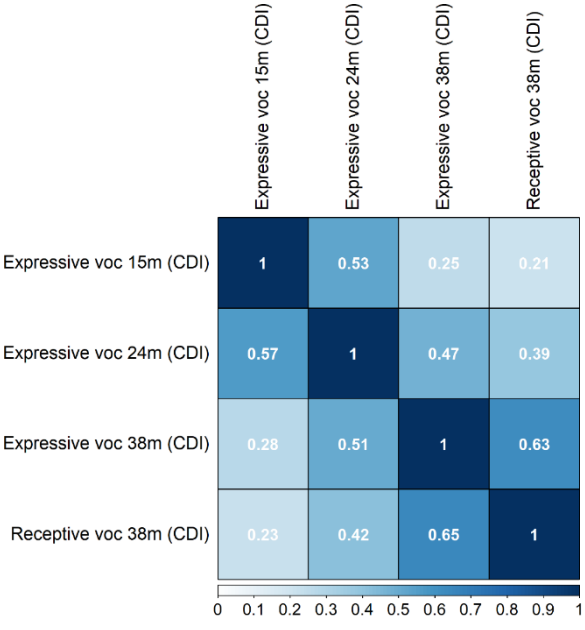
Factorial co-heritability reflects the proportion of the total SNP-h^2 estimated for a mid-childhood ability explained by a specific genetic factor. SEs were derived using the Delta method and *P*-values based on a Wald-test assuming normality (Supporting Methods). a. Proportion of genetic influences for expressive vocabulary at 15 months with respect to the total mid-childhood ability SNP-h^2: $a_{51}*a_{51}$ / ($a_{51}*a_{51} + a_{52}*a_{52} + a_{53}*a_{53} + a_{54}*a_{54} + a_{55}*a_{55}$). b. Proportion of genetic influences for expressive vocabulary at 24 months with respect to the total mid-childhood ability SNP-h^2: $a_{52}*a_{52}$ / ($a_{51}*a_{51} + a_{52}*a_{52} + a_{53}*a_{53} + a_{54}*a_{54} + a_{55}*a_{55}$). c. Proportion of genetic influences for expressive vocabulary at 38 months with respect to the total mid-childhood ability SNP-h^2: $a_{53}*a_{53}$ / ($a_{51}*a_{51} + a_{52}*a_{52} + a_{53}*a_{53} + a_{54}*a_{54} + a_{55}*a_{55}$). d. Proportion of genetic influences for receptive vocabulary at 38 months with respect to the total mid-childhood ability SNP-h^2: $a_{54}*a_{54}$ / ($a_{51}*a_{51} + a_{52}*a_{52} + a_{53}*a_{53} + a_{54}*a_{54} + a_{55}*a_{55}$). e. Proportion of genetic influences for mid-childhood LRA with respect to the total mid-childhood ability SNP-h^2: $a_{55}*a_{55}$ / ($a_{51}*a_{51} + a_{52}*a_{52} + a_{53}*a_{53} + a_{54}*a_{54} + a_{55}*a_{55}$). Abbreviations: a, accuracy; c, comprehension; LRA, language- and literacy-related ability; PIQ, performance intelligence quotient; VIQ, verbal intelligence quotient; voc, vocabulary; WORD, WISC-III, Wechsler Intelligence Scale for Children III

S10 Table. Bivariate heritability for early-life vocabulary measures and mid-childhood reading, verbal intelligence and performance intelligence

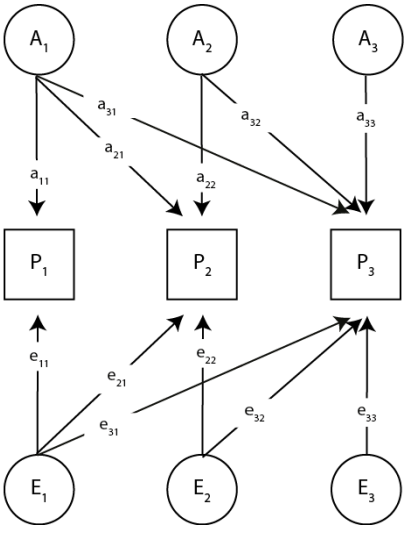
Measure	Bivariate heritability							
	Expressive voc 15 months		Expressive voc 24 months		Expressive voc 38 months		Receptive voc 38 months	
	Estimate (SE)	*P*	Estimate (SE)	*P*	Estimate (SE)	*P*	Estimate (SE)	*P*
Reading a/c 7 (WORD)	-0.24(0.34)	0.48	0.29(0.20)	0.13	0.36(0.25)	0.15	0.87(0.21)	$3x10^{-5}$
VIQ 8 (WISC-III)	-0.19(0.49)	0.69	0.54(0.19)	0.004	0.60(0.24)	0.01	0.88(0.16)	$8x10^{-8}$
PIQ 8 (WISC-III)	-0.88(0.94)	0.35	-0.12(0.36)	0.74	0.17(0.44)	0.70	0.68(0.27)	0.01

Bivariate heritability, reflecting the proportion of the phenotypic covariance that is accounted for by the genetic covariance. Standard errors (SEs) were approximated by the SE of the genetic covariance divided by the phenotypic covariance (as the SE of the phenotypic covariance is small) and *P*-values are based on a Wald-test, assuming normality (Supporting Methods). Abbreviations: a, accuracy; c, comprehension; PIQ, performance intelligence quotient; VIQ, verbal intelligence quotient; voc, vocabulary; WORD, WISC-III, Wechsler Intelligence Scale for Children III

Supporting Figures



S1 Fig: Phenotypic correlations among early-life vocabulary measures. Phenotypic correlations among untransformed (lower triangle) and rank-transformed (upper triangle) measures with sufficient evidence for SNP-h^2 (P>0.05) were estimated with Spearman's rank and Pearson correlation coefficients respectively. All phenotypic correlation coefficients passed the significance threshold of P<0.05. Abbreviations: CDI, Communicative Development Inventory; m, months; voc, vocabulary



S2 Fig: Path diagram for a trivariate trait. The variance/covariance structure of multivariate trait consisting of three standardised measures P1, P2 and P3 can be described using a Cholesky decomposition consisting of three genetic factors (A1, A2 and A3) and three residual factors (E1, E2 and E3), shown here with genetic and residual factor loadings (path coefficients). The observed phenotypic measures are represented by squares, while all latent genetic and residual factors are represented by a circle. Single headed arrows ('paths') denote causal relationships between variables and are shown for genetic factor loadings (a) and residual factor loadings (e). Note that the variance of latent variables is constrained to unit variance, this is omitted from the diagrams to improve clarity.

Supporting References

1. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111–127 (2013).
2. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* **42**, 97–110 (2013).
3. *WORD, Wechsler Objective Reading Dimensions Manual*. (Psychological Corporation, 1993).
4. Elliott, C. D., Salerno, J. D., Dumont, R. & Willis, J. O. The Differential Ability Scales—Second Edition. in *Contemporary intellectual assessment: Theories, tests, and issues, 4th ed* 360–382 (The Guilford Press, 2018).
5. Wechsler, D., Golombok, S. & Rust, J. *WISC-III UK Wechsler Intelligence Scale for Children – Third Edition UK Manual*. (The Psychological Corporation, 1992).
6. Sattler, J. M. *Assessment of children: Cognitive applications, 4th ed*. (Jerome M Sattler Publisher, 2001).
7. Guilmette, T. J., Kennedy, M. L. & Queally, P. T. A Comparison of the WISC-III and the Otis-Lennon School Ability Test with Students Referred for Learning Disabilities. *Journal of Psychoeducational Assessment* **19**, 239–244 (2001).
8. Canivez, G. L. & Watkins, M. W. Long-term stability of the Wechsler Intelligence Scale for Children—Third Edition. *Psychological Assessment* **10**, 285–291 (1998).
9. Cherny, S. S. Cholesky Decomposition. in *Encyclopedia of Statistics in Behavioral Science* (John Wiley & Sons, Ltd, 2005).
10. Neale, M. & Maes, H. H. M. *Methodology for genetic studies of twins and families.* (Kluwer Academic Publishers, 2004).
11. Martin, N. G. & Eaves, L. J. The genetical analysis of covariance structure. *Heredity (Edinb)* **38**, 79–95 (1977).
12. Falconer, P. D. S. & Mackay, P. T. F. C. *Introduction to Quantitative Genetics*. (Longman, 1995).
13. Janssens, M. J. J. Co-heritability: Its relation to correlated response, linkage, and pleiotropy in cases of polygenic inheritance. *Euphytica* **28**, 601–608 (1979).
14. Verhoef, E., Shapland, C. Y., Fisher, S. E., Dale, P. S. & Pourcain, B. S. The developmental origins of genetic factors influencing language and literacy: Associations with early-childhood vocabulary. *Journal of Child Psychology and Psychiatry* (2020) doi:10.1111/jcpp.13327

.

Chapter 5

Genome-wide association meta-analysis of expressive and receptive vocabulary from infancy to early childhood

Abstract

Expressive and receptive vocabulary are widely used measures to assess language development in young children and have a complex underlying genetic architecture. Here, I conduct a meta- genome-wide association study (meta-GWAS) of expressive and receptive vocabulary between the ages of 15 and 38 months, the largest such study to date based on 37,913 observations and 17,298 individuals, within the Early Genetics and Life Course Epidemiology (EAGLE) Consortium. Meta-analyses were performed for early-phase expressive vocabulary (15-18 months, N=8,799), late-phase expressive vocabulary (24-38 months, N=16,615), and late-phase receptive vocabulary (24-38 months, N=6,291), as well as a combination thereof. These analyses involved children of European descent across seven independent population-based or community-based cohorts. There was no evidence for single-nucleotide polymorphism (SNP)-vocabulary associations at the genome-wide significance threshold adjusted for the number of independent traits studied ($P \leq 1.79 \times 10^{-8}$). The strongest association was observed for early-phase expressive vocabulary and rs9854781 ($P = 4 \times 10^{-8}$), a SNP located in an intergenic region at chr3p12.3 near *ROBO2*. This variant is in high linkage disequilibrium (LD) with rs764282 (LD-r^2=0.78), a genome-wide significant signal in a previous GWAS investigating samples that largely overlap with those in the current study. Multi-trait analysis of genome-wide association (MTAG) across genetically correlated vocabulary traits increased the statistical power, equivalent to analysing 26,206 individuals, but did not identify evidence for single variant genetic associations ($P \leq 5 \times 10^{-8}$). Single-trait GWAS summary statistics captured low SNP-heritability (SNP-h^2) for both early-phase (SNP-h^2=0.12(SE=0.05)) and late-phase (SNP-h^2=0.09(SE=0.03)) expressive vocabulary, while SNP-h^2 based on summary statistics for late-phase receptive vocabulary was consistent with zero (SNP-h^2=0.07(SE=0.08)), possibly due to low power. Genetic correlation analyses provided evidence for positive genetic links of late-phase expressive vocabulary with mid-childhood to early-adulthood reading ability (r_g=0.58(SE=0.19), P=0.003), adult educational attainment (r_g=0.23(SE=0.06), $P = 5 \times 10^{-4}$), and intelligence across the lifespan (r_g=0.25(SE=0.08), $P = 3 \times 10^{-4}$). Thus, genetic factors contributing to later-life cognition can already be partially tagged by measures of expressive vocabulary collected at the age of two years.

5.1. Introduction

Language development in infants and toddlers is often studied with measures of expressive and receptive vocabulary[1,2]. These constructs relate to the ability of infants to produce and understand language, respectively, and can be relatively easily (albeit indirectly) assessed through parental reports. The first spoken words, representing one of the milestones in language development, typically emerge between the ages of 10 to 15 months[2]. Once children reach a vocabulary size of ~50 words at an age of 12 to 18 months, there is often a period of rapid vocabulary growth around 16 to 22 months of age[3], resulting in an expressive vocabulary size between 100 and 600 words at ~24 months[4]. Around this age, children typically start to use word combinations for the first time, marking another milestone in language development: the onset of grammar[5,6]. Receptive vocabulary precedes expressive vocabulary during early development with an emergence already at the age of six to nine months[7]. Consequently, the number of words understood is usually larger than the number of words produced, with median values of 169 and 40 words for receptive and expressive vocabulary size at the age of 16 months, respectively[5].

Children within the same population show large individual differences in early vocabulary development, which are known to be modestly heritable[8–11]. Twin heritability (twin-h^2) estimates based on different subsets of a large community-based sample from the UK, including up to 11,466 individuals, range between 10% and 25% for expressive vocabulary assessed at 24 and 36 months of age[8–10]. Similar estimates were observed in studies of large samples of unrelated individuals[8]. Meta-analyses of single-nucleotide polymorphism heritability (SNP-h^2) estimates across community- and population-based samples from the UK, the Netherlands and Australia reported SNP-h^2 estimates of 13% and 14% for expressive vocabulary assessed at 15-18 months (N=8,022) and 24-30 months (N=9,966)[8], respectively. Receptive vocabulary at 14 months of age was also shown to be modestly heritable, assessed in 378 twin pairs from the US, with 28% of the phenotypic variation being accounted for by genetic influences[12].

During the course of infancy to early childhood, there is evidence for both stability and change in the genetic factors underlying expressive vocabulary. Genetic correlations for measures of expressive vocabulary between 15 and 36 months of age, assessed in UK samples of both unrelated children and twins, ranged from 0.48 to 0.68[8,9], suggesting moderate genetic stability. In addition, twin research on a different subset of the same UK twin sample, applying latent factor structural equation models, reported that 28%, 3% and 20% of the variance in expressive language skills at 24, 36 and 48 months of age could be attributed to age-specific genetic influences, respectively[13]. Age-specific genetic influences may also exist at the SNP level, supported by a previous meta-genome-wide association study (meta-GWAS) on expressive vocabulary across English and Dutch speaking community- and population-based cohorts, analysing up to 10,819

individuals[8]. The reported genome-wide association signal for expressive vocabulary at rs7642482, a single-nucleotide polymorphism (SNP) located near *ROBO2*, was detectable in children aged 15-18 months, but attenuated in toddlers aged 24-30 months of age[8], with non-overlapping 95%-confidence intervals.

Individual differences in early language skills are predictive of later-life outcomes, such as reading proficiency[14–16], suggesting shared genetic factors. At the level of individual SNPs, rs7642482 was found to be associated with mid-childhood reading speed at the nominal level (P<0.05), while no such relationships were seen with mid-childhood phonological memory, verbal intelligence or reading comprehension[8]. At the polygenic level, there is evidence for a moderate genetic correlation (r_g=0.36) between a latent factor of early expressive language (2, 3 and 4 years) and a latent factor of mid-childhood reading (7, 9 and 10 years), based on studying a community-based cohort of UK twins[13]. Similarly, using GWAS data, moderate to strong genetic links were also reported between reading abilities (7-13 years) and receptive vocabulary (38 months), with genetic correlations ranging from 0.58 to 0.92 in a sample of up to 6,092 unrelated children from the UK (chapter 3)[17].

Genetic influences underlying early vocabulary may also overlap with genetic factors that are involved in childhood-onset neurodevelopmental disorders, such as Attention-Deficit/Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD). Children with ADHD often experience difficulties with mastering language and literacy skills[18–20] and poor language skills in ADHD children at the age of three years of age were found to be predictive of inattention and hyperactive symptoms two years later in life[21]. For children diagnosed with ASD the phenotypic spectrum is wider, including both children who have little or no spontanous spoken language by the time they reach school age[22] as well as 'high-functioning' individuals who often experience little problems in the language domain[23]. Although little is known about genetic links between vocabulary assessed during the first three years of life and genetic risk for ADHD or ASD, there is evidence for genetic overlap between ADHD and mid-childhood/early-adolescent language- and literacy-related abilities, primarily related to reading[24–27], as shown by twin and molecular studies in unrelated individuals (see also chapter 6). Finally, language development might be biologically coupled to anthroprometric growth. For example, birth weight and length are important predictors of developmental milestones during the first few years of life[28–30]. Low birth weight in particular has been associated with impairments in both expressive and receptive language skills from infancy to late-childhood[28,29]. In addition, measures of head circumference are used to monitor brain development[31] and are highly correlated with brain volume assessed using MRI in infancy and childhood[32,33].

The goal of this study is to gain further knowledge of the genetic factors that are associated with early vocabulary development by performing a meta-GWAS for expressive and receptive vocabulary assessed between 15 to 38 months of age. The

current study extends a previous GWAS effort[8], by (i) increasing the total number of children studied by ~50%, (ii) using a high-density genomic imputation reference panel from the Haplotype Reference Consortium[34] (HRC), thereby allowing for a more detailed study of lower frequency genetic variants within the allele frequency range of 0.5%-1%, and (iii) including receptive vocabulary, in addition to expressive vocabulary scores. Here, I aim to identify novel SNPs contributing to variation in expressive and receptive vocabulary, while allowing for age- and ability-specific effects through a stratified design. In addition, I apply a multivariate analysis approach across vocabulary measures to maximise the statistical power to detect SNP-vocabulary associations. Follow-up analyses based on the derived GWAS summary statistics include gene-based genome-wide association, gene-set and gene-property analyses. Finally, I study genetic links of early vocabulary with several cognition-related later-life outcomes, anthropometric traits reflecting growth and childhood-onset neurodevelopmental disorders.

5.2. Methods

Phenotype selection and study design

Cohorts with information on quantitative vocabulary scores during the first three years of life and genome-wide genotypes were invited to participate in the current meta-GWAS, which was conducted within the framework of the he Early Genetics and Life Course Epidemiology (EAGLE) consortium[35] (https://www.eagle-consortium.org/working-groups/behaviour-and-cognition/early-language/). Vocabulary scores were assessed between 15 and 38 months of age and analysed as part of two developmental stages to allow for age-specific genetic influences, an early phase (15-18 months) and a late phase (24-38 months). The early phase reflects a developmental period during which children produce their first words, usually in isolation[2], whereas during the late phase children start to use word combinations and more complex grammatical structures[5,6]. Scores for receptive vocabulary were included for the late phase (24-38 months) only, as parents tend to underestimate receptive vocabulary in children below the age of two years in comparison to a direct assessment of child receptive language using a preferential looking task[36]. Furthermore, there was little evidence for SNP-h^2 of receptive vocabulary assessed at 15 months based on analyses of individual-level genotype data from the Avon Longitudinal Study of Parents And Children (ALSPAC) cohort, although this does not prevent to existence of individual SNP signals (chapter 4). The reliability and validity for parental assessments of expressive vocabulary is generally high, showing correlations with direct assessments of over 0.70[37,38].

Up to seven population-based or community-based studies participated in the analyses, of which two had longitudinal vocabulary assessments (Table 1). Observations for early-phase expressive vocabulary and late-phase receptive vocabulary did not

include longitudinal assessments of the same individuals and were thus independent, whereas late-phase expressive vocabulary analysis included vocabulary assessments of the same children within ALSPAC at two different ages (24 months and 38 months, Table 1). The number of individuals and observations included in each meta-GWAS are shown in Figure 1. Ethical approval was obtained by the local research ethics committee for each participating study, and all parents and/or legal guardians provided written informed consent.

Vocabulary scores were ascertained by parental report using age-specific word lists that were adapted from the MacArthur Communicative Development Inventory (CDI)[10,39–43] or the Language Development Survey (LDS)[44] (Table 1). The CDIs were originally developed to assess language and communication development in young children[42], whereas the LDS aims to identify children with language delays[44]. During the early phase (15-18 months), expressive vocabulary was assessed using an abbreviated

Table 1: Overview of participating cohorts

Cohort	Trait	Psychological Instrument	Raw trait score (SD)	Age (SD) in months	N individuals (males)
ALSPAC	Early-phase EV	MacArthur CDI:Words & Gestures	14.34 17.84)	15.42(0.98)	6,741(3,445)
	Late-phase EV	MacArthur CDI:Words & Sentences	64.10(35.20)	24.39(1.02)	6,208(3,197)
			113.28(17.5)	38.48(1.19)	6,291(3,226)
	Late-phase RV	MacArthur CDI:Words & Sentences	109.66(23.78)	38.48(1.19)	6,291(3,226)
BIS	Late-phase EV	MCDI:UKSF	78.31(20.08)	29.62(1.92)	383(210)
COPSAC	Late-phase EV	MacArthur CDI:Words & Sentences	253.00(158.12)	24.18(0.28)	487(256)
GenR	Early-phase EV	N-CDI-2A	17.51(17.05)	18.36(0.96)	2,058(1,054)
	Late-phase EV	LDS	245.86(53.67)	31.32(2.04)	1,825(937)
LSAC	Late-phase EV	MCDI	56.95(23.60)	33.51(2.51)	1,134(558)
Raine	Late-phase EV	LDS	185.60(83.44)	25.52(1.74)	980(504)
TEDS	Late-phase EV	MCDI	48.66(24.79)	24.48(1.20)	5,515(2,665)

Expressive and receptive vocabulary were assessed between 15-38 months of age using parental questionnaires in seven independent cohorts, resulting in a total of 37,913 observations from 17,298 individuals. Per cohort, psychological instrument, mean raw trait score and age, with corresponding standard errors, as well as sample size are reported. Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; BIS, Barwon Infant Study; CDI, Communicative Development Inventory; COPSAC, Copenhagen Prospective Studies on Asthma in Childhood; EV, expressive vocabulary; GenR, Generation Rotterdam; LDS; Language Development Survey; LSAC, Longitudinal Study of Australian Children; MCDI, MacArthur Communicative Development Inventory; Raine, the Western Australian Pregnancy Cohort; RV, receptive vocabulary; TEDS, Twins Early Development Study; UKSF, UK short form.

form of the MacArthur CDI:Words & Gestures[39] in the ALSPAC cohort (N=6,741). Early-phase expressive vocabulary was defined by this instrument as the total number of words a child could "say and understand" (in contrast to "understand" only), and thus jointly represents expressive and receptive vocabulary. Within Generation Rotterdam (GenR, N=2,058), expressive vocabulary was assessed at ~18 months of age using a Dutch adaptation of the short-form version of the MacArthur CDI (N-CDI-2A)[40]. This form included the response "say" in addition to "say and understand", so early-phase expressive vocabulary was defined as the number of words that fell into either of these categories.

During the late phase (24-38 months), expressive vocabulary was assessed with an abbreviated version of the MacArthur CDI:Words & Sentences[42] in ALSPAC (24 months: N=6,208; 38 months: N=6,291), the corresponding Danish adaptation[43] in Copenhagen Prospective Studies on Asthma in Childhood (COPSAC, N=487), and using the LDS[44] in GenR (N=1,825) and the Western Australian Pregnancy Cohort (Raine, N=980). Adapted forms of the MacArthur CDI[10,45] (MCDI) were used to assess expressive vocabulary in the Barwon Infant Study (BIS, N=383), the Longitudinal Study of Australian Children (LSAC, N=1,134), and the Twins Early Development Study (TEDS, N=5,515). For CDI vocabulary assessments[10,42,43,45], expressive vocabulary was defined as the number of words a child "says" or "says and understands". For LDS vocabulary assessments[44], expressive vocabulary was defined as the total number of words spontaneously produced by a child from a given list of words. The LDS and CDI have high concurrent validity, with a correlation of 0.95 on total vocabulary scores[46]. Late-phase receptive vocabulary scores were only available in ALSPAC at 38 months (N=6,291, Table 1) and assessed using an abbreviated form of the MacArthur CDI:Words & Sentences[42]. The receptive vocabulary score was calculated as the number of words a child could understand, regardless of whether they also produced the word, encoded as "understand" plus "say and understand".

Vocabulary scores were primarily assessed in English (ALSPAC, BIS, LSAC, Raine and TEDS), but also included Danish (COPSAC, Danish adaptation of the MacArthur CDI:Words & Sentences[43]) and Dutch (GenR, N-CDI-2A[40]) CDI versions. Research has shown that children follow similar patterns of language acquisition across different languages[47] and that CDI vocabulary assessments are comparable across different cultures, including English, Dutch and Danish[48].

The current study extends a previous meta-GWAS that focused on expressive vocabulary during early development, based on 8,889 and 10,819 individuals for early- and late-phase expressive vocabulary respectively[8]. I extended the age range for the late phase from 30 to 38 months of age, based on strong genetic correlations between vocabulary measures assessed at 24 and 38 months in the ALSPAC sample (Table S1) and augmented the late phase to measures of receptive vocabulary. In addition, I applied MTAG to allow for repetitive assessments of the same children at different ages and

analysed imputed genotyping information based on the HRC r1.1. reference panel[34] for all participating cohorts (see "Genotyping and imputation"). Compared to the previous meta-GWAS effort[8], three additional population-based cohorts (BIS, COPSAC and LSAC) participated in the late-phase expressive vocabulary analyses (N_{total}=2,004; Table 1). Furthermore, observations of expressive vocabulary at 38 months within ALSPAC children were included (N=6,291) and the sample size for TEDS was increased by 3,788 children due to the inclusion of related individuals. Together, this lead to a boost in statistical power that corresponded to analysing 19,926 individuals for late-phase expressive vocabulary, about double the sample size available in the previous meta-GWAS effort[8].

Genotyping and imputation

Genotyping within each cohort was conducted using high-density SNP arrays (Table 2). Quality control parameters for autosomal markers included individual call rate, SNP call rate, minor allele frequency, and deviations from Hardy-Weinberg equilibrium and are reported for each cohort in Table 2. In total, between 440,476 and 608,517 high-quality autosomal genotyped markers were subsequently imputed against a HRC r1.1 reference panel[34] using either the Sanger imputation server (EAGLE2[49] v2.0.5 and PBWT[50] software, https://imputation.sanger.ac.uk/) or the Michigan imputation server[51] (Minimac 3 and Shapeit v2.r790, https://imputationserver.sph.umich.edu/) (Table 2). X chromosomal markers were not included in the current study, as genotype probabilities derived from the Sanger imputation server for males were mis-interpreted by SNPTEST GWAS analysis software. This resulted in incorrect association analyses for X chromosomal markers in males, due to artificial heterozygote instead of haploid genotype readings.

Independent number of vocabulary measures

The number of independent vocabulary measures assessed in this study was estimated using Matrix Spectral Decomposition (matSpD)[52,53]. As the polygenic signal captured by GWAS summary statistics for late-phase receptive vocabulary was too low (SNP-h^2 Z-score < 1.5) to yield reliable results with Linkage Disequilibrium Score (LDSC)[54], it was not possible to estimate the number of independent traits based on genetic correlations derived from single-trait meta-analyses summary statistics (stage I). Instead, bivariate genetic correlations were estimated between four early vocabulary measures from the ALSPAC sample (expressive vocabulary at 15, 24 and 38 months, as well as receptive vocabulary at 38 months) using individual-level genotype data and bivariate Genome-based Restricted Maximum Likelihood (GREML) analyses[55], a powerful approach based on individual-level genotype data[56,57], as implemented in Genome-wide Complex Trait Analysis (GCTA) software[58]. A genetic relationship matrix (GRM)[58] was

created with PLINK[59] using individuals with a genetic relationship <0.05 ($N_{individuals}$≤6,092) and directly genotyped SNPs only (N_{SNPs}=465,740). Based on the resulting genetic correlation matrix (Table S1), matSpD estimated 2.79 independent effective vocabulary traits. SNP-h^2 estimates for ALSPAC vocabulary traits had overlapping 95%-confidence intervals with LDSC-based SNP-h^2 for early-phase expressive vocabulary, late-phase expressive vocabulary and late-phase receptive vocabulary (see "SNP-heritability estimations").

Single variant association analysis

Within each cohort, vocabulary scores were adjusted for age, sex, age^2 and their interaction effects, as well as ancestry-informative principal components (that differed by cohort) and other study-specific covariates defined by the local GWAS analyst. Adjusted scores were then rank-transformed to achieve normality and to allow for comparisons of genetic effects across different psychological instruments. SNP-vocabulary associations were estimated within each cohort using a linear regression of rank-transformed residuals on posterior genotype probability using SNPTEST[60], Proabel[61], and GEMMA[62] software (Table 2), assuming an additive genetic model, except for the LSAC cohort. For LSAC, a linear regression of rank-transformed residualised vocabulary scores on bestguess genotypes was performed with PLINK 1.9[63], using imputed markers (INFO>0.3) due to data availability (Table 2). GWAS analyses of twin samples were performed using GEMMA[62] following a linear mixed-model approach. This method accounts for relatedness among individuals using a GRM derived from high-quality directly genotyped markers. To capture strong relatedness, GRM off-diagonal elements with values <0.05 were set to zero[64,65].

Prior to the meta-analyses, GWAS summary statistics from all cohorts underwent extensive quality control using the EasyQC R package[66]. Variants that had a low (i) imputation quality (INFO<0.6 for SNPTEST, PLINK and GEMMA association analyses and INFO<0.5 for Proabel association analyses), (ii) minor allele count (MAC≤10), or (iii) effect allele frequency (EAF≤0.005 or EAF≥0.995) were excluded. In addition, marker names were harmonised and alleles were aligned against HRC r1.1 reference data. Variants with missing or mismatching alleles were dropped, as well as all insertions/deletions, duplicate SNPs and multi-allelic SNPs. Finally, variants with an EAF that deviated >0.2 from the frequency in the HRC r1.1. reference data were excluded.

Within an initial step, single-trait meta-analyses were performed for early-phase expressive vocabulary, late-phase expressive vocabulary and late-phase receptive vocabulary (stage I, Figure 1). For early-phase expressive vocabulary and late-phase receptive vocabulary all observations were independent and fixed-effect meta-analyses were carried out, as implemented in METAL software[67]. This approach includes a meta-analysis across effect size estimates reported by each individual study, weighted by the inverse of the corresponding standard error[67]. Late-phase expressive vocabulary scores

included longitudinal assessments of the same ALSPAC children at 24 and 38 months (Table 1). Consequently, fixed-effect meta-analyses were first carried out excluding ALSPAC expressive vocabulary at 38 months to ensure independence of GWAS summary statistics. The derived METAL output was then jointly analysed with the GWAS results for ALSPAC expressive vocabulary at 38 months using multi-trait analysis of genome-wide association (MTAG)[68], allowing for sample overlap, resulting in combined genome-wide association summary statistics for all late-phase expressive vocabulary observations. MTAG exploits genetic relationships among traits and provides a generalised inverse-variance-weighted meta-analysis estimate by integrating GWAS summary statistics across different traits, while allowing for overlapping samples[68]. In order to increase the statistical power to detect single SNP associations, I conducted a multi-trait meta-analysis across all studied vocabulary traits using MTAG[68] (stage II, Figure 1). Analyses were carried out using late-phase expressive vocabulary, the most powerful measure, as outcome. Thus, in addition to three single-trait meta-analyses as part of stage I, I also carried out a multi-trait vocabulary analysis in stage II (Figure 1).

MTAG results derived based on low powered traits (mean χ^2 statistic < 1.02), such as ALSPAC expressive vocabulary at 38 months and late-phase receptive vocabulary, could lead to biased MTAG estimates and an increased FDR rate[68]. Therefore, sensitivity analyses were conducted for both MTAG analyses based on the less powerful fixed-effect meta-analyses for late-phase expressive vocabulary excluding ALSPAC expressive vocabulary at 38 months.

Association analyses were applied with genomic control[69] for variant discovery and without genomic control for follow-up analyses, including LDSC regression and correlation analyses that contain a more accurate and powerful correction factor than genomic control[70]. The sample size and the number of high-quality SNPs included in each meta-analysis can be found in Figure 1.

Locus discovery and annotation

Genome-wide associations were identified from GWAS summary statistic data using Functional Mapping and Annotation of genetic associations[71] software (FUMA, v1.3.6). First, the study-wide GWAS *P*-value threshold for SNP-vocabulary associations was defined at 1.79×10^{-8}. This threshold reflects the genome-wide significance threshold of 5×10^{-8} adjusted for 2.79 independent vocabulary measures analysed in this study, which was estimated based on genetic correlations among vocabulary measures in the ALSPAC sample. Next, independent SNPs were identified based on LD-$r^2 \leq 0.6$ using the 1000 Genomes Phase 3 European reference panel (release 20130502). For SNP signals passing the genome-wide significance threshold without adjustment for the number of independent traits studied ($P < 5 \times 10^{-8}$), allele frequencies were derived from the GnomAD database[72] v2.1.1. based on European (non-Finnish) control samples. To investigate whether vocabulary associated SNPs have pleiotropic effects, I performed a screen

across 2,986 publically available GWASs as included in Phenoscanner[73,74] (v2). For this, the significance threshold was determined at $P<1x10^{-5}$, reflecting a correction for the number of traits included in the screen (0.05/2,986).

Next, I mapped independent genome-wide SNP signals to genes using two different approaches, as implemented within FUMA[71] (v1.3.6). Firstly, genetic variants were mapped to genes using positional mapping, based on a maximum physical distance of 10kb from the SNP to the gene (hg19). Secondly, SNPs were mapped to genes using eQTL mapping. A SNP was mapped to a gene if it mapped within 1Mb of the gene and was known to influence the gene's expression. For these analyses, eQTL information on blood (Blood eQTL browser[75] and BIOS QTL browser[76]) and 13 brain tissues (GTEx v8[77]) were used. The false discovery rate (FDR) for SNP-gene expression associations was defined at 0.05.

Gene-based genome-wide association studies

As part of follow-up analyses, gene-based GWASs were conducted with MAGMA according to a SNP-wide mean model[78], as implemented within FUMA software[71] (v1.3.6a). SNPs were mapped to genes using positional mapping based on the 1000 Genomes Phase 3 European reference panel (release 20130502) and a 0kb window, consistent with default MAGMA settings. In total, SNPs were mapped to 18,828-18,896 protein coding genes, adjusted for the estimated number of analysed traits, corresponding to an adjusted genome-wide gene-based significance threshold of 9.48×10^{-7} (0.05/18,896 genes/2.79 independent vocabulary measures). Gene-based GWAS results subsequently served as input for gene-set and gene-property analyses (see next section).

Table 2: Overview of genotyping, imputation and analysis software

Cohort		ALSPAC	BIS	COPSAC	GenR	LSAC	Raine	TEDs
Genotyping platform		Illumina HumanHap550 quad chip	Illumina Global Screening Array platform	Illumina Infinium HumanOmni ExomeExpress	Illumina 610K	Illumina Infinium® Global Screening Array-24 v1.0	Illuminia Human660W Quad BeadChip	AffymetrixGeneChip (Affy) 6.0 and Human OmniExpress Exome-8v1.2 (OEE)
Quality control	MAF	≥0.01	≥0.01	≥0.01	≥0.01	≥0.01	≥0.01	≥0.01
	SNP Call rate	≥0.99	≥0.99	≥0.95	≥0.95	≥0.95	≥0.95	≥0.99
	HWE	$\geq 5x10^{-7}$	$\geq 5x10^{-7}$	$\geq 1x10^{-5}$	$\geq 1x10^{-5}$	$\geq 5x10^{-7}$	$\geq 1x10^{-6}$	$\geq 1x10^{-4}$
	Individual call rate	≥0.97	≥0.97	≥0.95	≥0.95	≥0.97	≥0.97	≥0.99
	N SNPs genotyped	440,476	451,479	566,755	477,033	468,271	517,183	Affy: 608,517 OEE: 502,434
Imputation	Platform	Sanger Imputation Server	Sanger Imputation Server	Sanger Imputation Server	Michigan Imputation Server	Sanger Imputation Server	Michigan Imputation Server	Sanger Imputation Server
	Reference panel	HRC (r1.1)	HRC (r1.1)	HRC (r1.1)	HRC (r1.1)	HRC (r1.1)	HRC (r1.1)	HRC (r1.1)
	N SNPs imputed	38,691,102	38,680,099	12,106,919	38,640,072	7,109,494	38,381,542	Affy: 22,236,473 OEE: 22,342,556
GWAS	Analysis software	SNPTEST	SNPTEST	SNPTEST	SNPTEST	PLINK	Proabel	GEMMA
	N SNPs	Early-phase EV: 8,663,580 Late-phase EV: 8,665,928 (24m) 8,667,217 (38m) Late-phase RV: 8,667,217	7,244,741	7,795,895	Early-phase EV: 8,610,574 Late-phase EV: 8,607,086	7,518,913	8,654,834	8,293,360

Genotyping data for each cohort was obtained using high-density SNP arrays. Standard genomic quality control procedures were applied and genotypes were imputed against the HRC r1.1. reference panel[34] using either the Sanger imputation server or Michigan imputation server[51]. Association analyses within cohorts of unrelated individuals were performed using SNPTEST[60], PLINK[63] and Proabel[61]. Genome-wide association analyses of related individuals were performed using GEMMA[62]. Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; BIS, Barwon Infant Study; COPSAC, Copenhagen Prospective Studies on Asthma in Childhood; EV, expressive vocabulary; GenR, Generation Rotterdam; GWAS, genome-wide association study; HRC, Haplotype Reference Consortium; HWE, Hardy-Weinberg Equilibrium; LSAC, Longitudinal Study of Australian Children; m, months; MAF, minor allele frequency; Raine, the Western Australian Pregnancy Cohort; RV, receptive vocabulary; SNPs, Single-Nucleotide Polymorphism; TEDS, Twins Early Development Study.

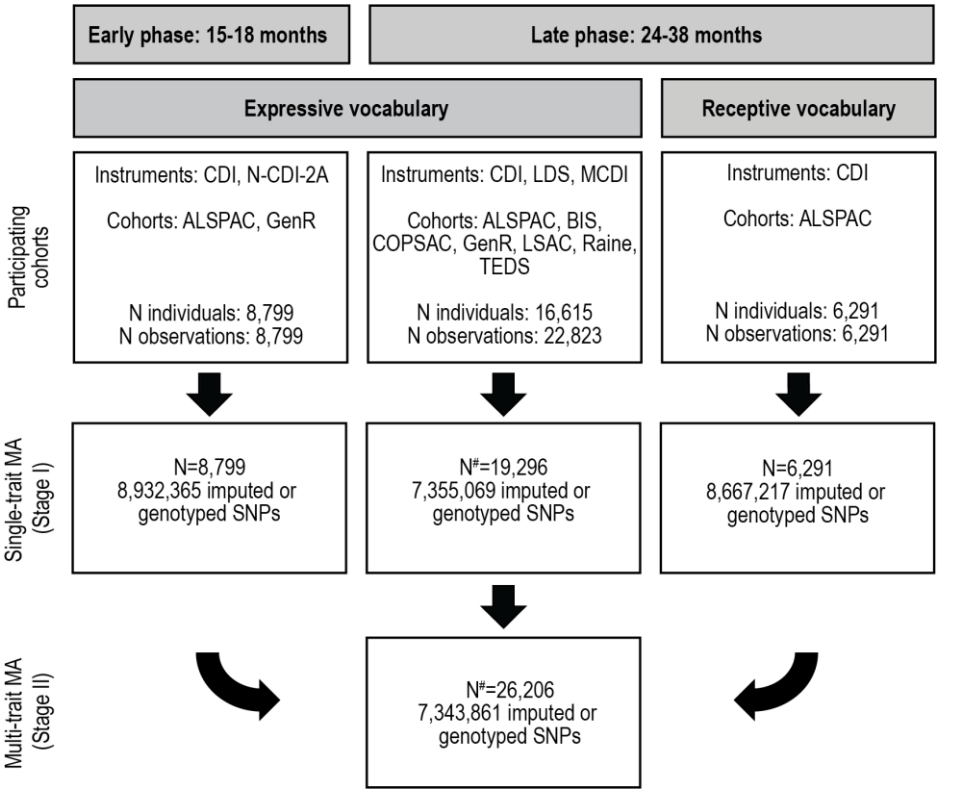


Figure 1: Meta-analysis study design. Vocabulary scores were assessed between 15-38 months of age and divided into an early phase (15-18 months) and late phase (24-38 months) to allow for age-specific genetic influences. Scores for receptive vocabulary were only included in the late-phase due to the low validity of parental reports in very young children. In total, three single-trait meta-analyses were conducted as part of stage I: early-phase expressive vocabulary, late-phase expressive vocabulary and late-phase receptive vocabulary. For each stage the participating cohorts, psychological instruments, total number of observations and individuals, as well as the total number of high-quality SNPs are provided. In order to increase statistical power, multi-trait analysis across early-phase expressive vocabulary, late-phase expressive vocabulary and late-phase receptive vocabulary was performed as part of stage II using multi-trait analysis of genome-wide association. # Estimated sample size based on the increase in mean χ^2 statistic using multi-trait analysis of genome-wide association. Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; BIS, Barwon Infant Study; CDI, Communicative Development Inventory; COPSAC, Copenhagen Prospective Studies on Asthma in Childhood; GenR, Generation Rotterdam; LDS; Language Development Survey; LSAC, Longitudinal Study of Australian Children; MA, meta-analysis; Raine, the Western Australian Pregnancy Cohort; TEDS, Twins Early Development Study.

Gene-set and gene-property analyses

To gain insight into the biological mechanisms tagged by common genetic variation related to early vocabulary development, MAGMA-based gene-set analyses[78] were performed, as implemented within FUMA software[71] (v1.3.6a). This competitive test was conditioned on gene size, gene density, and the inverse of the mean minor allele count in the gene[78]. Association was investigated with up to 10,286 gene-sets and GO terms that were derived from MsigDB[79] v6.2 and contained between 10 and 200 genes to avoid bias related to gene-set size[80]. The multiple-testing-adjusted threshold was defined at $P<1.74\times10^{-6}$ (0.05/10,286 gene-sets/2.79 independent vocabulary measures).

In addition to gene-set analyses, MAGMA[78] gene-property analyses were performed in FUMA[71] (v1.3.6a) to assess whether common genetic variation related to vocabulary was enriched for expression in certain tissues and/or developmental periods. For this, gene expression data from 30 broad tissue types and 54 specific tissues derived from the GTEx v8 RNA-sequencing database[77], as well as gene expression data for 29 different age groupings and 11 developmental stages from the BrainSpan[81] were obtained. The multiple-testing-adjusted threshold was $P<1.45\text{x}10^{-4}$, accounting for the total number of gene expression data sets and independent vocabulary measures (0.05/124/2.79).

SNP-heritability estimations

To estimate SNP-h^2 for early vocabulary measures as captured by GWAS summary statistics, I applied LDSC regression analyses[70]. LDSC regression estimates the proportion of phenotypic variance tagged by SNPs on genotyping arrays, by regressing genome-wide χ^2-statistics on the amount of genetic variation captured by each SNP[70]. The intercept of this regression minus one is an estimator of the mean contribution of confounding bias to the inflation in the mean χ^2-statistic, which is a more powerful and accurate correction compared to genomic control[70].

For comparison, SNP-h^2 was also estimated for early vocabulary measures available in the ALSPAC sample using GREML[58], a powerful approach based on individual-level genotype data[56,57]. This methodology is implemented within GCTA software[58] and was applied using the same GRM as included in bivariate GREML analyses to estimate the independent number of traits analysed (see "Independent number of vocabulary measures").

LD Score genetic correlations

To estimate bivariate genetic correlations (r_g) based on summary statistics, I applied unconstrained LDSC correlation[82] analyses that allowed for sample overlap. This involves a regression of the product of test statistics on LD-score, reflecting the sum of

LD-r^2 for a specific SNP measured with all other SNPs[70], and captures the extent of shared genetic influences between phenotypes assessed in different samples[82]. The intercept of this regression captures potential sample overlap and shared population stratification, resulting in unbiased genetic correlation estimates[82]. I applied LDSC correlation analyses to vocabulary summary statistics only with SNP-h^2 of Z-score > 1.5 to exclude low-powered summary statistics[54]. First, I assessed genetic overlap between early-phase and late-phase receptive vocabulary. Second, I estimated genetic correlations of vocabulary measures with respect to multiple cognition-related later-life outcomes, infancy and childhood anthropometric traits, as well as childhood-onset neurodevelopmental disorders. Cognition-related later-life outcomes included reading performance (8-22 years, N=13,027, Supplementary Information), childhood intelligence[83] (6-18 years, N=12,441), intelligence[84] (5-98 years, N=279,930), and EA[85] (>30 years, N=766,345). For anthropometric traits, I studied birth length[86] (N=28,459), birth weight[87] (N=286,879), infant head circumference[88] (6-30 months, N=10,768) and childhood head circumference[89] (6-9 years, N=10,600). Finally, I investigated genetic overlap between early vocabulary and childhood-onset neurodevelopmental disorders, such as ADHD[90] (N=53,293; N_{cases}=19,099) and ASD[91] (N=46,350; N_{cases}=18,381). ADHD and ASD summary statistics were obtained from the largest GWASs on European individuals available to date[90,91], as well as data reflecting ADHD symptoms in children[92] (<13 years, N=17,666).

All analyses were performed with LDSC software[70,82] and based on the set of well-imputed HapMap3 SNPs and a European reference panel of LD scores[82]. The multiple-testing threshold for LDSC bivariate genetic correlations was defined at ($P \leq 0.005$), reflecting a correction for 8.85 independent vocabulary, cognition-related later-life, anthropometric and childhood-onset neurodevelopmental disorder measures, estimated using matSpD[52,53] and LDSC bivariate genetic correlations (Table S3) (0.05/8.85).

5.3. Results

Single-trait and multi-trait genome-wide association analyses

Single-trait genome-wide association analyses were performed for early-phase expressive vocabulary (15-18 months, N=8,799), late-phase expressive vocabulary (24-38 months, N=16,615) and late-phase receptive vocabulary (24-38 months, N=6,291) (Figure 1). Applying this study design allowed me to distinguish between the production of words in isolation[2] (early phase), followed by the use of word combinations and more complex grammatical structures[5,6] (late phase), as well as between word production and word understanding. In total, I studied up to 37,913 observations from 17,298 children of European origin, collected within seven population-based and community-based

cohorts (Table 1), and between seven and nine million imputed or genotyped SNPs (Figure 1).

Across the three single-trait meta-GWASs, there was no evidence for SNP association with vocabulary size at the multiple-testing-adjusted genome-wide significance level ($P<1.79x10^{-8}$, Figure 2a-c). For early-phase expressive vocabulary, a single GWAS signal (rs9854781) at chr3p12.3 passed the genome-wide significance threshold ($P<5x10^{-8}$, Figure 2a). The T-allele at rs9854781 was associated with an increase of 0.10 standard deviation units in rank-transformed expressive vocabulary scores at 15-18 months of age (β=0.10(SE=0.02), $P=4x10^{-8}$; Table 3). This SNP resides within an intergenic region of the short arm of chromosome 3, ~20 kb downstream of the 3' end of *ROBO2*. rs9854781 is in high LD with rs764282 (LD-r^2=0.78), a known signal for early-phase expressive vocabulary based on a previous meta-GWAS consisting of largely overlapping samples, but an older imputation reference panel[8]. Beyond vocabulary, no phenotypic associations could be identified for rs9854781 passing the multiple-testing corrected threshold of $P<1x10^{-5}$ when performing a screen of 2,986 traits from publicly available GWASs. To functionally characterise the identified genetic variant, I performed eQTL mapping based on SNP-gene expression associations from both blood and brain tissues as implemented in FUMA software. This analysis did not yield evidence for eQTL effects of rs9854781 on *ROBO2* or other protein-coding genes located within 1Mb of rs9854781 ($P>0.05$, data not shown).

Both expressive vocabulary traits had low SNP-h^2, with common genetic variants explaining a small proportion of the phenotypic variation in early-phase (LDSC SNP-h^2=0.12, SE=0.05, Table 4) and late-phase expressive vocabulary (LDSC SNP-h^2=0.09, SE=0.03, Table 3). The heritability of late-phase receptive vocabulary could not be reliably estimated (LDSC SNP-h^2=0.07, SE=0.08, Table 4) and late-phase receptive vocabulary summary statistics captured too little polygenic signal to warrant LDSC genetic correlation analyses (SNP-h^2 Z-score < 1.5)[54]. A GREML-based analysis of SNP-h^2, based on largely the same individuals as included in the GWAS but using individual-level genotype data, provided evidence for low SNP-h^2 of late-phase receptive vocabulary (SNP-h^2=0.12, SE=0.06, Table S2). Consequently, genetic correlation analyses including late-phase receptive vocabulary were performed with bivariate GREML[55] using direct genotypes instead of summary-statistic-based LDSC analyses, given their larger power[56,57].

Genetic correlations among the three early vocabulary single-trait meta-analyses were high, suggesting similarity in underlying genetic architectures. The genetic correlation between early- and late-phase expressive vocabulary was 0.74 (SE=0.21, $P=4x10^{-4}$), as estimated using LDSC. GREML-based genetic correlations of late-phase receptive vocabulary with both available measures of late-phase expressive vocabulary in ALSPAC were 0.85(SE=0.25, P=0.004) and 0.86 (SE=0.15, P=0.004) (Table S1). Consequently, to maximise power for variant discovery, I combined the three single-trait

Table 3: Genome-wide associations with early-phase expressive vocabulary ($P<5x10^{-8}$)

Variant	chr:bp (hg19)	EA	EAF	Fixed-effect meta-analysis			ALSPAC		GenR	
				β(SE)	*P*	P_{het}	β(SE)	*P*	β(SE)	*P*
rs9854781	3:77718732	T	0.78	0.10(0.02)	$4.3x10^{-8}$	0.95	0.10(0.02)	$1.7x10^{-6}$	0.10(0.04)	0.006

Evidence for association at the genome-wide significant level ($P<5x10^{-8}$) was observed for early-phase expressive vocabulary (stage I) using fixed-effect meta-analysis correcting for genomic control. Effect heterogeneity across cohorts was assessed using Cochran's Q-test. Abbreviations: ALSPAC; Avon Longitudinal Study of Parents and Children; bp, base pair position; chr, chromosome; EA, effect allele; EAF, effect allele frequency; GenR, Generation R.

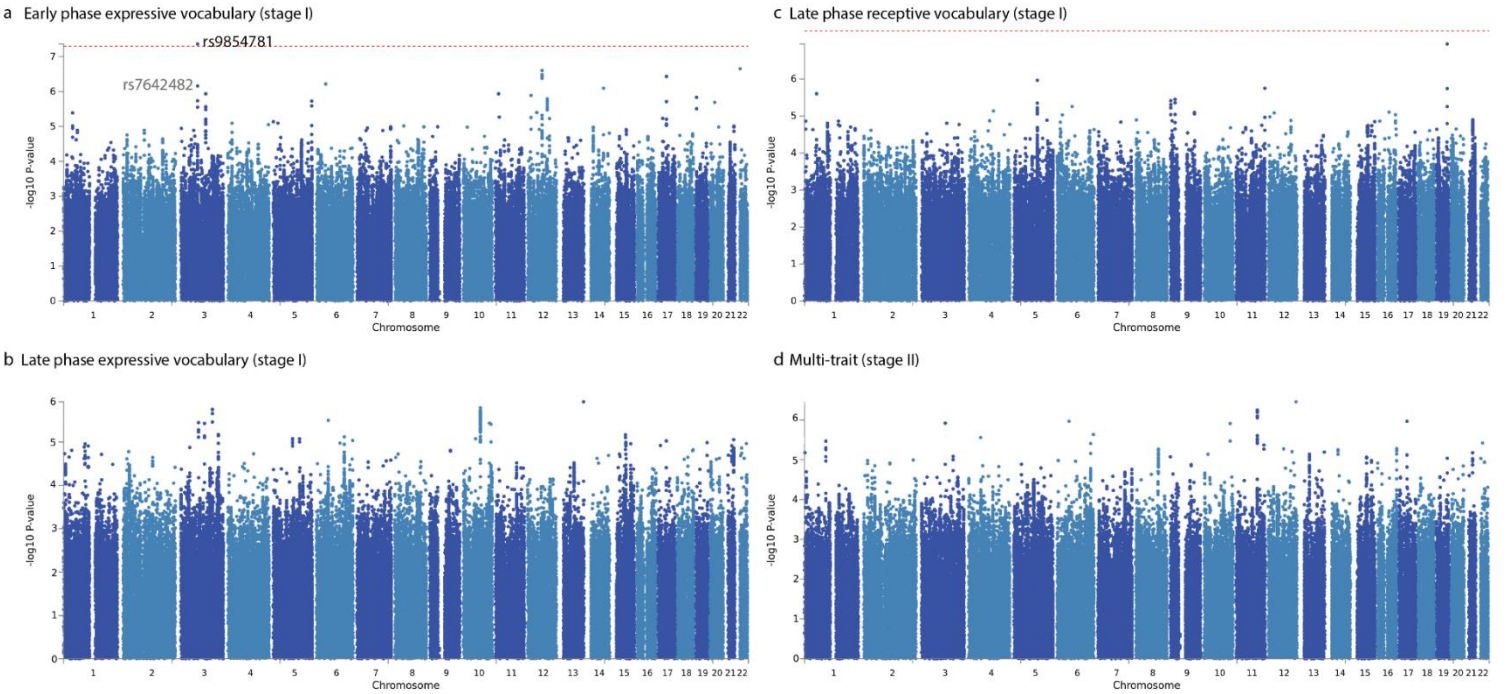


Figure 2: Genome-wide association results for early-phase expressive vocabulary, late-phase expressive vocabulary and receptive vocabulary. Genome-wide association with **(a)** early-phase expressive vocabulary (N=8,799); **(b)** late-phase expressive vocabulary (N=16,615); **(c)** receptive vocabulary (N=6,291) estimated using single-trait meta-analyses as part of stage I. **(d)** Multi-trait genome-wide association results as estimated using MTAG across single-trait vocabulary GWAS results **(a-c)**. No associations passed the genome-wide significance threshold adjusted for the number of independent traits studied of 1.79×10^{-8}. The dashed red line represents the unadjusted genome-wide significance threshold of 5×10^{-8}, variants passing this threshold are labelled in black. Previously associated genetic variants at $P < 5 \times 10^{-8}$ are labelled in gray. Genomic positions are shown according to NCBI Build 37.

meta-analyses as part of a stage II multi-trait meta-analysis (Figure 1). Multi-trait vocabulary MTAG analysis increased both sample size and SNP-h^2 Z-score (estimated-N=26,206, SNP-h^2 Z-score=4.73, Table 4). Despite this power increase, MTAG analysis did not identify further SNP associations passing the unadjusted genome-wide significance level (Figure 2d).

To confirm the reliability of MTAG estimates, MTAG association results for late-phase expressive vocabulary (stage I) and multi-trait vocabulary (stage II) were compared with output derived from fixed-effect meta-analysis for late-phase expressive vocabulary excluding ALSPAC at 38 months. Across shared SNP signals (N=7,343,861) MTAG beta coefficients and standard errors correlated >0.78 and >0.97 with corresponding estimates derived from the fixed-effect meta-analysis. For highly associated SNPs (P<5x10^{-6}, N=37), correlations increased to >0.99 for both beta coefficients and corresponding standard errors. This demonstrates the robustness of MTAG estimates, despite high estimated maximum FDRs (late-phase expressive vocabulary: maxFDR=0.42; multi-trait vocabulary: maxFDR=0.36).

Table 4: SNP-heritability of vocabulary measures based on summary statistics

meta-GWAS	Trait	SNP-h^2 (SE)	SNP-h^2 Z-score	Lambda GC	Mean χ^2	Intercept(SE)	N
Stage I	Early-phase EV	0.12 (0.05)	2.31	1.03	1.03	1.01 (0.01)	8,799
	Late phase EV	0.09 (0.03)	3.44	1.04	1.03	1.00 (0.01)	19,296‡
	Late phase RV	0.07 (0.08)	0.90	1.01	1.01	1.00 (0.01)	6,291
Stage II	Multi-trait	0.12 (0.03)	4.73	1.03	1.04	1.00 (0.01)	26,206‡

As part of stage I, single-trait meta-analyses were carried out for early-phase expressive vocabulary, late-phase expressive vocabulary and late-phase receptive vocabulary. Stage II consisted of multi-trait vocabulary meta-analysis, retrieved by combining data from stage I. Detailed information for each meta-GWAS is provided in Figure 1. SNP-heritability, the intercept of the regression slope, lambda GC and mean χ^2 were estimated with LDSC regression analysis. SNP-h^2 Z-scores were calculated by dividing SNP-h^2 by its standard error. ‡ Estimated sample size based on the increase in mean χ^2 statistic using multi-trait analysis of genome-wide association. Abbreviations: EV, expressive vocabulary; GWAS, genome-wide association study; RV, receptive vocabulary.

Gene-based genome-wide association analyses

Gene-based analyses have increased statistical power to detect associations with a trait of interest compared to GWASs at the SNP level, due to the combination of effects from multiple SNPs and the reduced multiple testing burden[93]. Here, I studied gene-based associations with vocabulary using three single-trait (stage I) and one multi-trait (stage II) meta-analysis summary statistics using MAGMA[78]. However, no gene-level associations passed the significance threshold adjusted for both the number of genes and independent vocabulary measures tested (P<9.48×10^{-7}, Figure 3).

Table 5: MAGMA gene-set and gene-property analyses

Analysis type		meta-GWAS			
		Stage I			Stage II
		Early-phase EV	Late-phase EV	Late-phase RV	Multi-trait
Gene-set	10,286 gene-sets	$P \geq 5x10^{-5}$	$P \geq 2x10^{-5}$	$P \geq 8x10^{-6}$	$P \geq 4x10^{-6}$
Gene-property	GTEx v8 30 broad tissue types	$P \geq 0.09$	$P \geq 0.05$	$P \geq 0.04$	$P \geq 0.04$
	GTEx v8 54 specific tissue types	$P \geq 0.14$	$P \geq 0.13$	$P \geq 0.02$	$P \geq 0.04$
	BrainSpan 29 ages	$P \geq 0.06$	$P \geq 4x10^{-4}$	$P \geq 0.12$	$P \geq 0.01$
	BrainSpan 11 developmental periods	$P \geq 0.07$	$P \geq 0.11$	$P \geq 0.12$	$P \geq 0.22$

MAGMA[78] gene-set and gene-property analyses were performed in FUMA[71] (v1.3.6a). Association with 10,286 gene-sets containing between 10 and 200 genes was tested and the significance threshold adjusted for multiple-testing was determined at $P \leq 1.75x10^{-6}$, correcting for both the number of gene-sets tested and the estimated number of independent traits studied. Gene-property analyses were based on gene expression data from 30 broad tissue types and 54 specific tissue types from the GTEx v8 RNA sequencing database[77]. In addition, gene expression data from 29 different age groupings and 11 developmental stages from the BrainSpan database[81] were utilised. The lowest *P*-value obtained for each association analyses is reported. Gene-property analyses were considered significant if they passed a multiple-testing-adjusted *P*-value threshold of $1.45x10^{-4}$

Gene-set and gene-property analyses

To identify putative biological pathways contributing to the polygenic architecture of early vocabulary, I studied association with 10,286 gene-sets, using MAGMA gene-set analyses[78] based on stage I and stage II vocabulary summary statistics. These analyses did not provide evidence for association at the multiple-testing-adjusted significance threshold of $1.74x10^{-6}$ (Table 5). The strongest association was observed using summary statistics from the most powerful multi-trait analysis for a gene-set consisting of 25 genes that are related to the aggregation, arrangement and bonding of a set of components to form an excitatory synapse (GO:1904861, β=0.81(SE=0.18), $P=4x10^{-6}$). Finally, MAGMA[78] gene-property analyses did not reveal evidence for an association of tissue- and/or age-specific gene expression patterns with early vocabulary measures studied (Table 5). A study of partitioned heritability according to functional category[94] was not feasible due to low power of the derived GWAS summary statistics. According to power calculations included in the original method description[94], sample sizes and SNP-h^2 estimates for vocabulary traits (Table 4) correspond to less than 20% power.

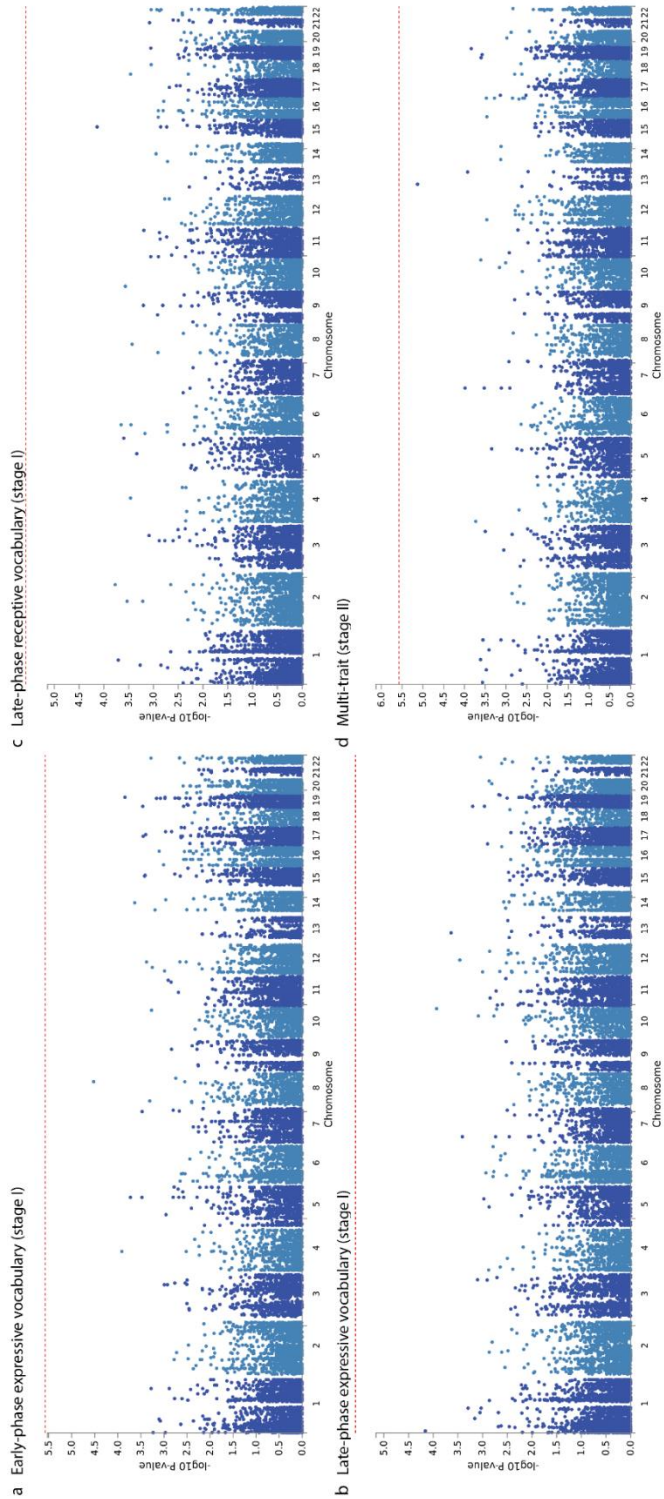


Figure 3: Gene-based GWASs for early-phase expressive vocabulary, late-phase expressive vocabulary and receptive vocabulary. Gene-based genome-wide association analyses for single-trait summary statistics of **(a)** early-phase expressive vocabulary, **(b)** late-phase expressive vocabulary, and **(c)** late-phase receptive vocabulary, as well as **(d)** multi-trait vocabulary summary statistics. The dashed red line represents the significance threshold adjusted for the total number of genes tested ($P<2.6x10^{-6}$). Genomic positions are shown according to NCBI Build 37.

LD Score genetic correlations

Finally, I investigated genetic links between early vocabulary and multiple cognition-related later-life outcomes, as well as infancy and childhood anthropometric traits and childhood-onset neurodevelopmental disorders. Analyses were conducted using single-trait GWAS summary statistics for both early- and late-phase expressive vocabulary as well as multi-trait vocabulary GWAS summary statistics. Late-phase receptive vocabulary summary statistics captured too little polygenic signal to be included in LDSC genetic correlation analyses (Table 4),as described above. For late-phase expressive vocabulary, but not early-phase expressive vocabulary I identified weak-to-moderate positive genetic links with cognition-related later-life outcomes, including reading, intelligence and EA that passed the multiple-testing-adjusted threshold (P<0.005, Figure 4). Genetic correlations of late-phase expressive vocabulary with intelligence across the lifespan (r_g=0.25(SE=0.08), P=3x10^{-4}) and adulthood EA (r_g=0.23(SE=0.06), P=5x10^{-4}) were weak. The genetic overlap of late-phase expressive vocabulary with mid-childhood to adolescence reading ability was moderate (r_g=0.58(SE=0.19), P=0.003) and 95%-confidence intervals overlapped with respective genetic correlation estimates for intelligence and EA (Figure 4). Studying genetic links based on multi-trait vocabulary summary statistics suggested overlap with intelligence (r_g=0.14(SE=0.05), P=0.006) and reading (r_g=0.35(SE=0.16), P=0.03) at the nominal level. I also observed evidence for positive genetic correlations of ADHD with early-phase expressive vocabulary (r_g=0.39(SE=0.14), P=0.006) and multi-trait vocabulary (r_g=0.18(SE=0.08), P=0.02) at the nominal level. However, there was little support for genetic overlap between vocabulary and the studied childhood-onset neurodevelopmental disorders or infant and childhood anthropometric traits at the study-wide threshold (Figure 4).

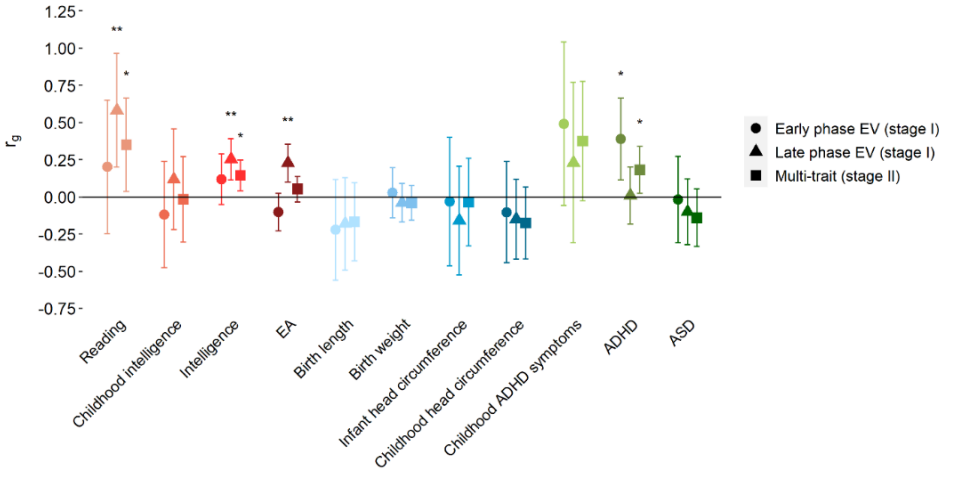


Figure 4: Genetic correlations of vocabulary with cognition-related later-life outcomes, infant and childhood anthropometric traits and childhood-onset neurodevelopmental disorders. Genetic correlations (r_g) were estimated using summary statistics and unconstrained LD score correlation (LDSC)[82]. Bars represent 95%-confidence intervals. Detailed information for each analysis stage is provided in Figure 1. Late-phase receptive vocabulary was not included in LDSC genetic correlation analyses due to its low inherent power (SNP-h^2 Z-score=0.90, Table 4). ** Estimate passing the multiple-testing adjusted *P*-value threshold of 0.005. * Estimate passing the nominal *P*-value threshold of 0.05. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; EA, educational attainment; EV, expressive vocabulary

5.4. Discussion

In this study, I conducted a meta-GWAS of expressive and receptive vocabulary size during infancy and early childhood. I confirmed the low SNP-h^2 for both early-phase expressive vocabulary and late-phase expressive vocabulary, as estimated by earlier studies[8], while SNP-h^2 for late-phase receptive vocabulary based on summary statistics was consistent with zero, possibly due to low power. Genetic correlation analyses suggested shared genetic influences underlying early expressive vocabulary across different developmental phases as well as across expressive and receptive vocabulary traits. In addition, expressive vocabulary size in toddlerhood showed weak-to-moderate polygenic links with cognition-related outcomes during later life, such as adult educational attainment and intelligence across the lifespan. These findings imply that part of the genetic influences related to later-life cognition and educational attainment can already be captured by genetic factors influencing expressive vocabulary at the age of two and three years. I did not observe evidence for single variant associations at the genome-wide significant threshold adjusted for testing of multiple phenotypes ($P<1.79\times10^{-8}$). However, association of rs9854781 with early-phase expressive vocabulary was observed at the unadjusted genome-wide level ($P<5\times10^{-8}$), confirming a previous GWAS signal[8].

The high genetic correlations across different developmental stages for expressive vocabulary suggested genetic stability, consistent with previous findings using both samples of twins and unrelated individuals[8,9,13]. With respect to cognition-related later life outcomes, the current findings support a role for genetic influences related to late-phase expressive vocabulary (24-38 months). Genetic correlation analyses provided evidence for weak-to-moderate positive genetic links between late-phase expressive vocabulary and subsequent reading ability, intelligence and educational attainment. These findings support the outcomes of observational studies reporting the predictive value of early language abilities for both accuracy and comprehension aspects of reading up to five years later[95], as well as academic achievement[96]. Interestingly, the current results imply that polygenic variation contributing to expressive vocabulary during toddlerhood, i.e. at an age where IQ tests cannot be administered yet, tag a subset of the genetic variation that is related to intelligence in later life. However, the estimated genetic correlations between expressive vocabulary and later cognition-related traits were only weak to modest, and thus the predictive power of late-phase expressive vocabulary measures is low. In contrast, genetic correlations between mid-childhood/early-adolescent literacy- and language-related skills and EA are considerably larger, with estimates ranging between 0.57 and 0.89[27] (chapter 6). In my analyses of common genetic variation, I did not find strong evidence for genetic links of early vocabulary with ASD or ADHD, or with the infant and childhood anthropometric traits investigated.

There is little currently known about genetic factors contributing to variation in receptive vocabulary. Scarce information on late-phase receptive vocabulary most likely reflects the absence of a receptive vocabulary scale in frequently used psychological instruments to assess early language development, such as the LDS[44] and CDI-WS[42]. In general, parents are thought to be poorer at judging their child's language comprehension compared to language production[97], as assessing receptive language skills requires that parents notice their children's non-verbal responses to words and is therefore deemed more subjective than assessing expressive language skills. A study of 25-month-old children, however, showed that parents are able to assess receptive vocabulary, with a correlation of 0.55 between parent report and child task performance, highlighting the feasibility of studies investigating early receptive language skills[38]. Indeed, genetic correlation analyses using GREML[55], a more powerful approach when individual-level genotype is available than LDSC[56,57], confirmed the utility of receptive vocabulary scores in toddlers, with strong genetic correlations between late-phase expressive and receptive vocabulary measures.

At the single variant level, a genome-wide significant association signal with early-phase expressive vocabulary (P=4x10^{-8}) was identified on chr3p12.3 at rs9854781, although it did not pass the corrected threshold for testing multiple phenotypes (P<1.79x10^{-8}). This association signal is consistent with findings from a previous meta-

GWAS effort (rs7642482) studying early-phase expressive vocabulary with genotypes from nearly identical samples derived from an earlier imputation reference panel[8]. rs9854781 is in high LD with rs7642482 (LD-r^2=0.78) and is located only 976 base pairs downstream of rs7642482. Both SNPs are located within an intergenic region near *ROBO2*, encoding *Homo sapiens* roundabout homologue 2 (Drosophila), which is an axon guidance receptor that binds to secreted SLIT ligands[98,99]. However, my follow-up *in silico* analyses, including eQTL mapping, did not uncover evidence supporting a functional role of rs9854781 on the expression of *ROBO2* or other protein-coding genes.

Despite an increase in statistical power for the detection of SNPs associated with vocabulary compared to the previous meta-GWAS effort[8], there was no evidence for association of SNPs with late-phase expressive vocabulary, late-phase-receptive vocabulary or multi-trait vocabulary ($P<5x10^{-8}$). The most powerful analyses (multi-trait vocabulary, stage II) had 99% power to detect association with a genetic variant explaining 0.3% of the trait variance (assuming an additive model and an increaser allele frequency of 0.1, with complete LD with marker and genetic risk variant)[100]. However, power to detect variants with smaller contributions to the trait variance is still only modest (e.g. 37% power to detect a genetic variant explaining 0.1% of the trait variance)[100]. Thus, the power of this study is still too low to identify genetic variants with very small effects on vocabulary development. Nonetheless, the derived summary statistics for early-phase expressive vocabulary, late-phase expressive vocabulary and multi-trait vocabulary captured sufficient polygenic signal (SNP-h^2 Z-score > 1.5)[54] to study genetic links with each other and with other complex heritable traits, using LDSC correlation.

This study has several strengths and limitations. Its strengths include maximising the statistical power to detect SNP-vocabulary associations by the application of MTAG analysis for genetically correlated traits. However, MTAG results derived based on low powered traits (mean χ^2 statistic < 1.02) could lead to biased MTAG estimates and an increased FDR rate[68]. Sensitivity analyses suggested that this scenario is unlikely and showed the validity of results. Beta coefficients and standard errors derived from MTAG analyses were strongly correlated with results from a fixed-effect meta-analysis for late-phase expressive vocabulary, despite the inclusion of some vocabulary data with a mean χ^2 statistic < 1.02 and maximum FDR rates of 0.42 and 0.36 for late-phase expressive vocabulary and multi-trait vocabulary, respectively. Residual errors between high- and low-powered studies are uncorrelated with each other and are thus unlikely to bias genetic links with later-life cognition-related abilities as reported in this study, even in the presence of inflated MTAG estimates. Another cautionary note related to our findings is that genetic links between samples of unrelated individuals reflect both direct and indirect genetic effects. The latter include environmental effects that are created based on parental genotypes and that will in turn influence offspring development[101]. A study comparing within- and between-family polygenic scoring prediction showed that

indirect effects may especially contribute to polygenic predictions of cognition-related traits[102]. Thus, the reported genetic links between late-phase expressive vocabulary and subsequent cognition-related traits may not only represent shared genetic variance, but also genotype-environment correlation.

The study described in this chapter represents analyses based on an intermediate freeze of the EAGLE early vocabulary meta-analysis, incorporating cohort data available at the time of writing. Future efforts will benefit from further increasing the sample size to boost statistical power, which may result in identification of more genome-wide significant associations, as observed for example for EA[85,103,104]. Furthermore, statistical power could be increased by applying MTAG analyses beyond early vocabulary measures. Such approaches could, for example, include reading ability, a trait that can be assessed later during development and was shown to be genetically correlated with late-phase expressive vocabulary. In addition, researchers should consider to include other aspects of early language development than vocabulary size in multi-trait analyses. This could include grammatical abilities, for which moderate-to-strong genetic correlations with vocabulary assessments at two and three years of age have been reported based on analyses of a UK twin sample[9,10]. Although broadening the definition of early language skills may increase the power to detect genetic factors underlying shared developmental processes, it may limit the identification of vocabulary-specific genetic influences. Finally, increased verbal abilities of girls compared to boys at 24 months[105] suggests an indirect or direct influence of sex chromosomes.

In summary, this study provided evidence for a contribution (albeit low) of common genetic variation to early- and late-phase expressive vocabulary scores. Genetic correlation analyses suggest a shared genetic aetiology both across early vocabulary measures and between late-phase expressive vocabulary and later-life cognition-related skills, such as reading and intelligence. Despite increased power of this meta-GWAS study compared to a previous effort[8], the statistical power to identify single variants with small effects on early vocabulary development is still low.

Acknowledgements

We are very thankful to all the cohorts that participated in this study, including the Avon Longitudinal Study of Parents and Children (ALSPAC), the Barwon Infant Study (BIS), the Copenhagen Prospective Studies on Asthma in Childhood (COPSAC), Generation Rotterdam (GenR), the Longitudinal Study of Australian Children (LSAC) and its CheckPoint module, the Western Australian Pregnancy Cohort (Raine) and the Twins Early Development Study (TEDS). We would especially like to thank all principal investigators and genome-wide association lead analysts involved in the analyses for the current study.

ALSPAC: We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and they will serve as guarantors for the contents of this paper. GWAS data were generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf).

BIS: We thank the BIS participants for the generous contribution they have made to this project. We also thank current and past staff for their efforts in recruiting and maintaining the cohort and in obtaining and processing the data and biospecimens. The establishment work and infrastructure for the BIS was provided by the Murdoch Children's Research Institute, Deakin University and Barwon Health. Subsequent funding was secured from the National Health and Medical Research Council of Australia, The Shepherd Foundation, The Jack Brockhoff Foundation, the Scobie Trust, the Shane O'Brien Memorial Asthma Foundation, the Our Women's Our Children's Fund Raising Committee Barwon Health, the Rotary Club of Geelong, the Ilhan Food Allergy Foundation, GMHBA Ltd, The Gandel Foundation, The Percy Baxter Charitable Trust, Perpetual Trustees and the Gwenyth Raymond Trust. Vanguard Investments Australia Ltd provided funding for analysis of plasticisers in biospecimens. In-kind support was provided by the Cotton on Foundation and CreativeForce. Research at Murdoch Children's Research Institute is supported by the Victorian Government's Operational Infrastructure Support Program. A-L. Ponsonby, P. Vuillermin, D. Burgner, V. Anderson and P. Sly receive NHMRC fellowship support. C. Symeonides was supported by a NHMRC PhD scholarship. The BIS Investigator Group includes John Carlin, Mimi Tang and Len Harrison.

COPSAC: We express our deepest gratitude to the children and families of the COPSAC cohort study for all their support and commitment. We acknowledge and

appreciate the unique efforts of the COPSAC research team. All funding received by COPSAC is listed on www.copsac.com. The Lundbeck Foundation (Grant no R16-A1694); The Ministry of Health (Grant no 903516); Danish Council for Strategic Research (Grant no 0603-00280B) and The Capital Region Research Foundation have provided core support to the COPSAC research center.

GenR: The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the Erasmus University Rotterdam, Faculty of Social Sciences, the Municipal Health Service Rotterdam area, and the Stichting Trombosedienst and Artsenlaboratorium Rijnmond (STAR), Rotterdam. We gratefully acknowledge the contribution of general practitioners, hospitals, midwives and pharmacies in Rotterdam. The Generation R Study is made possible by financial support from: Erasmus Medical Center, Rotterdam, and the Netherlands Organization for Health Research and Development (ZonMw).

LSAC: The study is conducted in partnership between the Department of Social Services, the Australian Institute of Family Studies, and the Australian Bureau of Statistics. Its biophysical module, the Child Health CheckPoint, was led by researchers at the Murdoch Children's Research Institute and collaborators. The findings and views reported are those of the authors. We thank the LSAC and CheckPoint (Child Health CheckPoint) study participants, staff, and students for their contributions.

Raine: This study was supported by the National Health and Medical Research Council of Australia [grant numbers 572613 and 40398] and the Canadian Institutes of Health Research [grant number MOP-82893]. The authors are grateful to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the NH&MRC for their long term funding to the study over the last 25 years and also the following institutes for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA) , Curtin University, the Raine Medical Research Foundation, the UWA Faculty of Medicine, Dentistry and Health Sciences, the Telethon Kids Institute, the Women's and Infant's Research Foundation (King Edward Memorial Hospital), Murdoch University, The University of Notre Dame (Australia), and Edith Cowan University. The authors gratefully acknowledge the assistance of the Western Australian DNA Bank (National Health and Medical Reserach Council of Australia National Enabling Facility). Andrew Whitehouse is funded by Senior Research Fellowship from the NHMRC (1077966). We would also like to acknowledge the Raine Study participants for their ongoing participation in the study, and the Raine Study Team for study co-ordination and data collection. This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and Government of Western Australia.

TEDS: We gratefully acknowledge the ongoing contribution of the participants in TEDS and their families. TEDS is supported by UK Medical Research Council Program

Grant MR/M021475/1 (and previously Grant G0901245) (to Robert Plomin), with additional support from National Institutes of Health Grant AG046938.

We would also like to thank dbGap for providing access to data from the Philadelphia Neurodevelopmental Cohort (PNC, accession number 12043). Support for the collection of the data for Philadelphia Neurodevelopment Cohort (PNC) was provided by grant RC2MH089983 awarded to Raquel Gur and RC2MH089924 awarded to Hakon Hakonarson. Subjects were recruited and genotyped through the Center for Applied Genomics (CAG) at The Children's Hospital in Philadelphia (CHOP). Phenotypic data collection occurred at the CAG/CHOP and at the Brain Behavior Laboratory, University of Pennsylvania. Cognitive assessment was performed using the Penn Computerized Neurocognitive Battery (CNB), which was customized and shortened for a pediatric population using both classical psychometric and item response theory methods. The samples released were all previously genotyped using Illumina or Affymetrix SNP-array platform by the Center for Applied Genomics at The Children's Hospital of Philadelphia.

This work also made use of data and samples generated by the 1958 Birth Cohort (NCDS), which is managed by the Centre for Longitudinal Studies at the UCL Institute of Education, funded by the Economic and Social Research Council (grant number ES/M001660/1). Access to these resources was enabled via the Wellcome Trust & MRC: 58FORWARDS grant [108439/Z/15/Z] (The 1958 Birth Cohort: Fostering new Opportunities for Research via Wider Access to Reliable Data and Samples). Before 2015 biomedical resources were maintained under the Wellcome Trust and Medical Research Council 58READIE Project (grant numbers WT095219MA and G1001799). Genotyping was undertaken as part of the Wellcome Trust Case-Control Consortium (WTCCC) under Wellcome Trust award 076113, and a full list of the investigators who contributed to the generation of the data is available at www.wtccc.org.uk. Data governance was provided by the METADAC data access committee, funded by ESRC, Wellcome, and MRC. (2015-2018: Grant Number MR/N01104X/1 2018-2020: Grant Number ES/S008349/1).

Finally, this research was facilitated by the Social Science Genetic Association Consortium (SSGAC), Early Growth Genetics (EGG) Consortium, Early Genetics and Life Course Epidemiology (EAGLE) Consortium, Psychiatric Genomics Consortium (PGC) and the Danish Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH) by providing access to genome-wide summary statistics.

References

1. Kennison, S. M. *Introduction to language development*. (SAGA, 2014).
2. Clark, E. V. *First language acquisition*. (Cambridge University Press, 2016).
3. Goldfield, B. A. & Reznick, J. S. Early lexical acquisition: rate, content, and the vocabulary spurt. *J Child Lang* **17**, 171–183 (1990).
4. Frank, M. C., Braginsky, M., Yurovsky, D. & Marchman, V. A. Wordbank: an open repository for developmental vocabulary data. *J Child Lang* **44**, 677–694 (2017).
5. Fenson, L. *et al.* Variability in early communicative development. *Monogr Soc Res Child Dev* **59**, 1–173; discussion 174-85 (1994).
6. Hoff, E. *Language Development*. (Cengage Learning, 2013).
7. Bergelson, E. & Swingley, D. At 6-9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3253–3258 (2012).
8. St Pourcain, B. *et al.* Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature communications* **5**, 4831 (2014).
9. Dionne, G., Dale, P. S., Boivin, M. & Plomin, R. Genetic Evidence for Bidirectional Effects of Early Lexical and Grammatical Development. *Child Development* **74**, 394–412 (2003).
10. Dale, P. S., Dionne, G., Eley, T. C. & Plomin, R. Lexical and grammatical development: a behavioural genetic perspective. *Journal of Child Language* **27**, 619–642 (2000).
11. Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev Sci* **15**, 233–49 (2012).
12. Reznick, J. S., Corley, R. & Robinson, J. A longitudinal twin study of intelligence in the second year. *Monogr Soc Res Child Dev* **62**, i–vi, 1–154; discussion 155-60 (1997).
13. Harlaar, N., Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. Why Do Preschool Language Abilities Correlate With Later Reading? A Twin Study. *J Speech Lang Hear Res* **51**, 688–705 (2008).
14. Lee, J. Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics* **32**, 69–92 (2011).
15. Bornstein, M. H., Hahn, C.-S., Putnick, D. L. & Suwalsky, J. T. D. Stability of Core Language Skill from Early Childhood to Adolescence: A Latent Variable Approach. *Child Development* **85**, 1346–1356 (2014).
16. Bleses, D., Makransky, G., Dale, P. S., Højen, A. & Ari, B. A. Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics* **37**, 1461–1476 (2016).
17. Verhoef, E., Shapland, C. Y., Fisher, S. E., Dale, P. S. & Pourcain, B. S. The amplification of genetic factors for early vocabulary during children's language and literacy development. *bioRxiv* 863118 (2019) doi:10.1101/863118.
18. Geurts, H. M. & Embrechts, M. Language profiles in ASD, SLI, and ADHD. *J Autism Dev Disord* **38**, 1931–1943 (2008).
19. Helland, W. A., Posserud, M.-B., Helland, T., Heimann, M. & Lundervold, A. J. Language Impairments in Children With ADHD and in Children With Reading Disorder. *Journal of Attention Disorders* 1087054712461530 (2012) doi:10.1177/1087054712461530.
20. Germanò, E., Gagliano, A. & Curatolo, P. Comorbidity of ADHD and dyslexia. *Dev Neuropsychol* **35**, 475–493 (2010).
21. Peyre, H. *et al.* Relationship between early language skills and the development of inattention/hyperactivity symptoms during the preschool period: Results of the EDEN mother-child cohort. *BMC Psychiatry* **16**, (2016).
22. Tager-Flusberg, H. *et al.* Handbook of autism and pervasive developmental disorders. (2005).
23. Ozonoff, S., South, M. & Miller, J. N. DSM-IV-Defined Asperger Syndrome: Cognitive, Behavioral and Early History Differentiation from High-Functioning Autism. *Autism* **4**, 29–46 (2000).
24. Martin, N. C., Levy, F., Pieka, J. & Hay, D. A. A Genetic Study of Attention Deficit Hyperactivity Disorder, Conduct Disorder, Oppositional Defiant Disorder and Reading Disability: Aetiological overlaps and implications. *International Journal of Disability, Development and Education* **53**, 21–34 (2006).
25. Willcutt, E. G., Pennington, B. F. & DeFries, J. C. Twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **96**, 293–301 (2000).

26. Willcutt, E. G., Pennington, B. F., Olson, R. K. & DeFries, J. C. Understanding comorbidity: A twin study of reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **144B**, 709–714 (2007).
27. Verhoef, E. *et al.* Disentangling polygenic associations between attention-deficit/hyperactivity disorder, educational attainment, literacy and language. *Translational Psychiatry* **9**, 35 (2019).
28. Madigan, S., Wade, M., Plamondon, A., Browne, D. & Jenkins, J. M. Birth Weight Variability and Language Development: Risk, Resilience, and Responsive Parenting. *J Pediatr Psychol* **40**, 869–877 (2015).
29. Barre, N., Morgan, A., Doyle, L. W. & Anderson, P. J. Language abilities in children who were very preterm and/or very low birth weight: a meta-analysis. *J. Pediatr.* **158**, 766-774.e1 (2011).
30. Flensborg-Madsen, T., Grønkjær, M. & Mortensen, E. L. Predictors of early life milestones: Results from the Copenhagen Perinatal Cohort. *BMC Pediatrics* **19**, 420 (2019).
31. Harris, S. R. Measuring head circumference: Update on infant microcephaly. *Can Fam Physician* **61**, 680–684 (2015).
32. Maunu, J. *et al.* Brain and Ventricles in Very Low Birth Weight Infants at Term: A Comparison Among Head Circumference, Ultrasound, and Magnetic Resonance Imaging. *Pediatrics* **123**, 617–626 (2009).
33. Bartholomeusz, H. H., Courchesne, E. & Karns, C. M. Relationship Between Head Circumference and Brain Volume in Healthy Normal Toddlers, Children, and Adults. *Neuropediatrics* **33**, 239–241 (2002).
34. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
35. Middeldorp, C. M. *et al.* The Early Growth Genetics (EGG) and EArly Genetics and Lifecourse Epidemiology (EAGLE) consortia: design, results and future prospects. *Eur J Epidemiol* **34**, 279–300 (2019).
36. Houston-Price, C., Mather, E. & Sakkalou, E. Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *J Child Lang* **34**, 701–724 (2007).
37. Dale, P. S. The validity of a parent report measure of vocabulary and syntax at 24 months. *J Speech Hear Res* **34**, 565–571 (1991).
38. Ring, E. D. & Fenson, L. The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language* **20**, 141–159 (2000).
39. Fenson, L. *et al*. Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics* **21**, 95–116 (2000).
40. Zink, I. & Lejaegere, M. N-CDI's: korte vormen, Aanpassing en hernormering van de MacArthur Short Form Vocabulary Checklist van Fenson et al.
41. Reznick, J. S. & Goldsmith, L. A multiple form word production checklist for assessing early language. *Journal of Child Language* **16**, 91–100 (1989).
42. Fenson, L. *et al. User's Guide and Technical Manual for the MacArthur Communicative Development Inventories*. (Singular Publishing, 1993).
43. Bleses, D. *et al.* The Danish Communicative Developmental Inventories: validity and main developmental trends. *J Child Lang* **35**, 651–669 (2008).
44. Rescorla Leslie. The Language Development Survey. *Journal of Speech and Hearing Disorders* **54**, 587–599 (1989).
45. Fenson, L. & Marchman, V. A. *MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual*. (Paul H. Brookes Publishing Company, 2007).
46. Rescorla, L., Ratner, N. B., Jusczyk, P. & Jusczyk, A. M. Concurrent validity of the language development survey: associations with the MacArthur-Bates communicative development inventories: words and sentences. *Am J Speech Lang Pathol* **14**, 156–163 (2005).
47. Dale, P. S. & Goodman, J. C. Commonality and Individual Differences in Vocabulary Growth. in *Beyond nature-nurture: Essays in honor of Elizabeth Bates* 41–78 (Lawrence Erlbaum Associates Publishers, 2005).
48. Bleses, D. *et al.* Early vocabulary development in Danish and other languages: a CDI-based comparison. *J Child Lang* **35**, 619–650 (2008).
49. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
50. Durbin, R. Efficient haplotype matching and storage using the positional Burrows-Wheeler transform (PBWT). *Bioinformatics* **30**, 1266–1272 (2014).

51. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
52. Nyholt, D. R. A Simple Correction for Multiple Testing for Single-Nucleotide Polymorphisms in Linkage Disequilibrium with Each Other. *Am J Hum Genet* **74**, 765–769 (2004).
53. Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)* **95**, 221–227 (2005).
54. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
55. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540–2542 (2012).
56. Ni, G., Moser, G., Wray, N. R. & Lee, S. H. Estimation of Genetic Correlation via Linkage Disequilibrium Score Regression and Genomic Restricted Maximum Likelihood. *Am J Hum Genet* **102**, 1185–1194 (2018).
57. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* **20**, 567–581 (2019).
58. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
59. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* **81**, 559–575 (2007).
60. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906–913 (2007).
61. Aulchenko, Y. S., Struchalin, M. V. & van Duijn, C. M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134 (2010).
62. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).
63. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, (2015).
64. Tucker, G. *et al.* Two-Variance-Component Model Improves Genetic Prediction in Family Datasets. *Am. J. Hum. Genet.* **97**, 677–690 (2015).
65. Zaitlen, N. *et al.* Using Extended Genealogy to Estimate Components of Heritability for 23 Quantitative and Dichotomous Traits. *PLoS Genet* **9**, (2013).
66. Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192–212 (2014).
67. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
68. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics* **50**, 229–237 (2018).
69. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
70. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* **47**, 291–5 (2015).
71. Watanabe, K., Taskesen, E., Bochoven, A. van & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications* **8**, 1826 (2017).
72. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
73. Kamat, M. A. *et al.* PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).
74. Staley, J. R. *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* **32**, 3207–3209 (2016).
75. Westra, H.-J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature Genetics* **45**, 1238–1243 (2013).
76. Zhernakova, D. V. *et al.* Identification of context-dependent expression quantitative trait loci in whole blood. *Nat. Genet.* **49**, 139–145 (2017).

77. Aguet, F. *et al.* The GTEx Consortium atlas of genetic regulatory effects across human tissues. *bioRxiv* 787903 (2019) doi:10.1101/787903.
78. Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology* **11**, e1004219 (2015).
79. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
80. Mooney, M. A. & Wilmot, B. Gene Set Analysis: A Step-By-Step Guide. *Am J Med Genet B Neuropsychiatr Genet* **168**, 517–527 (2015).
81. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, (2018).
82. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236–1241 (2015).
83. Benyamin, B. *et al.* Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. *Mol Psychiatry* **19**, 253–258 (2014).
84. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics* **50**, 912–919 (2018).
85. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* **50**, 1112–1121 (2018).
86. van der Valk, R. J. P. *et al.* A novel common variant in DCST2 is associated with length in early life and height in adulthood. *Hum. Mol. Genet.* **24**, 1155–1168 (2015).
87. Warrington, N. M. *et al.* Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nat. Genet.* **51**, 804–814 (2019).
88. Taal, H. R. *et al.* Common variants at 12q15 and 12q24 are associated with infant head circumference. *Nat. Genet.* **44**, 532–538 (2012).
89. Haworth, S. *et al.* Low-frequency variation in TP53 has large effects on head circumference and intracranial volume. *Nature Communications* **10**, 357 (2019).
90. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics* 1 (2018) doi:10.1038/s41588-018-0269-7.
91. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics* **51**, 431 (2019).
92. Middeldorp, C. M. *et al.* A Genome-Wide Association Meta-Analysis of Attention-Deficit/Hyperactivity Disorder Symptoms in Population-Based Pediatric Cohorts. *J Am Acad Child Adolesc Psychiatry* **55**, 896-905.e6 (2016).
93. Liu, J. Z. *et al.* A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* **87**, 139–145 (2010).
94. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228–35 (2015).
95. Duff, F. J., Reen, G., Plunkett, K. & Nation, K. Do infant vocabulary skills predict school-age language and literacy outcomes? *J Child Psychol Psychiatry* **56**, 848–856 (2015).
96. Hohm, E., Jennen-Steinmetz, C., Schmidt, M. H. & Laucht, M. Language development at ten months. Predictive of language outcome and school achievement ten years later? *Eur Child Adolesc Psychiatry* **16**, 149–56 (2007).
97. Tomasello, M. & Mervis, C. B. The Instrument Is Great, but Measuring Comprehension Is Still a Problem. *Monographs of the Society for Research in Child Development* **59**, 174–179 (1994).
98. Fricke, C., Lee, J.-S., Geiger-Rudolph, S., Bonhoeffer, F. & Chien, C.-B. astray, a Zebrafish roundabout Homolog Required for Retinal Axon Guidance. *Science* **292**, 507–510 (2001).
99. Simpson, J. H., Kidd, T., Bland, K. S. & Goodman, C. S. Short-Range and Long-Range Guidance by Slit and Its Robo Receptors: Robo and Robo2 Play Distinct Roles in Midline Guidance. *Neuron* **28**, 753–766 (2000).
100. Purcell, S., Cherny, S. S. & Sham, P. C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003).
101. Kong, A. *et al.* The nature of nurture: Effects of parental genotypes. *Science* **359**, 424–428 (2018).
102. Selzam, S. *et al.* Comparing Within- and Between-Family Polygenic Score Prediction. *The American Journal of Human Genetics* **105**, 351–363 (2019).

103. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–42 (2016).
104. Rietveld, C. A. *et al.* GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science* **340**, 1467–1471 (2013).
105. Galsworthy, M. J., Dionne, G., Dale, P. S. & Plomin, R. Sex differences in early verbal and non-verbal cognitive development. *Developmental Science* **3**, 206–215 (2000).

Supplementary Materials

Supplementary Methods

Genome-wide association summary statistics on reading

GWAS summary statistics on reading (N=13,027) were derived by conducting a fixed-effect meta-analysis of reading abilities as assessed in the Avon Longitudinal Study of Parents and Children[1,2] (ALSPAC, N=4,247), the 1958 Birth Cohort[3] (1958BC, N=4,638) and Philadelphia Neurodevelopmental Cohort[4,5] (PNC, N=4,142). Within ALSPAC, word reading speed was assessed at 13 years using the Test of Word Reading Efficiency[6] (TOWRE). Within the PNC, reading accuracy was assessed in participants between 8 and 22 years of age using the reading items of the Wide Range Achievement Test[7] and within the 1958BC, reading comprehension was assessed at 11 years of age using a study-specific reading comprehension test designed to parallel the Watts-Vernon test of reading ability. Here, the child was required to choose from a selection of five words the word that appropriately completed the sentence. There were 35 questions in total and the reliability coefficient of this test is 0.82. Reading scores were adjusted for sex, age, age^2, the first two principal components and study-specific covariates such as batch, if applicable. For each cohort, genome-wide genotyping data were imputed against the HRC r1.1 reference panel[8] and association tests were performed using SNPTEST[9] (version 2.5.2). Finally, a fixed-effect meta-analysis across all three cohorts was performed using METAL[10] (N=13,027).

Supplementary Tables

Table S1: Bivariate genetic correlations among ALSPAC vocabulary traits

	Expressive vocabulary 15m (CDI)	Expressive vocabulary 24m (CDI)	Expressive vocabulary 38m (CDI)	Receptive vocabulary 38m (CDI)
Expressive vocabulary 15m (CDI)	1			
Expressive vocabulary 24m (CDI)	0.57 (0.23)	1		
Expressive vocabulary 38m (CDI)	0.39 (0.27)	0.74 (0.16)	1	
Receptive vocabulary 38m (CDI)	-0.01 (0.36)	0.85 (0.25)	0.86 (0.15)	1

Bivariate genetic correlations among vocabulary traits assessed within the Avon Longitudinal Study of Parents and Children study were assessed using a genetic-relationship-matrix of unrelated individuals (genetic relatedness < 0.05) estimated based on directly genotyped markers, as implemented within GCTA[11]. Corresponding standard errors are shown in brackets. Abbreviations: CDI, Communicative Development Inventory; m, months.

Table S2: SNP heritability estimates for ALSPAC vocabulary measures based on individual-level genotype data

ALSPAC vocabulary measure	SNP-h^2(SE)	SNP-h^2 Z-score	N
Expressive vocabulary 15m (CDI)	0.11(0.05)	3.14	6,524
Expressive vocabulary 24m (CDI)	0.16(0.06)	2.77	6,014
Expressive vocabulary 38m (CDI)	0.18(0.06)	2.06	6,092
Receptive vocabulary 38m (CDI)	0.12(0.06)	2.01	6,092

SNP-heritability estimates were estimated based on rank-transformed scores, directly genotyped SNPs and individuals with a genetic relationship of <0.05 using Genome-based Restricted Maximum Likelihood (GREML) analyses as implemented in genome-wide complex trait analysis (GCTA) software. SNP-h^2 Z-scores were calculated by dividing SNP-h^2 by its standard error. Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; CDI, Communicative Development Inventory; GCTA, genome-wide complex trait analysis; m, months.

Table S3: Genetic correlations among traits included in LDSC correlation analyses

	Early-phase EV (stage I)	Late phase EV (stage I)	Multi-trait vocabulary (stage II)	Reading	Childhood intelligence	Intelligence	EA	Birth length	Birth weight	Infant head circumference	Childhood head circumference	Childhood ADHD symptoms	ADHD
Late phase EV (stage I)	0.74 (0.21)												
Multi-trait vocabulary (stage II)	0.97 (0.15)	0.92 (0.04)											
Reading	0.20 (0.23)	0.58 (0.19)	0.35 (0.16)										
Childhood intelligence	-0.12 (0.18)	0.12 (0.17)	-0.02 (0.15)	0.89 (0.16)									
Intelligence	0.12 (0.09)	0.25 (0.07)	0.14 (0.05)	0.91 (0.1)	0.81 (0.07)								
EA	-0.10 (0.06)	0.23 (0.06)	0.05 (0.04)	0.85 (0.08)	0.74 (0.06)	0.72 (0.01)							
Birth length	-0.22 (0.17)	-0.18 (0.16)	-0.17 (0.13)	0.16 (0.12)	0.04 (0.10)	0.12 (0.04)	0.13 (0.04)						
Birth weight	0.03 (0.09)	-0.04 (0.07)	-0.04 (0.06)	0.11 (0.06)	0.10 (0.06)	0.11 (0.02)	0.11 (0.02)	0.77 (0.05)					
Infant head circumference	-0.03 (0.22)	-0.16 (0.19)	-0.04 (0.15)	0.45 (0.15)	0.38 (0.14)	0.24 (0.06)	0.24 (0.04)	0.55 (0.12)	0.37 (0.07)				
Childhood head circumference	-0.10 (0.17)	-0.15 (0.14)	-0.18 (0.12)	0.39 (0.12)	0.23 (0.12)	0.23 (0.04)	0.26 (0.04)	0.49 (0.10)	0.44 (0.05)	1.40 (0.13)			
Childhood ADHD symptoms	0.49 (0.28)	0.23 (0.28)	0.37 (0.21)	-0.47 (0.22)	-0.32 (0.20)	-0.57 (0.11)	-0.59 (0.12)	-0.16 (0.15)	-0.11 (0.08)	-0.46 (0.19)	-0.15 (0.17)		
ADHD	0.39 (0.14)	0.01 (0.10)	0.18 (0.08)	-0.44 (0.09)	-0.41 (0.08)	-0.36 (0.03)	-0.52 (0.03)	-0.21 (0.07)	-0.10 (0.03)	-0.14 (0.09)	-0.23 (0.07)	0.95 (0.22)	
ASD	-0.02 (0.15)	-0.10 (0.11)	-0.14 (0.10)	0.19 (0.09)	0.23 (0.07)	0.22 (0.03)	0.21 (0.03)	0.13 (0.09)	0.01 (0.04)	0.07 (0.09)	0.08 (0.07)	0.32 (0.14)	0.34 (0.05)

Genetic correlations were estimated with unconstrained LD-score correlation analyses[12]. Standard errors are provided within brackets. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; EA, educational attainment; EV, expressive vocabulary

Supplementary References

1. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* **42**, 97–110 (2013).
2. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111–127 (2013).
3. Power, C. & Elliott, J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol* **35**, 34–41 (2006).
4. Satterthwaite, T. D. *et al.* The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *Neuroimage* **124**, 1115–1119 (2016).
5. Glessner, J. T. et al. Strong synaptic transmission impact by copy number variations in schizophrenia. Proc. Natl. Acad. Sci. 107, 10584–10589 (2010)
6. *TOWRE, Test of Word Reading Efficiency: Examiner's Manual*. (PRO-ED, 1999).
7. Gur, R. C. *et al.* Neurocognitive growth charting in psychosis spectrum youths. *JAMA Psychiatry* **71**, 366–374 (2014).
8. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
9. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906–913 (2007).
10. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
11. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
12. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236–1241 (2015).

Chapter 6

Disentangling polygenic associations between Attention-Deficit/Hyperactivity Disorder, educational attainment, literacy and language

This chapter has been published as:
Verhoef, E., Demontis, D., Burgess, S., Shapland, C.Y., Dale, P.S., Okbay, A., Neale, B.M., Faraone, S.V., iPSYCH-Broad-PGC ADHD Consortium, Stergiakouli, E., Davey Smith, G., Fisher, S.E., Børglum, A.D., St Pourcain, B. Disentangling polygenic associations between attention-deficit/hyperactivity disorder, educational attainment, literacy and language. *Translational Psychiatry* **9**, 35 (2019). doi: 10.1038/s41398-018-0324-2

Abstract

Interpreting polygenic overlap between ADHD and both literacy- and language-related impairments is challenging as genetic associations might be influenced by indirectly shared genetic factors. Here, we investigate genetic overlap between polygenic ADHD risk and multiple literacy- and language-related abilities (LRAs), as assessed in UK children (N≤5,919), accounting for genetically predictable educational attainment (EA). Genome-wide summary statistics on clinical ADHD and years of schooling were obtained from large consortia (N≤326,041). Our findings show that ADHD-polygenic scores (ADHD-PGS) were inversely associated with LRAs in ALSPAC, most consistently with reading-related abilities, and explained ≤1.6% phenotypic variation. These polygenic links were then dissected into both ADHD effects shared with and independent of EA, using multivariable regressions (MVR). Conditional on EA, polygenic ADHD risk remained associated with multiple reading and/or spelling abilities, phonemic awareness and verbal intelligence, but not listening comprehension and non-word repetition. Using conservative ADHD-instruments (*P*-threshold<$5x10^{-8}$), this corresponded, for example, to a 0.35 SD decrease in pooled reading performance per log-odds in ADHD-liability (*P*=$9.2x10^{-5}$). Using subthreshold ADHD-instruments (*P*-threshold<0.0015), these effects became smaller, with a 0.03 SD decrease per log-odds in ADHD risk (*P*=$1.4x10^{-6}$), although the predictive accuracy increased. However, polygenic ADHD-effects shared with EA were of equal strength and at least equal magnitude compared to those independent of EA, for all LRAs studied, and detectable using subthreshold instruments. Thus, ADHD-related polygenic links with LRAs are to a large extent due to shared genetic effects with EA, although there is evidence for an ADHD-specific association profile, independent of EA, that primarily involves literacy-related impairments.

6.1. Introduction

Children with Attention-Deficit/Hyperactivity Disorder (ADHD) often experience difficulties mastering literacy- and language-related abilities (LRAs)[1–3]. It has been estimated that up to 40% of children diagnosed with clinical ADHD also suffer from reading disability (RD, also known as developmental dyslexia) and vice versa[4]. The spectrum of affected LRAs in ADHD may, however, also include writing[5,6], spelling[7,8], syntactic[9,10] and phonological[9,10] abilities. Both clinical ADHD and RD are complex childhood-onset neurodevelopmental conditions that affect about 5% and 7% of the general population respectively[11,12]. ADHD is characterised by hyperactive, inattentive and impulsive symptoms[13], whereas decoding and/or reading comprehension deficits are prominent in individuals with RD[14].

To interpret the comorbidity of ADHD and RD, a multiple-deficit model including shared underlying aetiologies has been proposed, involving both genetic and environmental influences[15]. This model is supported by twin studies suggesting that the co-occurrence of ADHD symptoms and reading deficits is, to a large extent, attributable to shared genetic influences[16–18]. Further twin research suggests that the genetic covariance between reading difficulties and ADHD is largely independent of genetic factors shared with IQ[19], although it is not known whether these findings extend to a wider spectrum of LRAs, beyond reading abilities. Furthermore, the interpretation of polygenic ADHD-LRA overlap using markers on genotyping arrays is more challenging. There is strong evidence that genetically predicted educational attainment (EA)[20] shares genetic variability with both ADHD[21] and reading abilities[22,23]. Genetically predicted EA is a genetic proxy of cognitive abilities, but also socioeconomic status[20] including, for example, associations with maternal smoking during pregnancy, parental smoking, household income or watching television[24]. Thus, observed genetic associations between ADHD and reading abilities may solely reflect shared genetic variation with EA, but not any other, more specific neuro-cognitive mechanisms. In other words, polygenic associations might be inflated or even induced[25] by genetically predictable traits that are related to both, ADHD and reading abilities (or other LRAs).

Here, we (a) study polygenic links between clinical ADHD and a wide range of population-ascertained literacy- and language-related measures as captured by common variation, (b) evaluate to what extent such links reflect a shared genetic basis with EA and (c) assess whether there is support for shared genetic factors between clinical ADHD and LRAs conditional on genetically predicted EA.

Studied ADHD polygenic scores (ADHD-PGS) are based on ADHD genome-wide association study (GWAS) summary statistics from two large independent ADHD samples, the Psychiatric Genomics Consortium (PGC) and the Danish Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), and a combination thereof. Associations between ADHD-PGS and a wide spectrum of population-based

literacy- and language-related measures related to reading, spelling, phonemic awareness, listening comprehension, non-word repetition and verbal intelligence skills, are examined in a sample of children from the UK Avon Longitudinal Study of Parents and Children (ALSPAC). Applying multivariable regression (MVR) techniques, analogous to Mendelian Randomization (MR) approaches[26], we report here disentangled associations between polygenic ADHD risk and LRA measures and estimate effects independent of and shared with genetically predicted years of schooling, using summary statistics from the Social Science Genetic Association Consortium (SSGAC).

6.2. Methods and Materials

Literacy- and language-related abilities in the general population

LRAs were assessed in children and adolescents from ALSPAC, a UK population-based longitudinal pregnancy-ascertained birth cohort (estimated birth date: 1991-1992, Supplementary Information)[27,28]. Ethical approval was obtained from the ALSPAC Law-and-Ethics Committee (IRB00003312) and the Local Research-Ethics Committees. Written informed consent was obtained from a parent or individual with parental responsibility and assent (and for older children consent) was obtained from the child participants.

Phenotype information: Thirteen measures capturing LRAs related to reading, spelling, phonemic awareness, listening comprehension, non-word repetition and verbal intelligence scores were assessed in 7 to 13 year-old ALSPAC participants (N≤5,919, Table 1) using both standardised and ALSPAC-specific instruments. Detailed descriptions of all LRA measures are available in Table 1 and the Supplementary Information.

All LRA scores were rank-transformed to allow for comparisons of genetic effects across different psychological instruments with different distributions (Supplementary Information). Phenotypic correlations, using Pearson-correlation coefficients, were comparable for untransformed and rank-transformed scores (Table S1). To account for multiple testing, we estimated the effective number of phenotypes studied using Matrix Spectral Decomposition[29] (MatSpD), revealing seven independent measures (experiment-wide error rate of 0.007).

For sensitivity analysis, we excluded 188 children with an ADHD diagnosis at age 7, based on the Development and Wellbeing Assessment (DAWBA)[30] (Supplementary Information).

Genetic analyses*:* ALSPAC participants were genotyped using the Illumina HumanHap550 quad chip genotyping platforms, and genotypes were called using the Illumina GenomeStudio software. Genotyping, imputation and genome-wide association analysis details are described in the Supplementary Information and Table 2.

Table 1: Literacy- and language-related abilities in the Avon Longitudinal Study of Parents and Children

LRA (psychological instrument)	Mean Score (SE)	Mean Age (SE)	N (%males)	LRA combinations
Reading accuracy and comprehension (WORD[69]), words	28.44 (9.24)	7.53 (0.31)	5,891 (50.6)	Reading; All LRAs
Reading accuracy (ALSPAC specific: NBO[70]), words	7.55 (2.44)	9.87 (0.32)	5,738 (49.3)	Reading; All LRAs
Reading speed‡ (NARA II[71]), passages	105.50 (12.47)	9.88 (0.32)	5,189 (49.1)	Reading; All LRAs
Reading accuracy‡ (NARA II[71]), passages	104.11 (13.62)	9.88 (0.32)	5,201 (49.1)	Reading; All LRAs
Reading speed (TOWRE[72]), words	82.58 (10.28)	13.83 (0.20)	4,247 (48.4)	Reading; All LRAs
Non-word reading accuracy (ALSPAC specific: NBO[70])	5.24 (2.48)	9.87 (0.32)	5,731 (49.2)	Reading; All LRAs
Non-word reading speed (TOWRE[72])	50.82 (9.38)	13.83 (0.20)	4,237 (48.3)	Reading; All LRAs
Spelling accuracy (ALSPAC specific: NB)	7.89 (4.39)	7.53 (0.31)	5,800 (50.2)	Spelling; All LRAs
Spelling accuracy (ALSPAC specific: NB)	10.27 (3.43)	9.87 (0.32)	5,728 (49.2)	Spelling; All LRAs
Phonemic awareness (AAT[73])	20.23 (9.51)	7.53 (0.31)	5,919 (50.6)	All LRAs
Listening comprehension (WOLD[74])	7.50 (1.96)	8.63 (0.30)	5,473 (49.9)	All LRAs
Non-word repetition (CNRep[75])	7.26 (2.51)	8.63 (0.30)	5,464 (49.9)	All LRAs
Verbal intelligence‡ (WISC-III[76])	107.85 (16.74)	8.64 (0.31)	5,456 (49.7)	All LRAs

Thirteen LRAs capturing aspects related to reading, spelling, phonemic awareness, listening comprehension, non-word repetition and verbal intelligence were assessed in 7 to 13 year-old ALSPAC participants using both standardised and ALSPAC-specific instruments (Supplementary Information). ‡ Scores were derived using age norms and adjusted for sex and principal components only before transformation. Abbreviations: LRAs, literacy- and language-related abilities; WORD, Wechsler Objective Reading Dimension; ALSPAC, Avon Longitudinal study of Parents and Children; NBO, ALSPAC-specific assessment developed by Nunes, Bryant and Olson; NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NB, ALSPAC-specific assessment developed by Nunes and Bryant; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; WISC-III, Wechsler Intelligence Scale for Children III.

Clinical ADHD summary statistics

Psychiatric Genomics Consortium (PGC). GWAS summary statistics were obtained from a mega-analysis of clinical ADHD[31], conducted by the PGC (4,163 cases and 12,040 controls/pseudo-controls) (Table 2, Supplementary Information, www.med.unc.edu/pgc/).

The Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH). An independent set of ADHD GWAS summary statistics were accessed through the Danish iPSYCH project[32] (14,584 ADHD cases, 22,492 controls) (Table 2, Supplementary Information), using samples from the Danish Neonatal Screening Biobank hosted by Statens Serum Institute[21,33].

Combined PGC and iPSYCH ADHD sample (PGC+iPSYCH). To maximise power, we also analysed meta-GWAS summary statistics from an ADHD sample containing both PGC and iPSYCH participants[21] (20,183 cases, 35,191 controls/pseudo-controls) (Table 2, www.med.unc.edu/pgc/) and its European-only subset (*PGC+iPSYCH(EUR),* 19,099 cases, 34,194 controls/pseudo-controls) (Table 2, www.med.unc.edu/pgc/).

Detailed sample descriptions are available in Table 2 and the Supplementary Information.

Educational attainment summary statistics

GWAS summary statistics for EA[20] (discovery and replication sample combined, excluding ALSPAC and 23andMe samples, N=326,041) were obtained from the SSGAC consortium. EA was assessed as years of schooling[20]. A detailed sample description is available in Table 2 and the Supplementary Information.

Genome-wide complex trait analysis

SNP-h^2 and genetic correlations (r_g) between LRAs were estimated using Restricted Maximum Likelihood (REML) analyses[34,35] as implemented in Genome-wide Complex Trait Analysis (GCTA) software[36], including individuals with a genetic relationship <0.05[34]. For this study, we selected only LRAs with evidence for SNP-h^2 and sample size N>4,000 (Table S2).

Linkage Disequilibrium Score regression and correlation

Linkage Disequilibrium Score (LDSC) regression[37] was used to distinguish confounding biases from polygenic influences by examining the LDSC regression intercept. Unconstrained LD-score correlation[38] analysis was applied to estimate r_g (Supplementary Information).

Table 2: Sample description

Phenotype	Sample	Source	Ethnicity	Imputation reference panel	N
LRAs	ALSPAC	General population	White European	HRC r1.1	≤ 5,891
ADHD	PGC	Clinical sample	Predominantly white European	HapMap phase 3	16,203 (N_{cases}=4,163)
	iPSYCH	Clinical sample	White European	1000 Genomes phase 3	37,076 (N_{cases}=14,584)
	PGC+iPSYCH (EUR)	Clinical sample	White European	1000 Genomes phase 3	53,293 (N_{cases}=19,099)
	PGC+iPSYCH	Clinical sample	Predominantly white European	1000 Genomes phase 3	55,374 (N_{cases}=20,183)
EA	SSGAC	Predominantly general population	White European	1000 Genomes phase 3[a]	326,041

Note that there is no overlap between LRA, ADHD and EA samples. a - Predominantly 1000 Genomes phase 3[20]. Abbreviations: LRAs, literacy- and language-related abilities; ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; ALSPAC, Avon Longitudinal study of Parents and Children; PGC, Psychiatric Genomics Consortium; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; EUR, European ancestry; SSGAC, Social Science Genetic Consortium; HRC, The Haplotype Reference Consortium

Polygenic scoring analyses

ADHD-PGS[39,40] were created in ALSPAC using the independent PGC and iPSYCH GWAS summary statistics, and, to maximise power, also for GWAS summary statistics from the combined PGC+iPSYCH sample (Supplementary Information). ADHD-PGS have been previously linked to ADHD symptoms in ALSPAC participants[41]. Rank-transformed LRAs were regressed on Z-standardised ADHD-PGS (aligned to measure risk-increasing alleles) using ordinary least square (OLS) regression (R:stats library, Rv3.2.0). The proportion of phenotypic variance explained is reported as OLS-regression-R^2. Beta-coefficients (β) for ADHD-PGS quantify here the change in standard deviation (SD) units of LRA performance per one SD increase in ADHD-PGS.

Multivariable regression analysis

To study the genetic association between ADHD and LRAs conditional on genetic influences shared with EA, we applied MVR. This technique is analogous to MR methodologies[26] and controls for collider bias[42] through the use of GWAS summary statistics. Technically, it involves the regression of regression estimates from independent samples on each other[26] (Supplementary Information). Within this study we use MVR without inferring causality due to violations of classical MR assumptions[26] (see below).

Genetic variant selection: To disentangle ADHD-LRA associations, we selected two sets of instruments from the most powerful ADHD GWAS summary statistics (PGC+iPSYCH). The first set contained genome-wide significant variants ($P<5x10^{-8}$, conservative). The second set included variants passing a more lenient *P*-value threshold ($P<0.0015$, subthreshold) to increase power, consistent with current guidelines for the selection of genetic instruments in MR (F-statistic<10)[43]. All sets included independent (PLINK[44] clumping: LD-$r^2<0.25$, ±500 kb), well imputed (INFO[45]>0.8) and common (EAF>0.01) variants. This resulted in 15 conservative and $2,689<N_{SNPs}\leq2,692$ subthreshold ADHD-instruments (Table S8).

Estimation of ADHD effects: We extracted regression estimates for selected ADHD-instruments (conservative and subthreshold) from ADHD (PGC+iPSYCH), EA (SSGAC) and 13 LRA (ALSPAC) GWAS summary statistics. Analysing each set of variants independently, regression estimates for individual LRA measures (β) were regressed on both ADHD (β as lnOR) and EA regression estimates (β) using an OLS regression framework (R:stats library, Rv3.2.0). Outcomes were 1) a MVR regression estimate quantifying the change in SD units of LRA performance per log odds increase in ADHD risk conditional on years of schooling (ADHD effect independent of EA), and 2) a MVR regression estimate quantifying the change in SD units of LRA performance per year of schooling as captured by ADHD instruments (ADHD effect shared with EA). Latter MVR regression estimates capture here shared genetic effects between ADHD, EA and LRAs, including 1) genetic confounding (i.e. genetically predictable EA influences both ADHD and LRAs), 2) mediation (i.e. genetically predictable ADHD influences LRA indirectly through EA) and 3) biological pleiotropy (i.e. ADHD risk variants affect ADHD and EA through independent biological pathways). As ADHD risk and EA are inversely genetically related with each other[21], they were reported to quantify change per missing year of schooling. To compare the magnitude of both MVR estimates, we also conducted analyses using fully standardised EA, ADHD and LRA regression estimates (Supplementary Information).

Finally, MVR regression estimates were meta-analysed and contrasted across reading-related, spelling-related and all LRA measures (excluding the composite measure verbal intelligence) (Table 1) using random-effects meta-regression, accounting for phenotypic correlations between LRAs (R:metafor library[46], Rv3.2.0; Supplementary Information).

Sensitivity analyses: As the directionality of the relationship between ADHD, EA and LRAs cannot be inferred in this study, we also examined the genetic association between EA and LRAs, conditional on ADHD, using MVR. Two sets of EA instruments (conservative and subthreshold, Table S8) were selected from EA (SSGAC) GWAS summary statistics, analogous to the selection of ADHD instruments, and MVR was conducted as described above. Note that we did not create LRA instrument sets, as GWAS summary statistics of LRAs were underpowered.

Attrition analysis

We carried out an *attrition* analysis in ALSPAC studying the genetic association between LRA-missingness and polygenic ADHD risk, using both polygenic scoring analyses and MVR (Supplementary Information).

6.3. Results

Genetic architecture of literacy- and language-related abilities and clinical ADHD

Phenotypic variation in literacy- and language-related measures (Table 1), including reading abilities (comprehension, accuracy and speed) assessed in words/passages and non-words, spelling abilities (accuracy), phonemic awareness, listening comprehension, non-word repetition and verbal intelligence scores, can be tagged by common variants, with SNP-h^2 estimates between 0.32 (SE=0.07, non-word repetition age 8) and 0.54 (SE=0.07, verbal intelligence age 8) (Table S2; GCTA- and LDSC-based estimations). Importantly, all LRAs are phenotypically (Table S1) and genetically (Table S3) moderately to strongly interrelated. The observed LDSC-based evidence for genetic liability of clinical ADHD within the PGC (LDSC-h^2=0.08(SE=0.03)), iPSYCH (LDSC-h^2=0.26(SE=0.02)) and PGC+iPSYCH samples (Table S4) is consistent with previous reports [21].

Association between ADHD polygenic risk scores and literacy- and language-related abilities

We observed robust evidence for an inverse genetic association between ADHD-PGS and reading accuracy/comprehension age 7 (PGC: OLS-R^2=0.1%, P=4.6x10^{-3}; iPSYCH: OLS-R^2=1.0%, P<1x10^{-10}), reading accuracy age 9 (PGC: OLS-R^2=0.1%, P=5.7x10^{-3}; iPSYCH: OLS-R^2=1.2%, P<1x10^{-10}), and spelling accuracy age 9 (PGC: OLS-R^2=0.2%, P=1.5x10^{-3}; iPSYCH: OLS-R^2=0.8%, P<1x10^{-10}) using independent ADHD discovery samples (Figure 1, Table S5). The strongest evidence for association was observed when ADHD discovery samples were combined (PGC+iPSYCH; Figure 1), including those of European ancestry only (PGC+iPSYCH(EUR)), with genetic trait-disorder overlap present for all LRAs studied (Table S5). For example, ADHD-PGS explain 1.49% phenotypic variation in reading accuracy age 9, translating into a genetic covariance of -0.11(95%-CI: -0.14;-0.09) (Supplementary Information). Polygenic scoring results are presented for a P-value threshold of 0.1, but other thresholds provided similar results (data not shown). Results were not affected by the exclusion of children with an ADHD diagnosis in ALSPAC (Table S6).

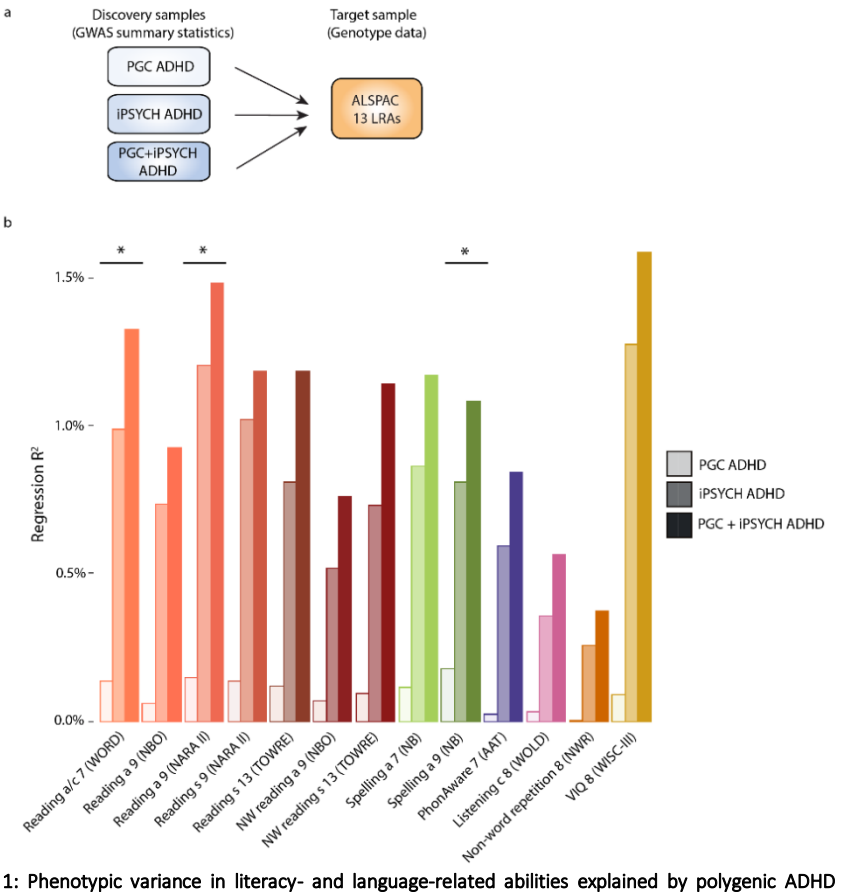


Figure 1: Phenotypic variance in literacy- and language-related abilities explained by polygenic ADHD risk. **(a)** Schematic representation of polygenic scoring analyses. ADHD polygenic scores were created in ALSPAC using PGC, iPSYCH and PGC+iPSYCH GWAS summary statistics. Rank-transformed LRAs were regressed on Z-standardised ADHD-PGS using ordinary least square regression. **(b)** Phenotypic variance in literacy- and language-related abilities explained by polygenic ADHD risk. * Evidence for association between LRAs and polygenic ADHD risk as observed in PGC ADHD, iPSYCH ADHD and PGC+iPSYCH ADHD samples. Note that all LRAs were associated with polygenic ADHD risk in iPSYCH ADHD and PGC+iPSYCH ADHD passing the experiment-wide error rate (P<0.007). Abbreviations: a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; PGC, Psychiatric Genomics Consortium; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; ADHD, Attention-Deficit/Hyperactivity Disorder.

Shared genetic liability between ADHD and LRA with EA

There was strong evidence for a moderate negative genetic correlation (r_g=-0.53(SE=0.03), P<1x10^{-10}) between genetically predicted ADHD, as captured by the largest ADHD discovery sample (PGC+iPSYCH), and EA (LDSC-h^2=0.11(SE=0.004)), consistent with previous findings[21]. Likewise, LRAs were moderately to highly positively correlated with EA (e.g. reading speed age 13 r_g=0.80(SE=0.22), P=3.0x10^{-4}; Table S7), as previously reported[22,23]. Additionally, two independent variants reached genome-wide significance for both ADHD[21] and EA[20], consistent with biological pleiotropy (i.e. single genetic loci influencing multiple traits)[47]. These findings indicate complex, potentially reciprocal cross-trait relationships (Figure 2a) and violate MR causal modelling assumptions[26]. Consequently, ADHD instruments are not valid MR instruments as they are not independent of EA.

Multivariable regression analyses

To disentangle the genetic overlap of polygenic ADHD risk with literacy- and language-related measures into ADHD genetic effects independent of and shared with EA, we applied MVR[26] using ADHD instruments based on the most powerful ADHD discovery sample (PGC+iPSYCH) (Figure 2b).

Using conservative ADHD instruments (Table S8), non-word reading accuracy at age 9 and pooled reading-related abilities were associated with polygenic ADHD risk, conditional on EA (Table 3). The latter translates into, for example, a decrease of 0.35 SD in pooled reading performance per log-odds increase in ADHD risk (β_{ADHD}=-0.35(SE=0.09), P=9.2x10^{-5}, P_{het}=0.19), an effect that was considerably stronger than for other LRAs (P_{mod}=0.011, Table S10).

Using subthreshold ADHD instruments (Table S8), polygenic ADHD effects on LRA performance, conditional on EA, were detectable for all reading- and spelling-related measures, phonemic awareness and verbal intelligence, but not other LRAs such as listening comprehension and non-word repetition (Table 3). Evidence was strongest for pooled reading and spelling abilities (Table 3, minimum P=1.1x10^{-8}). However, observable effects were smaller in magnitude compared to those captured by conservative ADHD instruments with, for example, a 0.03 SD decrease in pooled reading performance per log-odds increase in ADHD risk (β_{ADHD}=-0.03(SE=0.01), P=1.4x10^{-6}, Table 3). Comparing ADHD-specific effects on both reading and spelling with ADHD-specific effects on all other LRAs provided evidence for effect differences (P_{mod}= 0.016), with stronger ADHD effects on literacy-related abilities, in particular spelling (Table S10).

Polygenic ADHD effects that are shared with EA were identified for all LRAs studied using subthreshold, but not conservative ADHD instruments (Table 3). This translates into, for example, a further 0.50 SD units decrease in pooled reading performance per missing school year (β_{EA}=-0.50(SE=0.09), P=4.9x10^{-8}, Table 3). Thus, the

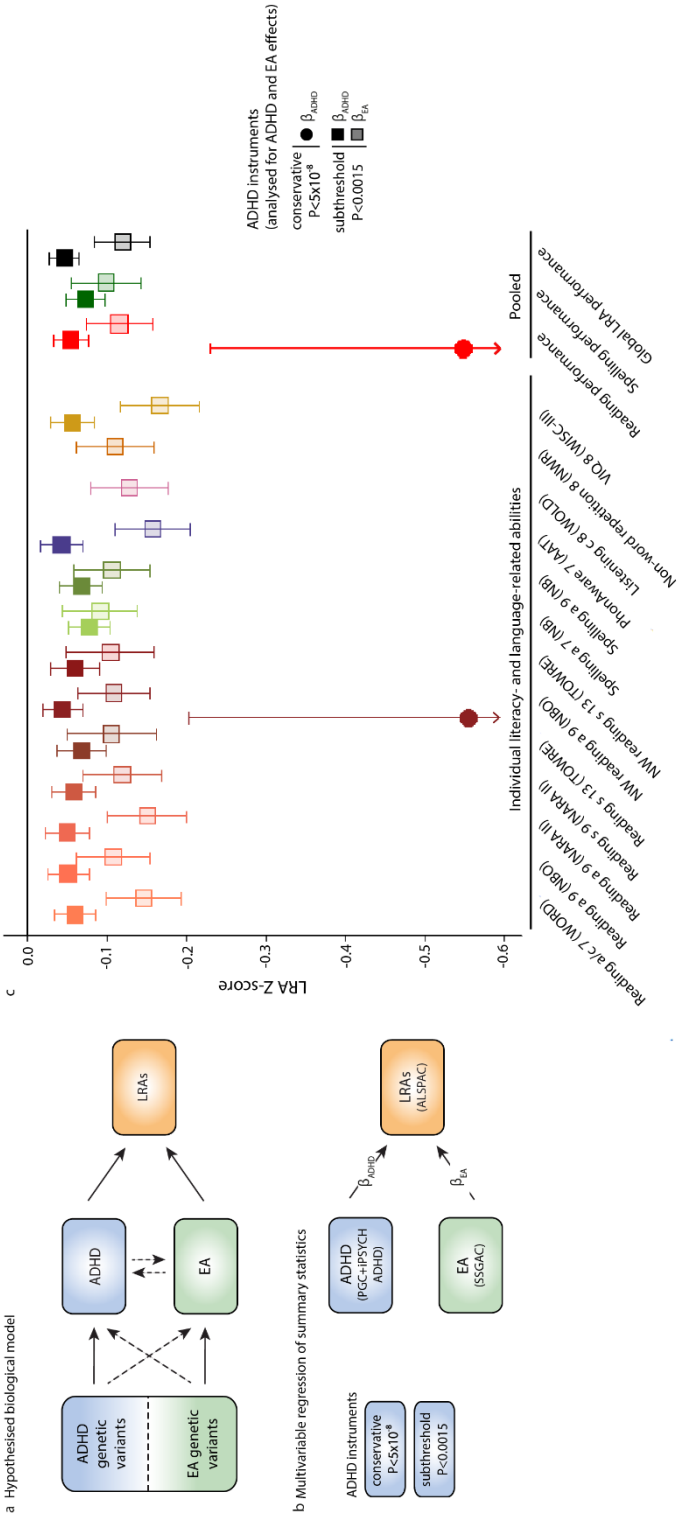

a Hypothesised biological model
ADHD genetic variants
EA genetic variants
ADHD
EA
LRAs
b Multivariable regression of summary statistics
ADHD instruments
conservative P<5x10⁻⁸
subthreshold P<0.0015
ADHD (PGC+iPSYCH ADHD)
EA (SSGAC)
LRAs (ALSPAC)
c
LRA Z-score
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
ADHD instruments (analysed for ADHD and EA effects)
conservative P<5x10⁻⁸
subthreshold P<0.0015
Individual literacy- and language-related abilities
Pooled
Reading a/c 7 (WORD)
Reading a 9 (NBO)
Reading a 9 (NARA II)
Reading s 9 (NARA II)
Reading s 13 (TOWRE)
NW reading a 9 (NBO)
NW reading s 13 (TOWRE)
Spelling a 7 (NB)
Spelling a 9 (NB)
PhonAware 7 (AAT)
Listening C 8 (WOLD)
Non-word repetition 8 (NWR)
VIQ 8 (WISC-III)
Reading performance
Spelling performance
Global LRA performance

Figure 2: Genetic relationships between ADHD, educational attainment and literacy- and language-related abilities. (a) Hypothesised biological model of genetic relationships between ADHD, EA, and LRAs reflecting complex, pleiotropic and reciprocal genetic links that prevent causal inferences. **(b)** Schematic MVR model assessing polygenic ADHD-LRA overlap independent of and shared with genetic effects for EA, **(c)** MVR estimates of ADHD-specific effects independent of EA and ADHD effects shared with EA on LRAs using standardised ADHD instruments: Sets of conservative ($P<5x10^{-8}$) and subthreshold ($P<0.0015$) ADHD instruments were extracted from ADHD (PGC+iPSYCH), EA (SSGAC) and LRAs (ALSPAC) GWAS summary statistics. ADHD-specific effects independent of EA (β_{ADHD}) and ADHD effects shared with EA (β_{EA}) on LRAs were estimated with MVRs. To compare the magnitude of β_{ADHD} and β_{EA}, MVR analyses were conducted using standardised regression estimates (Supplementary Methods). β_{ADHD} estimates measure the change in LRA Z-score per Z-score in ADHD liability. β_{EA} estimates measure the change in LRA Z-scores per Z-score in missing school years. MVR estimates based on raw genetic effect estimates are provided in Table 3. Pooled estimates for reading, spelling and global LRA measures (Table 1) were obtained through random-effects meta-regression. Only effects passing the experiment-wide significance threshold ($P<0.007$) are shown with corresponding 95% confidence intervals. There is no causality inferred. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; LRAs, literacy and language-related abilities; PGC, Psychiatric Genomics Consortium; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; SSGAC, Science Genetic Association Consortium; ALSPAC, Avon Longitudinal Study of Parents And Children; MVR, multivariable regression.

Table 3: Multivariable regression analysis of polygenic associations between ADHD and literacy- and language-related abilities (raw estimates)

LRAs	ADHD (β_{ADHD}) (ADHD-specific effects independent of EA)						EA (β_{EA}) (EA genetic effects of ADHD-associated variants)[1]					
	Conservative instruments (P_{thr}<5x10^{-8})			Subthreshold instruments (P_{thr}<0.0015)			Conservative instruments (P_{thr}<5x10^{-8})			Subthreshold instruments (P_{thr}<0.0015)		
	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}
Reading a/c 7 (WORD)	-0.18 (0.13)	0.20	-	-0.03 (0.01)	1.4x10^{-5}	-	-0.09 (1.19)	0.94	-	-0.63(0.10)	1.2x10^{-9}	-
Reading a 9 (NBO)	-0.28 (0.09)	0.01	-	-0.03 (0.01)	2.1x10^{-4}	-	1.33 (0.84)	0.14	-	-0.46 (0.10)	6.3x10^{-6}	-
Reading s 9 (NARA II)	-0.29 (0.13)	0.04	-	-0.03 (0.01)	3.0x10^{-5}	-	1.55 (1.16)	0.21	-	-0.51 (0.11)	2.5x10^{-6}	-
Reading a 9 (NARA II)	-0.34 (0.12)	0.01	-	-0.03 (0.01)	3.8x10^{-4}	-	1.57 (1.11)	0.18	-	-0.64 (0.11)	4.1x10^{-9}	-
Reading s 13 (TOWRE)	-0.36 (0.15)	0.03	-	-0.04 (0.01)	2.7x10^{-5}	-	2.06 (1.35)	0.15	-	-0.46 (0.12)	1.8x10^{-4}	-
NW reading a 9 (NBO)	-0.37 (0.09)	0.002	-	-0.03 (0.01)	2.4x10^{-4}	-	1.54 (0.85)	0.09	-	-0.46 (0.10)	6.1x10^{-6}	-
NW reading s 13 (TOWRE)	-0.21 (0.17)	0.25	-	-0.03 (0.01)	2.4x10^{-4}	-	1.43 (1.58)	0.38	-	-0.45 (0.12)	1.9x10^{-4}	-
Spelling a 7 (NB)	-0.05 (0.12)	0.72	-	-0.05 (0.01)	1.4x10^{-5}	-	-1.12 (1.13)	0.34	-	-0.63 (0.10)	1.2x10^{-9}	-
Spelling a 9 (NB)	-0.23 (0.08)	0.01	-	-0.04 (0.01)	8.7x10^{-7}	-	1.43 (0.72)	0.38	-	-0.45 (0.10)	1.3x10^{-5}	-
PhonAware 7 (AAT)	-0.22 (0.15)	0.16	-	-0.02 (0.01)	0.002	-	-0.31 (1.33)	0.82	-	-0.67 (0.10)	<1x10^{-10}	-
Listening c 8 (WOLD)	-0.03 (0.08)	0.75	-	-0.02 (0.01)	0.02	-	-1.47 (0.70)	0.06	-	-0.55 (0.11)	4.5x10^{-7}	-
Non-word repetition 8 (CNRep)	-0.14 (0.14)	0.34	-	-0.02 (0.01)	0.01	-	-0.16 (1.30)	0.90	-	-0.47 (0.11)	1.5x10^{-5}	-
VIQ 8 (WISC-III)	-0.23 (0.15)	0.16	-	-0.03 (0.01)	5.0x10^{-5}	-	-0.36 (1.37)	0.80	-	-0.71 (0.11)	<1x10^{-10}	-
Pooled reading	-0.35 (0.09)	9.2x10^{-5}	0.19	-0.03 (0.01)	1.4x10^{-6}	0.79	1.67 (0.78)	0.03	0.31	-0.50 (0.09)	4.9x10^{-8}	0.09
Pooled spelling	-0.18 (0.11)	0.10	0.03	-0.04 (0.01)	1.1x10^{-8}	0.28	0.40 (1.25)	0.75	0.001	-0.42 (0.10)	1.3x10^{-5}	0.39
Pooled LRAs	-0.18 (0.07)	0.01	0.005	-0.03 (0.01)	1.9x10^{-6}	0.05	0.29 (0.69)	0.67	0.001	-0.49 (0.08)	<1x10^{-10}	0.004

Sets of conservative ($P<5x10^{-8}$) and subthreshold ($P<0.0015$) ADHD instruments were extracted from ADHD (PGC+iPSYCH), EA (SSGAC) and LRAs (ALSPAC) GWAS summary statistics. ADHD-specific effects independent of EA (β_{ADHD}) and ADHD effects shared with EA (β_{EA}) on LRAs were estimated with MVRs (Figure 2b). ADHD effects shared with EA were assessed through EA genetic effect estimates of ADHD-associated variants and presented with respect to missing school years. β_{ADHD} quantifies the change in LRA Z-score per log odds increase in ADHD liability. β_{EA} quantifies the change in LRA Z-score per missing year of schooling. Pooled estimates for reading, spelling and global LRAs (Table 1) were obtained through random-effects meta-regression. Evidence for effect heterogeneity (P_{het}) was monitored through Cochran's Q-test. Effects passing the experiment-wide significance threshold ($P<0.007$) are depicted in grey. To compare effect sizes of β_{ADHD} and β_{EA}, MVR was carried out using standardised genetic effect estimates for which results are provided in Figure 2c. 1. ADHD genetic effects shared with EA as assessed through EA genetic effect estimates of ADHD-associated variants. Abbreviations: LRAs, literacy- and language-related abilities; ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; P_{thr}, *P*-value threshold; P_{het}, heterogeneity *P*-value; a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness, AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; MVR, Multivariable regression.

observed association between polygenic ADHD risk and listening comprehension and non-word repetition is fully attributable to genetic effects shared with EA (Table 3). Contrary to ADHD-specific effects, ADHD effects shared with EA showed no evidence for effect differences between literacy-related versus other LRAs (*P*=0.31). Conducting MVR with fully standardised estimates showed that ADHD effects shared with EA were as large as or even larger compared to ADHD-specific effects (Figure 2c, Table S9).

Using an analogous approach, we disentangled the genetic overlap between polygenic EA and LRAs into genetic EA effects independent of and shared with ADHD, based on EA instruments (Figure S1). There was strong evidence for EA effects shared with ADHD using subthreshold, but not conservative EA instruments (Table S11). The magnitude of ADHD genetic effects shared with EA, captured by ADHD genetic instruments, compared to the magnitude of EA genetic effects shared with ADHD, captured by EA instruments, was largely consistent with each other in fully standardised analyses (Table S9 and S11).

There was little evidence supporting the inclusion of regression intercepts in MVR that would imply additional genetic effect variation in LRAs estimates, not yet captured by either ADHD or EA effect estimates, based on the selected instruments. Therefore, all MVRs were performed using constrained intercepts[26].

6.4. Discussion

This study identified strong and replicated evidence for an inverse association between polygenic ADHD risk and multiple population-based LRAs using a polygenic scoring approach. However, these associations involve shared genetic variation with genetically predictable EA. Accurate modelling of polygenic links using MVR techniques, conditional on EA, revealed an ADHD-specific association profile that primarily involves literacy-related impairments. Once shared genetic effects with EA were accounted for, polygenic ADHD risk was most strongly inversely associated with reading and/or spelling abilities, in addition to phonemic awareness and verbal intelligence, but not listening comprehension and non-word repetition abilities. Importantly, genetic overlap between polygenic ADHD risk and all of the LRAs studied was inflated by genetic effects shared with EA.

Using independent ADHD discovery samples, these findings show that genetic overlap between ADHD and literacy-related impairments observed in twin and family studies[16–18] can be extended to genetic associations, as captured by common variation in general population samples. The identified association profile suggests that not only reading-related abilities (including both word and nonword reading skills), but also phonological and spelling-related abilities share genetic aetiology with ADHD. These interrelated LRAs may, as hypothesised for RD, arise from a phonological

impairment[48,49], which affects decoding and reading skills[50], but also spelling abilities[51]. However, reading abilities can, once developed, also shape phonological skills[52].

In addition, this study suggests that genetic associations between polygenic ADHD and LRAs reflect, at least partially, shared genetic influences with genetically predictable EA and that, equally likely, genetic associations between polygenic EA and LRAs share genetic influences with ADHD. The magnitude of these shared effects, modelled with different MVR approaches, was comparable with each other. This is consistent with reciprocal genetic influences between EA and ADHD (Figure 2a) and supports an intergenerational multiple-deficit model proposed for reading disability[15,53]. Children growing up in disadvantaged environments, genetically predictable through polygenic EA scores[54], might be more vulnerable to psychiatric illness including ADHD[55] that affects, in turn, their LRAs. In addition, adolescents with ADHD might be more likely to leave school at an earlier age, with lower LRA performance and EA, and pass on an increased genetic load to their own children[56].

Here, we demonstrate that disentangling multivariate genetic interrelationships between ADHD, EA and LRAs using MVR can aid the interpretation of genetic overlap, while controlling for collider bias[42]. However, using MVR, the detection of these polygenic associations was strongly governed by the choice of genetic variants. Conservative ADHD instruments identified large ADHD-specific effects on reading as a domain and little evidence for genetic effects that are shared with EA, although they had limited power[57]. They comprised 15 independent SNPs only, including variation within *FOXP2,* a gene that has been implicated in childhood apraxia of speech and expressive and receptive language impairments (http://omim.org/entry/602081)[58]. On the other hand, subthreshold instruments, including thousands of variants, tagged ADHD-specific polygenic links with LRAs (conditional on EA) with smaller effects, but with higher predictive accuracy. However, these instruments also captured shared genetic effects with EA, affecting polygenic links between ADHD and all of the LRAs studied. These shared genetic influences were of equal strength and at least equal magnitude compared to ADHD-LRA associations independent of EA. Contrary, a previous twin study showed that the genetic covariance between ADHD and reading difficulties was largely independent of genetic effects shared with IQ[19], suggesting that our findings may also reflect socio-economic influences. Thus, in order to improve reading and, more generally, literacy-related deficits in children with ADHD, there is potentially a need for further intervention programmes targeting EA-independent underlying neurocognitive deficits, beyond general training programmes aiming at schooling outcomes[59].

In general, our findings are consistent with an omnigenic[60] model of complex trait architectures, compatible with a general factor model of psychopathology[61], including ADHD[62]. The omnigenic model construes that only the largest-effect variants will reflect ADHD specificity, and may thus tag the most trait-specific associations between ADHD and reading, independent of EA. The majority of variants, however, will capture

pleiotropic (omnigenic) influences pointing to highly interconnected neural networks[60] that give rise to genetic confounding. Consequently, the majority of subthreshold variants, captured by both ADHD and EA subthreshold variants, are likely to represent highly powerful cross-trait genetic predictors that may enhance and induce genetic overlap.

Finally, the methodological framework within this work has not only relevance for studies investigating polygenic links between ADHD and LRAs, but for many studies examining multivariate trait interrelationships that involve shared genetic effects with a genetically predictable confounder. Specifically, our findings suggest that lower variant selection thresholds can introduce genetic variance sharing that is unspecific and needs to be accounted for before identified associations can be interpreted in terms of underlying mechanisms, including shared genetic aetiologies. This is especially important as current guidelines for studying polygenic links with allelic scores recommend aggregating genetic variants across less stringent significance thresholds to maximise genetic association between discovery and target samples[63,64].

This study has several limitations. Firstly, ALSPAC, as other cohort studies, suffers from attrition[65,66]. Sensitivity analyses showed that this is unlikely to bias our findings based on conservative instruments. However, links identified using subthreshold ADHD variants, might have been underestimated given that individuals with a genetic predisposition to ADHD (but also smoking initiation, higher body mass index, neuroticism, schizophrenia and depression) are more likely to drop out[66]. Secondly, the strength of the genetic overlap between polygenic ADHD risk and LRAs may vary according to ADHD symptom domain levels, implicating especially inattentiveness[67], as well as the nature of the literacy- or language-related ability involved (as we observed evidence for effect heterogeneity when combining all LRAs). It is conceivable that also other verbal abilities, not investigated in this study, such as grammar, expressive vocabulary or pragmatic skills, may genetically overlap with ADHD. Furthermore, we only studied the extent to which shared genetic variance with EA affects the genetic association between ADHD and LRAs. However, we found little evidence for the presence of additional unaccounted for genetic influences using these instruments, i.e. effects that are not yet captured by either genetically predicted ADHD or EA. Finally, the power of available LRA GWAS summary statistics is still too low to generate genetic instruments supporting reverse models. Larger and more detailed clinical and population-based samples, as well as extensive multivariate variance analyses of the spectrum of LRAs (that are currently computationally expensive[68]) will help to further characterise the overlap between ADHD and literacy- and language-related cognitive processes.

Conclusion

Polygenic associations of clinical ADHD and a range of LRAs are to a large extent attributable to genetic effects that are also shared with EA, especially when investigated with genetic variants typically selected for polygenic scoring approaches. Adjusting for these unspecific genetic effects reveals an ADHD-specific association profile that primarily involves literacy-related impairments.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and the Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. The ALSPAC GWAS data was generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. This research was facilitated by the Social Science Genetic Association Consortium (SSGAC) and Psychiatric Genomics Consortium (PGC) by providing access to genome-wide summary statistics. This publication is the work of the authors and EV and BSTP will serve as guarantors for the contents of this paper. EV, CYS, BSTP and SEF are supported by the Max Planck Society. The iPSYCH project is funded by the Lundbeck Foundation (grant no R102-A9118 and R155-2014-1724) and the universities and university hospitals of Aarhus and Copenhagen. Genotyping of iPSYCH and PGC samples was supported by grants from the Lundbeck Foundation, the Stanley Foundation, the Simons Foundation (SFARI 311789 to Mark J Daly), and NIMH (5U01MH094432-02 to Mark J Daly). The Danish National Biobank resource was supported by the Novo Nordisk Foundation. Data handling and analysis was supported by NIMH (1U01MH109514-01 to Michael O'Donovan and Anders D Børglum). High-performance computer capacity for handling and statistical analysis of iPSYCH data on the GenomeDK HPC facility was provided by the Centre for Integrative Sequencing, iSEQ, Aarhus University, Denmark (grant to Anders D Børglum).

References

1. Geurts, H. M. & Embrechts, M. Language profiles in ASD, SLI, and ADHD. *J. Autism Dev. Disord.* **38**, 1931–1943 (2008).
2. Helland, W. A., Posserud, M.-B., Helland, T., Heimann, M. & Lundervold, A. J. Language Impairments in Children With ADHD and in Children With Reading Disorder. *J. Atten. Disord.* 1087054712461530 (2012) doi:10.1177/1087054712461530.
3. Germanò, E., Gagliano, A. & Curatolo, P. Comorbidity of ADHD and dyslexia. *Dev. Neuropsychol.* **35**, 475–493 (2010).
4. Maughan, B. & Carroll, J. Literacy and mental disorders. *Curr. Opin. Psychiatry* **19**, 350–354 (2006).
5. Mayes, S. D. & Calhoun, S. L. Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child Neuropsychol. J. Norm. Abnorm. Dev. Child. Adolesc.* **13**, 469–493 (2007).
6. Mayes, S. D. & Calhoun, S. L. Frequency of reading, math, and writing disabilities in children with clinical disorders. *Learn. Individ. Differ.* **16**, 145–157 (2006).
7. Czamara, D. *et al.* Children with ADHD Symptoms Have a Higher Risk for Reading, Spelling and Math Difficulties in the GINIplus and LISAplus Cohort Studies. *PLOS ONE* **8**, e63859 (2013).
8. Åsberg, J., Kopp, S., Berg-Kelly, K. & Gillberg, C. Reading comprehension, word decoding and spelling in girls with autism spectrum disorders (ASD) or attention-deficit/hyperactivity disorder (AD/HD): performance and predictors. *Int. J. Lang. Commun. Disord.* **45**, 61–71 (2010).
9. Green, B. C., Johnson, K. A. & Bretherton, L. Pragmatic language difficulties in children with hyperactivity and attention problems: an integrated review. *Int. J. Lang. Commun. Disord.* **49**, 15–29 (2014).
10. Hawkins, E., Gathercole, S., Astle, D. & Holmes, J. Language Problems and ADHD Symptoms: How Specific Are the Links? *Brain Sci.* **6**, (2016).
11. Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J. & Rohde, L. A. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am. J. Psychiatry* **164**, 942–948 (2007).
12. Peterson, R. L. & Pennington, B. F. Developmental dyslexia. *Lancet Lond. Engl.* **379**, 1997–2007 (2012).
13. Thapar, A. & Cooper, M. Attention deficit hyperactivity disorder. *The Lancet* **387**, 1240–1250 (2016).
14. Gough, P. B. & Tunmer, W. E. Decoding, Reading, and Reading Disability. *Remedial Spec. Educ.* **7**, 6–10 (1986).
15. Pennington, B. F. From single to multiple deficit models of developmental disorders. *Cognition* **101**, 385–413 (2006).
16. Martin, N. C., Levy, F., Pieka, J. & Hay, D. A. A Genetic Study of Attention Deficit Hyperactivity Disorder, Conduct Disorder, Oppositional Defiant Disorder and Reading Disability: Aetiological overlaps and implications. *Int. J. Disabil. Dev. Educ.* **53**, 21–34 (2006).
17. Willcutt, E. G., Pennington, B. F. & DeFries, J. C. Twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **96**, 293–301 (2000).
18. Willcutt, E. G., Pennington, B. F., Olson, R. K. & DeFries, J. C. Understanding comorbidity: A twin study of reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **144B**, 709–714 (2007).
19. Paloyelis, Y., Rijsdijk, F., Wood, A. C., Asherson, P. & Kuntsi, J. The Genetic Association Between ADHD Symptoms and Reading Difficulties: The Role of Inattentiveness and IQ. *J. Abnorm. Child Psychol.* **38**, 1083–1095 (2010).
20. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–42 (2016).
21. Demontis, D. *et al.* Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv* 145581 (2017) doi:10.1101/145581.
22. Selzam, S. *et al.* Genome-Wide Polygenic Scores Predict Reading Performance Throughout the School Years. *Sci. Stud. Read.* 1–16 (2017) doi:10.1080/10888438.2017.1299152.
23. Luciano, M. *et al.* Single Nucleotide Polymorphisms Associated with Reading Ability Show Connection to Socio-Economic Outcomes. *Behav. Genet.* **47**, 469–479 (2017).
24. Krapohl, E. *et al.* Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proc. Natl. Acad. Sci.* 201707178 (2017) doi:10.1073/pnas.1707178114.

25. Krapohl, E. *et al.* Multi-polygenic score approach to trait prediction. *Mol. Psychiatry* **23**, 1368–1374 (2018).
26. Burgess, S. *et al.* Dissecting Causal Pathways Using Mendelian Randomization with Summarized Genetic Data: Application to Age at Menarche and Risk of Breast Cancer. *Genetics* genetics.300191.2017 (2017) doi:10.1534/genetics.117.300191.
27. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
28. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
29. Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* **95**, 221–227 (2005).
30. Goodman, R., Ford, T., Richards, H., Gatward, R. & Meltzer, H. The Development and Well-Being Assessment: Description and Initial Validation of an Integrated Assessment of Child and Adolescent Psychopathology. *J. Child Psychol. Psychiatry* **41**, 645–655 (2000).
31. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45**, 984–994 (2013).
32. Mors, O., Perto, G. P. & Mortensen, P. B. The Danish Psychiatric Central Research Register. *Scand. J. Public Health* **39**, 54–57 (2011).
33. Pedersen, C. B. *et al.* The iPSYCH2012 case–cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol. Psychiatry* **23**, 6–14 (2018).
34. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
35. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinforma. Oxf. Engl.* **28**, 2540–2542 (2012).
36. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
37. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
38. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
39. International Schizophrenia Consortium *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
40. Dudbridge, F. Power and Predictive Accuracy of Polygenic Risk Scores. *PLOS Genet* **9**, e1003348 (2013).
41. Stergiakouli, E. *et al.* Shared genetic influences between dimensional ASD and ADHD symptoms during child and adolescent development. *Mol. Autism* **8**, 18 (2017).
42. Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am. J. Hum. Genet.* **96**, 329–339 (2015).
43. Pierce, B. L., Ahsan, H. & VanderWeele, T. J. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int. J. Epidemiol.* **40**, 740–752 (2011).
44. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
45. Howie, B. N., Donnelly, P. & Marchini, J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLOS Genet.* **5**, e1000529 (2009).
46. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 1–48 (2010).
47. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* **23**, R89–R98 (2014).
48. Carroll, J. M. & Snowling, M. J. Language and phonological skills in children at high risk of reading difficulties. *J. Child Psychol. Psychiatry* **45**, 631–640 (2004).
49. Hulme, C. & Snowling, M. J. *Developmental Disorders of Language Learning and Cognition*. (Wiley-Blackwell, 2009).
50. Rack, J., Snowling, M. & Olson, R. The Nonword Reading Deficit in Developmental Dyslexia - a Review. *Read. Res. Q.* **27**, 28–53 (1992).

51. Treiman, R. *Beginning to Spell: A Study of First-Grade Children*. (Oxford University Press, 1992).
52. Kamhi, A. G. & Catts, H. W. *Language and Reading Disabilities*. (Pearson, 2011).
53. van Bergen, E., van der Leij, A. & de Jong, P. F. The intergenerational multiple deficit model and the case of dyslexia. *Front. Hum. Neurosci.* **8**, 346 (2014).
54. Ayorech, Z., Krapohl, E., Plomin, R. & Stumm, S. von. Genetic Influence on Intergenerational Educational Attainment. *Psychol. Sci.* **28**, 1302–1310 (2017).
55. Reiss, F. Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. *Soc. Sci. Med. 1982* **90**, 24–31 (2013).
56. Russell, G., Ford, T., Rosenberg, R. & Kelly, S. The association of attention deficit hyperactivity disorder with socioeconomic disadvantage: alternative explanations and evidence. *J. Child Psychol. Psychiatry* **55**, 436–445 (2014).
57. Schatzkin, A. *et al.* Mendelian Randomization: How It Can—and Cannot—Help Confirm Causal Relations between Nutrition and Cancer. *Cancer Prev. Res. (Phila. Pa.)* **2**, 104–113 (2009).
58. Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519–23 (2001).
59. Daley, D. & Birchwood, J. ADHD and academic performance: why does ADHD impact on academic performance and what can be done to support ADHD children in the classroom? *Child Care Health Dev.* **36**, 455–464 (2010).
60. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177–1186 (2017).
61. Lahey, B. B. *et al.* Is There a General Factor of Prevalent Psychopathology during Adulthood? *J. Abnorm. Psychol.* **121**, 971–977 (2012).
62. Franke, B. What's in a Name: the 'Omnigenic' Model as a Point of Departure for Polygenic Psychiatric Disorders. *JPBS* (2017) doi:10.20900/jpbs.20170014(S1-S8).
63. Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466–1468 (2015).
64. Wray, N. R. *et al.* Research Review: Polygenic methods and their application to psychiatric traits. *J. Child Psychol. Psychiatry* **55**, 1068–1087 (2014).
65. Martin, J. *et al.* Association of Genetic Risk for Schizophrenia With Nonparticipation Over Time in a Population-Based Cohort Study. *Am. J. Epidemiol.* **183**, 1149–1158 (2016).
66. Taylor, A. *et al.* The molecular genetics of participation in the Avon Longitudinal Study of Parents and Children. *bioRxiv* 206698 (2017) doi:10.1101/206698.
67. Greven, C. U., Harlaar, N., Dale, P. S. & Plomin, R. Genetic Overlap between ADHD Symptoms and Reading is largely Driven by Inattentiveness rather than Hyperactivity-Impulsivity. *J. Can. Acad. Child Adolesc. Psychiatry* **20**, 6–14 (2011).
68. St Pourcain, B. *et al.* Developmental changes within the genetic architecture of social communication behaviour: A multivariate study of genetic variance in unrelated individuals. *Biol. Psychiatry* **83**, 598–606 (2017).

Supplementary Materials

Supplementary methods & analyses

ALSPAC description

ALSPAC recruited 14541 pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992. 14,541 is the initial number of pregnancies for which the mother enrolled in the ALSPAC study and had either returned at least one questionnaire or attended a "Children in Focus" clinic by 19/07/99. Of these initial pregnancies, there were a total of 14,676 fetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age.

When the oldest children were approximately 7 years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join the study originally. As a result, when considering variables collected from the age of seven onwards (and potentially abstracted from obstetric notes) there are data available for more than the 14,541 pregnancies mentioned above.

The number of new pregnancies not in the initial sample (known as Phase I enrolment) that are currently represented on the built files and reflecting enrolment status at the age of 18 is 706 (452 and 254 recruited during Phases II and III respectively), resulting in an additional 713 children being enrolled. The phases of enrolment are described in more detail in the cohort profile paper[1].

The total sample size for analyses using any data collected after the age of seven is therefore 15,247 pregnancies, resulting in 15,458 fetuses. Of this total sample of 15,458 fetuses, 14,775 were livebirths and 14,701 were alive at 1 year of age.

A 10% sample of the ALSPAC cohort, known as the Children in Focus (CiF) group, attended clinics at the University of Bristol at various time intervals between 4 to 61 months of age. The CiF group were chosen at random from the last 6 months of ALSPAC births (1,432 families attended at least one clinic). Excluded were those mothers who had moved out of the area or were lost to follow-up, and those partaking in another study of infant development in Avon.

The study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/).

Literacy- and language-related measures in ALSPAC

Reading accuracy and comprehension age 7 (WORD)

Pictures and words were used to assess decoding and word reading. The child was shown a series of four pictures. Each picture had four short, simple words underneath it. The child was asked to point to the word which had the same beginning or ending sound as the picture. This was then followed by a series of three pictures, each with four words beneath, each starting with the same letter as the picture. The child was asked to point to the word that correctly named the picture. Basic reading was assessed using the basic reading subtest of the Wechsler Objective Reading Dimensions (WORD)[2]. In short, the child was asked to read aloud a series of 48 unconnected words which increased in difficulty. The task was stopped after the child made six consecutive errors. The reading accuracy and comprehension score was computed as the sum of the number of items that the child read/responded to correctly.

Reading accuracy age 9 (NBO)

Reading accuracy was assessed by asking the child to read out ten real words selected from a larger selection of words as described by Nunes, Bryant and Olsen (NBO)[3]. The test - retest reliability of word reading was 0.80. The correlation with the Schonell Word Reading Task[4] was 0.85. The reading accuracy score was computed as the sum of the number of items that the child read correctly.

Reading speed and reading accuracy age 9 (NARA II)

Children's reading skills were assessed with the revised Neale Analysis of Reading Ability (NARA II)[5]. The child was asked to read a passage from a booklet. The tester recorded the time it took the child to read the passage, and noted any errors made by the child. All scores were standardised by age.

Reading speed age 13 (TOWRE)

The Test of Word Reading Efficiency (TOWRE)[6] contains a word part to assess sight word efficiency. The child had 45 seconds to read as many words as possible. Words that a child skipped, or got wrong were marked by the tester. The reading speed score was computed as the sum of the number of correct words a child finished on.

Spelling accuracy age 7 (NB)

Spelling accuracy was assessed by asking a child to spell a series of 15 words. The words were chosen specifically for this age group after piloting on several hundred children (Nunes and Bryant, ALSPAC-specific measure). The words included regular and irregular words of differing frequencies, and were put in an order of increasing difficulty. For each word, the tester first read the word out alone to the child, then within a specific

sentence incorporating the word, and finally alone again. The child was asked to write down the spelling. The spelling accuracy score is computed as the number of words spelt correctly.

Spelling accuracy age 9 (NB)

The format to assess spelling accuracy at age 9 was similar to that at age 7. However, the series of 15 words that a child was asked to spell were adjusted to match the age group of 9. The spelling accuracy score is computed as the number of words spelt correctly.

Non-word reading accuracy age 9 (NBO)

This was assessed by asking the child to read out loud ten non-words, selected from a larger selection of non-words taken from research conducted by Nunes and colleagues[3]. The test - retest reliability of the non-word reading task was 0.73. The correlation with the Schonell Word Reading Task[4] was 0.73. The tester emphasised to the child that the words ware made-up, and asked the child to read all the non-words in the way that they thought they should be read. The decoding accuracy score was computed as the sum of the number of items the child read correctly.

Non-word reading speed age 13 (TOWRE)

The Test of Word Reading Efficiency (TOWRE)[6] contains a non-word part to assess decoding efficiency. The child had 45 seconds to read as many non-words as possible. Words that a child skipped, or got wrong were marked by the tester. The decoding speed score is computed as the sum of the number of correct non-words a child finished on.

Phonemic awareness age 7 (AAT)

Phonemic awareness was assessed using the Auditory Analysis Test (AAT)[7]. The task contained two practice and 40 test items of increasing difficulty. For each item, the child was first asked to repeat the word, and then produce it again but with part of the word (a phoneme or a number of phonemes) removed. There were seven omission categories included: omission of a first, medial or final syllable, omission of the initial, omission of the final consonant of a one syllable word, and omission of the first consonant or consonant blend of a medial consonant. The words from different categories were mixed. The phonemic awareness score is computed as the sum of correct responses.

Listening comprehension age 8 (WOLD)

A subset of the Wechsler Objective Language Dimensions (WOLD)[8] test was used to assess listening comprehension. The tester read out loud a paragraph about a picture, shown to the child. After that, the child answers ten questions on what they have heard.

The listening comprehension scores were calculated as the sum of the items that the child got correct.

Non-word repetition age 8 (CNRep)

An adaptation of the Children's Test of Nonword Repetition (CNRep)[9] was used to assess phonological working memory. The test comprised twelve nonsense words, four each of 3, 4 and 5 syllables. The words were conforming to English rules for sound combinations. The child was asked to listen to each word and repeat each item. If there was no phonological deviation from the target form, the repetition attempt was scored as correct. The non-word repetition score was computed as the sum of the number of correct non-words.

Verbal intelligence age 8 (WISC-III)

The Wechsler Intelligence Scale for Children (WISC-III)[10] was used to assess cognitive function. A short form of the measure was employed where alternate items were used for all subtests, with the exception of the coding subtest. The WISC-III comprises ten subtests five of which are verbal subtests: information, similarities, arithmetic, vocabulary, comprehension, and can be used to construct a verbal intelligence score. Raw scores were calculated according to the items used in the alternate item form of the WISC. The total age-scaled scores for the verbal scale were calculated using the look-up tables provided in the WISC manual. All scores were pro-rated.

ALSPAC trait transformation

LRAs were residualised for sex, age and the two most significant ancestry-informative principal components[11], and then rank-transformed unless they were derived using age-specific norms. These scores were adjusted for sex and principal components only before transformation.

ADHD cases within ALSPAC

The Development and Wellbeing Assessment (DAWBA) was used to assess psychological disorders. It is a validated instrument combining structured and semi-structured questions related to DSM and ICD diagnostic criteria[12]. DAWBA was collected through questionnaires posted to parents and teachers, and responses were reviewed by trained clinical raters who assigned diagnoses according to the DSM-IV[13]. Information from both parents and teachers were combined to assign a diagnosis (similar to a clinical setting).

ALSPAC genome-wide analyses

Standard genomic quality control including gender mismatch, heterozygosity, individual missingness, insufficient sample replication, population stratification, minor allele frequency (MAF), SNP call rate, Hardy-Weinberg equilibrium[14], and cryptic relatedness was performed using PLINK[15] (v1.07). After quality control 8 981 children and 465,740 SNPs were imputed to a HRC r1.1 reference panel[16] using the Sanger imputation server (EAGLE2[17] v2.0.5 and PBWT[18] software, https://imputation.sanger.ac.uk/). Genome-wide association analysis summary statistics for all literacy- and language-related measures were generated by regressing rank-transformed residuals on posterior genotype probabilities, assuming an additive genetic model, as implemented in SNPTEST (version 2.5.2) software[19] (without genomic control-based correction[20]).

Clinical ADHD summary statistics

Psychiatric Genomics Consortium (PGC)

ADHD cases (age 5 to 17 years) met diagnostic criteria for either clinical ADHD or hyperkinetic disorder (Diagnostic and Statistical Manual of Mental Disorders (DSM-III[13], DSM-IV[13], DSM-IV-TR[13]) or the International Classification of Diseases (ICD-10[21]).

The Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH)

ADHD cases were diagnosed according to ICD-1021, and identified using the Danish Psychiatric Central Research Register[22]. Controls were randomly selected from the same nationwide birth cohort and did not have a diagnosis of ADHD (F90.0) or moderate-severe mental retardation (F71-F79)[23,24]. Genotyping was performed using the Illumina Infinium PsychArray BeadChip and genotypes were imputed to a 1000 Genomes template[25] (Phase3, release 02-05-2013). Genotyping, quality control, imputation and genetic association analysis were carried out using the Ricopili pipeline with standard PGC settings[26].

Educational attainment summary statistics

Educational attainment (EA) was coded according to the International Standard Classification of Education (1997) scale[27] and analysed as a quantitative variable defined as an individual's years of schooling. Participants were >30 years of age at the time of assessment and of European ancestry. Genome-wide data were predominantly imputed to a 1000 genomes project[25] version 3 reference panel as described previously[28].

Linkage Disequilibrium score regression and correlation

Linkage Disequilibrium score (LDSC) regression[29] can estimate the cumulative effect of genetic variants as tagged by common genetic markers (SNP-h^2) to phenotypic variation, based on GWAS statistics, and distinguishes confounding from polygenic influences in genome-wide analyses[21]. To estimate LDSC-h^2, genome-wide χ^2-statistics

are regressed on the amount of genetic variation captured by each SNP (linkage disequilibrium score, LD score)[29], while the intercept of this regression minus one is an estimator of the mean contribution of confounding bias to the inflation in the mean χ^2-statistic[21] . We estimated the SNP-h^2 for LRAs and EA on the observed scale and ADHD SNP-h^2 on the liability scale (assuming a population prevalence of 5% for ADHD[30]).

In extension, LD score correlation[31] analysis can be applied to estimate genetic correlations (r_g) between genetic variants in distinct samples as a regression of the product of test statistics on LD score. All analyses were performed with LDSC software[29,31] and based on the set of well-imputed HapMap3 SNPs[32] and a European reference panel of LD scores[31]. Unconstrained LD-score correlation[31] analysis was applied to estimate genetic correlations (r_g) between LRAs and EA, as well as ADHD and EA, based on summary statistics from all thirteen LRAs (ALSPAC), EA (SSGAC)[28], and ADHD (PGC+iPSYCH)[23].

Polygenic scoring analyses

Consistent with current guidelines[33], autosomal ADHD GWAS signals were clumped (LD-r^2>0.25, ±500 kb) with PLINK[15] software. Polygenic scores for ADHD were constructed based on *P*-value thresholds of 0.001, 0.01, 0.05, 0.10, 0.3, 0.5, 0.7, 0.9 and 1. Only imputed markers with high imputation quality (INFO[34]>0.8, 95% posterior genotyping probability >0.9 and minor allele frequency >0.005) in ALSPAC were used to generate polygenic scores.

To further illustrate the strength of the genetic overlap, we translated the fitted polygenetic model into the estimated genetic covariance. For this we used reading accuracy at age 9 (NARA II) as an example, where ADHD-PGS explained up to 1.2% of the phenotypic variation when based on the iPSYCH discovery sample. Genetic covariance was estimated from PGS results for reading accuracy at age 9 (NARA II) using Avengeme software[35]. Input parameters were consistent with PGS analyses for a *P*-value threshold of 0.1, including 63,968 SNPs, a discovery sample of N=55,734, a target sample of N=5,201 and an association Z-score of -8.85 (Table S5). We furthermore assumed an ADHD prevalence of 0.05[30], a case sampling proportion of 0.36, an ADHD heritability on the liability scale of 0.21 (Table S4) and 95% null SNPs in the discovery sample.

Overlap in genome-wide significant association signals between ADHD and educational attainment

To identify independent genetic variants passing the genome-wide significance threshold in both the ADHD (PGC+iPSYCH) and EA (SSGAC) GWASs, we first selected all SNPs with $P<5x10^{-8}$ in either GWAS summary statistics. Next, we clumped (LD-r^2>0.25, ±500 kb) these variants with PLINK [15] for ADHD and EA GWASs individually. Both approaches resulted in two independent overlapping variants.

Multivariable regression

In multivariable regression (MVR) analyses, genetic effects for the outcome are regressed on the genetic effects for risk factor and covariates, using a multivariable weighted regression model[36]. The model allows for association between risk factor and outcome (risk-factor specific associations), but also covariate and outcome, based on the same set of instruments, where latter capture pleiotropic effects. However, the model does not imply causal inferences as modelling assumptions are not met.

Using effect estimates from GWAS summary statistics ($\hat{\beta}$), the specific effect of a risk factor (θ_s) conditional on a covariate (θ_C) on the outcome can be estimated using the weighted regression, omitting the intercept term:

$$\hat{\beta}_{outcome} = \theta_s \hat{\beta}_{riskfactor} + \theta_C \hat{\beta}_{\text{cov}\,ariate,} \; weights = se(\hat{\beta}_{outcome})^{-2} \text{ (equation 1)}$$

For a valid estimation of effects the following assumptions need to be met: 1) the genetic variants are associated with the risk factor and covariate, 2) the genetic variants are not associated with confounders, and 3) there is no pathway from any genetic variant to the outcome except via the risk factor and/or confounder[36].

In order to investigate whether the third assumption holds, we investigated the evidence for a regression intercept in an unconstrained regression model. An intercept consistent with zero (i.e. within the 95% confidence interval) suggests that there is no evidence for additional pleiotropic effects. Since the intercept is, however, sensitive to the defined direction of effect, unconstrained models were fitted twice, once with genetic variants aligned according to the risk factor and once according to the covariate.

Meta-regression across literacy- and language-related abilities

The genetic and phenotypic inter-relatedness among LRAs needs to be accounted for when combining ADHD effects (shared with and independent of EA) across multiple LRA combinations. A variance/covariance matrix across the correlated LRAs was approximated analogous to models accounting for correlated phylogenetic histories[37]. The matrix was based on the observed phenotypic correlation matrix using rank-transformed measures (Table S1) and weighted by the standard errors of the estimated polygenic association. The model included, for each LRA combination, one random intercept. Evidence for polygenic effect heterogeneity was assessed using Cochran's Q-test.

MVR estimates for defined LRA combinations were contrasted with each other by conducting moderator analyses as part of a random-effects meta-regression model (R:metafor library[38], Rv3.2.0) across the entire set of LRAs studied (at a significance level of 0.05). This includes a dummy coded contrast, the moderator (mod), that is added to the model in order to explain heterogeneity in effect estimates, and an assessment of the remaining residual effect heterogeneity (res het). Note that verbal intelligence

quotient scores were excluded from the analysis, as they represent a composite measure.

Standardisation of genetic effect estimates

To compare the magnitude of regression estimates capturing polygenic ADHD effects on LRAs independent of and shared with EA, we calculated standardised SNP effects for ADHD (PGC+iPSYCH), EA (SSGAC) and LRAs (ALSPAC) from GWAS summary statistics.

Standardised regression coefficients for SNP j with minor allele frequency MAF_j and sample size N_j were calculated as[27]

$$\hat{\beta}_j = \frac{z_j \times \widehat{\sigma_y}}{\sqrt{N_j \times \hat{\sigma}^2_{x_j}}} \quad \text{(equation 2)}$$

where z_j is the Wald test statistic, $\hat{\sigma}_y$ is the standard deviation of the phenotype y, and $\hat{\sigma}^2_{x_j}$ corresponds to the SNP variance 2 x MAF_j x (1-MAF_j). The corresponding standard error was calculated as[27]

$$SE(\hat{\beta}_j) = \sqrt{\frac{1}{N_j}\left(\frac{\hat{\sigma}^2_y - \hat{\beta}^2_j \cdot \hat{\sigma}^2_{xj}}{\hat{\sigma}^2_{xj}}\right)} \quad \text{(equation 3)}$$

We estimated $\hat{\sigma}_y$ for each component of the multivariable regression model as follows.

Literacy- and language-related abilities

All population-based linguistic traits were rank-transformed (continuous traits) and thus $\hat{\sigma}_y$ equals to one.

Educational attainment

The combined $\hat{\sigma}_y$ for EA, measured as continuous score in years of schooling, was pooled from 63 cohorts[28,39]

$$\hat{\sigma}_y = \sqrt{\frac{\sum_{i=1}^{k}(n_i - 1)\cdot \sigma_i^2}{\left(\sum_{i=1}^{k} n_i\right) - k}} \quad \text{(equation 4)}$$

with sample size n_i , sample variance σ_i^2 , and number of studies k .

ADHD

The estimated standard deviation of the liability of ADHD $\hat{\sigma}_{y*}$, taking the prevalence of the disorder K (0.05 for ADHD[30]) and sample prevalence P (0.364 for PGC+iPSYCH ADHD[23]) into account, was calculated as[40]

$$\hat{\sigma}_{y*} = \sqrt{1 + \frac{z_{trans}(P-K)}{K(1-K)} \cdot \left(t - \frac{z_{trans}(P-K)}{K(1-K)} \right)} \qquad \text{(equation 5)}$$

where t is the truncation threshold (derived as the inverse standard normal distribution function at $1-K$ and z_{trans} the height of the standard normal probability density function at t.

Attrition analysis in ALSPAC

A dichotomous variable indicating missingness (coded as 1) for reading accuracy and comprehension at age 7 (WORD) was created for each child that was alive at the age of one year and had genotype data available (N_{total}=8,095; $N_{missing}$=2,204). Genome-wide association analysis summary statistics were generated by logistic regression of sample drop-out on posterior genotype probabilities, assuming an additive genetic model, as implemented in SNPTEST (version 2.5.2) software[19] (without genomic control-based correction[20]).

ADHD polygenic scores were regressed on data missingness for reading accuracy and comprehension at age 7 (WORD) using logistic regression (R:stats library, Rv3.2.3) to estimate odds ratios and their standard errors per one-standard-deviation increase in polygenic scores (Methods). McFadden's pseudo-R^2 values for logistic regression were estimated (R:pscl library, Rv3.2.3), analogous to the ordinary least square regression R^2.

Utilising a multivariable regression approach[36] (see Methods), analysis were subsequently carried out conditional on genetically predicted educational attainment (EA), a potential correlate for non-participation[41], using GWAS summary statistics. We dissected ADHD polygenic influences on LRA-missingness into genetic effects that are independent of EA and genetic effects that are shared with EA, using both conservative (*P*-threshold $P<5\times10^{-8}$) and subthreshold ADHD variants (*P*-threshold $P<0.0015$), as described for multivariable regression analyses (see Methods).

We identified a positive genetic association between sample-dropout and ADHD-PGS (PGC+iPSYCH, OR=1.03(SE=0.005), $P=1.4\times10^{-8}$, Table S12), consistent with previous studies[42]. Disentangling this polygenic link using MVR (Table S13) based on ADHD subthreshold instruments showed a 0.05 increase in log-odds sample drop-out per log-odds increase in liability to ADHD, independent of EA (log(OR)=0.05(SE=0.01), $P=3.7\times10^{-4}$). Additionally, a further 1.04 increase in log odds sample drop-out per missing year of schooling (log(OR)=1.04(SE=0.19),$P=7.3\times10^{-8}$) was observed, suggesting also ADHD effects shared with EA.

Web resources

ALSPAC data dictionary:
http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/
ALSPAC variable catalogue:
http://www.bristol.ac.uk/alspac/researchers/access/
PGC: http://www.med.unc.edu/pgc
PLINK: https://www.cog-genomics.org/plink2
HRC: http://www.haplotype-reference-consortium.org/
SANGER IMPUTATION SERVER: https://imputation.sanger.ac.uk/
SNPTEST: https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html
LDSC: https://github.com/bulik/ldsc
GCTA: http://cnsgenomics.com/software/gcta/#Overview
GCTA power: http://cnsgenomics.com/shiny/gctaPower/
R: https://www.r-project.org/
METAFOR: http://www.metafor-project.org/doku.php

Supplementary Tables

	Reading a/c 7 (WORD)	Reading a 9 (NBO)	Reading s 9 (NARA II)	Reading a 9 (NARA II)	Reading s 13 (TOWRE)	NW reading a 9 (NBO)	NW reading s 13 (TOWRE)	Spelling a 7 (NB)	Spelling a 9 (NB)	PhonAware 7 (AAT)	Listening c 8 (WOLD)	Non-word repetition 8 (CNRep)	VIQ 8 (WISC-III)
Reading a/c 7 (WORD)	1	0.72	0.71	0.83	0.54	0.66	0.59	0.84	0.75	0.70	0.27	0.42	0.55
Reading a 9 (NBO)	0.72	1	0.63	0.76	0.50	0.72	0.56	0.67	0.71	0.57	0.23	0.36	0.45
Reading s 9 (NARA II)	0.71	0.63	1	0.74	0.62	0.56	0.65	0.63	0.65	0.48	0.30	0.34	0.50
Reading a 9 (NARA II)	0.82	0.77	0.74	1	0.59	0.70	0.65	0.75	0.78	0.64	0.29	0.44	0.55
Reading s 13 (TOWRE)	0.53	0.50	0.62	0.60	1	0.45	0.81	0.48	0.52	0.41	0.24	0.29	0.42
NW reading a 9 (NBO)	0.66	0.72	0.56	0.70	0.45	1	0.53	0.63	0.66	0.55	0.17	0.33	0.39
NW reading s 13 (TOWRE)	0.58	0.57	0.65	0.65	0.81	0.53	1	0.54	0.58	0.45	0.20	0.28	0.38
Spelling a 7 (NB)	0.84	0.66	0.62	0.74	0.47	0.62	0.53	1	0.77	0.66	0.19	0.35	0.46
Spelling a 9 (NB)	0.74	0.71	0.65	0.78	0.51	0.66	0.58	0.75	1	0.59	0.21	0.36	0.47
PhonAware 7 (AAT)	0.70	0.57	0.48	0.64	0.40	0.54	0.45	0.66	0.58	1	0.19	0.38	0.45
Listening c 8 (WOLD)	0.27	0.23	0.31	0.29	0.24	0.17	0.20	0.20	0.20	0.19	1	0.23	0.44
Non-word repetition 8 (CNRep)	0.42	0.36	0.34	0.43	0.29	0.32	0.28	0.35	0.36	0.38	0.23	1	0.37
VIQ 8 (WISC-III)	0.54	0.45	0.51	0.56	0.42	0.39	0.39	0.45	0.47	0.45	0.43	0.37	1

Table S1: Phenotypic correlations among literacy- and language-related measures

Phenotypic correlations are depicted as Pearson's correlation coefficients. The upper triangle represents phenotypic correlations based on rank-transformed measures, whereas the lower triangle represents phenotypic correlations based on untransformed measures. Abbreviations: a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness, AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III

Table S2: SNP-heritability estimates for literacy- and language-related measures

LRAs	LDSC					REML	
	SNP-h² (SE)	Mean chi-square	N	λ_{GC}	Intercept (SE)	SNP-h² (SE)	N[1]
Reading a/c 7 (WORD)	0.35 (0.08)	1.05	5 891	1.05	1.01 (0.01)	0.42 (0.06)	5 723
Reading a 9 (NBO)	0.31 (0.09)	1.05	5 738	1.05	1.01 (0.01)	0.46 (0.06)	5 574
Reading s 9 (NARA II)	0.32 (0.10)	1.05	5 189	1.03	1.01 (0.01)	0.45 (0.07)	5 037
Reading a 9 (NARA II)	0.43 (0.10)	1.05	5 201	1.05	1.01 (0.01)	0.50 (0.07)	5 048
Reading s 13 (TOWRE)	0.24 (0.11)	1.03	4 247	1.03	1.01 (0.01)	0.40 (0.09)	4 131
NW reading a 9 (NBO)	0.27 (0.09)	1.04	5 731	1.04	1.01 (0.01)	0.32 (0.06)	5 569
NW reading s 13 (TOWRE)	0.18 (0.11)	1.03	4 237	1.03	1.01 (0.01)	0.38 (0.09)	4 121
Spelling a 7 (NB)	0.29 (0.08)	1.04	5 800	1.04	1.00 (0.01)	0.32 (0.06)	5 637
Spelling a 9 (NB)	0.25 (0.09)	1.04	5 728	1.04	1.01 (0.01)	0.38 (0.06)	5 564
PhonAware 7 (AAT)	0.34 (0.08)	1.05	5 919	1.04	1.00 (0.01)	0.39 (0.06)	5 749
Listening c 8 (WOLD)	0.23 (0.08)	1.03	5 473	1.03	1.01 (0.01)	0.32 (0.07)	5 324
Non-word repetition 8 (CNRep)	0.14 (0.09)	1.03	5 464	1.04	1.01 (0.01)	0.32 (0.07)	5 315
VIQ 8 (WISC-III)	0.42 (0.10)	1.06	5 456	1.06	1.02 (0.01)	0.54 (0.07)	5 305

1. Differences compared with the total sample N are due to excluding individuals with a genetic relationship of ≥0.05[43]. Abbreviations: LRAs, literacy and language-related abilities; LDSC, LD score regression; REML, Restricted Maximum Likelihood (REML) analyses as implemented in genome-wide complex trait analysis software; N, sample size; λ_{GC}, lambda GC; a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III

Table S3: Genetic correlations among literacy- and language-related measures

	Reading a/c 7 (WORD)	Reading a 9 (NBO)	Reading s 9 (NARA II)	Reading a 9 (NARA II)	Reading s 13 (TOWRE)	NW reading a 9 (NBO)	NW reading s (TOWRE)	Spelling a 7 (NB)	Spelling a 9 (NB)	PhonAware 7 (AAT)	Listening c 8 (WOLD)	Non-word repetition 8 (CNRep)	VIQ 8 (WISC-III)
Reading a/c 7 (WORD)	1												
Reading a 9 (NBO)	0.88 (0.04)	1											
Reading s 9 (NARA II)	0.97 (0.05)	0.84 (0.06)	1										
Reading a 9 (NARA II)	1.00 (0.03)	0.89 (0.04)	0.93 (0.04)	1									
Reading s 13 (TOWRE)	0.81 (0.09)	0.67 (0.09)	0.94 (0.08)	0.90 (0.08)	1								
NW reading a 9 (NBO)	0.92 (0.06)	0.98 (0.05)	0.81 (0.08)	0.93 (0.05)	0.68 (0.11)	1							
NW reading s (TOWRE)	0.88 (0.08)	0.92 (0.08)	0.92 (0.07)	0.94 (0.07)	0.87 (0.04)	0.96 (0.11)	1						
Spelling a 7 (NB)	0.96 (0.03)	0.87 (0.06)	0.82 (0.07)	0.96 (0.05)	0.71 (0.11)	0.93 (0.08)	0.86 (0.11)	1					
Spelling a 9 (NB)	0.94 (0.04)	0.95 (0.04)	0.90 (0.06)	0.96 (0.03)	0.83 (0.09)	0.46 (0.14)	0.49 (0.15)	0.97 (0.04)	1				
PhonAware 7 (AAT)	0.97 (0.04)	0.86 (0.07)	0.86 (0.09)	0.97 (0.06)	0.77 (0.12)	0.93 (0.08)	0.86 (0.11)	0.98 (0.05)	0.87 (0.10)	1			
Listening c 8 (WOLD)	0.74 (0.11)	0.49 (0.12)	0.80 (0.12)	0.67 (0.11)	0.58 (0.14)	0.46 (0.14)	0.49 (0.15)	0.73 (0.14)	0.55 (0.13)	0.64 (0.13)	1		
Non-word repetition 8 (CNRep)	0.68 (0.10)	0.61 (0.61)	0.69 (0.12)	0.73 (0.10)	0.41 (0.14)	0.72 (0.12)	0.40 (0.15)	0.66 (0.12)	0.81 (0.11)	0.70 (0.11)	0.54 (0.14)	1	
VIQ 8 (WISC-III)	0.91 (0.06)	0.66 (0.07)	0.86 (0.07)	0.84 (0.06)	0.80 (0.10)	0.68 (0.09)	0.74 (0.10)	0.81 (0.08)	0.72 (0.08)	0.77 (0.07)	0.88 (0.09)	0.71 (0.09)	1

Genetic correlations were calculated based on rank-transformed scores using Restricted Maximum Likelihood (REML) analyses as implemented in genome-wide complex trait analysis software, based on samples of individuals with a genetic relationship of <0.05. Standard errors are provided in brackets. Abbreviations: a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Non-Word Repetition test; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III

Table S4: SNP-heritability estimates for clinical ADHD and educational attainment

Phenotype	Sample	SNP-h^2 (SE)	λ_{GC}	Intercept (SE)
ADHD	PGC	0.08 (0.03)	1.00	0.99 (0.01)
	iPSYCH	0.26 (0.02)	1.23	1.03 (0.01)[b]
	PGC+iPSYCH (EUR)	0.22 (0.01)	1.25	1.04 (0.01)[b]
	PGC+iPSYCH	0.21 (0.01)	1.25	1.04 (0.01)[b]
EA	SSGAC	0.11 (0.004)	1.47	0.96 (0.01)

SNP-heritability was estimated with LDSC regression analysis. SNP-heritability estimates for ADHD samples were calculated on a liability scale assuming a population prevalence of 0.05[30]. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; λ_{GC}, lambda GC; PGC, Psychiatric Genomics Consortium; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; EUR, European ancestry; SSGAC, Social Science Genetic Association Consortium.

Table S5: Association of polygenic ADHD risk scores with literacy- and language-related measures

LRAs	ADHD sample	β (SE)	*P*	R^2 (%)
Reading a/c 7 (WORD)	PGC	-0.04 (0.01)	$4.6x10^{-3}$	0.14
	iPSYCH	-0.10 (0.01)	$<1x10^{-10}$	0.99
	PGC+iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.33
	PGC+iPSYCH(EUR)	-0.12 (0.01)	$<1x10^{-10}$	1.42
Reading a 9 (NBO)	PGC	-0.02 (0.01)	$6.4x10^{-2}$	0.06
	iPSYCH	-0.09 (0.01)	$<1x10^{-10}$	0.73
	PGC+iPSYCH	-0.10 (0.01)	$<1x10^{-10}$	0.93
	PGC+iPSYCH(EUR)	-0.10 (0.01)	$<1x10^{-10}$	0.92
Reading s 9 (NARA II)	PGC	-0.04 (0.01)	$7.6x10^{-3}$	0.14
	iPSYCH	-0.10 (0.01)	$<1x10^{-10}$	1.02
	PGC+iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.18
	PGC+iPSYCH(EUR)	-0.11 (0.01)	$<1x10^{-10}$	1.28
Reading a 9 (NARA II)	PGC	-0.04 (0.01)	$5.7x10^{-3}$	0.15
	iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.20
	PGC+iPSYCH	-0.12 (0.01)	$<1x10^{-10}$	1.49
	PGC+iPSYCH(EUR)	-0.12 (0.01)	$<1x10^{-10}$	1.47
Reading s 13 (TOWRE)	PGC	-0.03 (0.02)	$2.5x10^{-2}$	0.12
	iPSYCH	-0.09 (0.02)	$4.2x10^{-9}$	0.81
	PGC+iPSYCH	-0.11 (0.02)	$<1x10^{-10}$	1.19
	PGC+iPSYCH(EUR)	-0.11 (0.02)	$<1x10^{-10}$	1.22
NW reading a 9 (NBO)	PGC	-0.03 (0.01)	$4.8x10^{-2}$	0.07
	iPSYCH	-0.07 (0.01)	$4.7x10^{-8}$	0.52
	PGC+iPSYCH	-0.09 (0.01)	$<1x10^{-10}$	0.76
	PGC+iPSYCH(EUR)	-0.09 (0.01)	$<1x10^{-10}$	0.77
NW reading s 13 (TOWRE)	PGC	-0.03 (0.02)	$4.7x10^{-2}$	0.09
	iPSYCH	-0.09 (0.02)	$2.6x10^{-8}$	0.73
	PGC+iPSYCH	-0.11 (0.02)	$<1x10^{-10}$	1.14
	PGC+iPSYCH(EUR)	-0.10 (0.02)	$<1x10^{-10}$	1.08
Spelling a 7 (NB)	PGC	-0.03 (0.01)	$9.6x10^{-3}$	0.12
	iPSYCH	-0.09 (0.01)	$<1x10^{-10}$	0.86
	PGC+iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.17
	PGC+iPSYCH(EUR)	-0.11 (0.01)	$<1x10^{-10}$	1.16
Spelling a 9 (NB)	PGC	-0.04 (0.01)	1.5x10-3	0.18
	iPSYCH	-0.09 (0.01)	<1x10-10	0.81
	PGC+iPSYCH	-0.10 (0.01)	<1x10-10	1.08
	PGC+iPSYCH(EUR)	-0.10 (0.01)	<1x10-10	1.06
PhonAware 7 (AAT)	PGC	-0.02 (0.01)	2.4x10-1	0.02
	iPSYCH	-0.08 (0.01)	2.9x10-9	0.59
	PGC+iPSYCH	-0.09 (0.01)	<1x10-10	0.84
	PGC+iPSYCH(EUR)	-0.10 (0.01)	<1x10-10	0.98
Listening c 8 (WOLD)	PGC	-0.02 (0.01)	1.9x10-1	0.03
	iPSYCH	-0.06 (0.01)	9.9x10-6	0.36
	PGC+iPSYCH	-0.08 (0.01)	2.5x10-8	0.57
	PGC+iPSYCH(EUR)	-0.08 (0.01)	9.2x10-10	0.68

Table S5: Association of polygenic ADHD risk scores with literacy- and language-related measures - continued

LRAs	ADHD sample	β (SE)	*P*	R^2 (%)
Non-word repetition 8 (CNRep)	PGC	-0.003 (0.01)	$8.1x10^{-1}$	0.001
	iPSYCH	-0.05 (0.01)	$1.7x10^{-4}$	0.26
	PGC+iPSYCH	-0.06 (0.01)	$6.3x10^{-6}$	0.37
	PGC+iPSYCH(EUR)	-0.07 (0.01)	$2.3x10^{-7}$	0.49
VIQ 8 (WISC-III)	PGC	-0.03 (0.01)	$2.8x10^{-2}$	0.09
	iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.28
	PGC+iPSYCH	-0.13 (0.01)	$<1x10^{-10}$	1.59
	PGC+iPSYCH(EUR)	-0.13 (0.01)	$<1x10^{-10}$	1.69

ADHD SNPs were selected from GWAS summary statistics based on a *P*-value threshold of 0.1, and alleles were aligned such that the effect allele increased ADHD risk. LRAs were regressed on polygenic ADHD risk scores using ordinary least square regression. Effects were considered as significant if they passed the experiment-wide significance threshold ($P<0.007$). Abbreviations: LRAs, literacy- and language-related abilities; ADHD, Attention-Deficit/Hyperactivity Disorder; R^2, OLS-regression R^2; a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; PGC, Psychiatric Genomics Consortium; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; EUR, European descent.

Table S6: Association of polygenic ADHD risk scores with literacy- and language-related measures excluding ADHD children

LRAs	ADHD sample	β (SE)	*P*	R^2 (%)
Reading a/c 7 (WORD)	PGC	-0.04 (0.01)	$3.5x10^{-3}$	0.15
	iPSYCH	-0.10 (0.01)	$<1x10^{-10}$	0.99
	PGC+iPSYCH	-0.12 (0.01)	$<1x10^{-10}$	1.33
Reading a 9 (NBO)	PGC	-0.02 (0.01)	0.09	0.05
	iPSYCH	-0.09 (0.01)	$<1x10^{-10}$	0.74
	PGC+iPSYCH	-0.10 (0.01)	$<1x10^{-10}$	0.93
Reading s 9 (NARA II)	PGC	-0.04 (0.01)	$7.3x10^{-3}$	0.14
	iPSYCH	-0.10 (0.01)	$<1x10^{-10}$	1.04
	PGC+iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.19
Reading a 9 (NARA II)	PGC	-0.04 (0.01)	$6.5x10^{-3}$	0.14
	iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.22
	PGC+iPSYCH	-0.12 (0.01)	$<1x10^{-10}$	1.50
Reading s 13 (TOWRE)	PGC	-0.03 (0.02)	0.03	0.11
	iPSYCH	-0.09 (0.02)	$2.8x10^{-9}$	0.84
	PGC+iPSYCH	-0.11 (0.02)	$<1x10^{-10}$	1.23
NW reading a 9 (NBO)	PGC	-0.03 (0.01)	0.05	0.07
	iPSYCH	-0.07 (0.01)	$4.2x10^{-8}$	0.53
	PGC+iPSYCH	-0.09 (0.01)	$<1x10^{-10}$	0.77
NW reading s 13 (TOWRE)	PGC	-0.03 (0.02)	0.05	0.09
	iPSYCH	-0.09 (0.02)	$2.5x10^{-8}$	0.74
	PGC+iPSYCH	-0.11 (0.02)	$<1x10^{-10}$	1.17
Spelling a 7 (NB)	PGC	-0.04 (0.01)	$7.9x10^{-3}$	0.12
	iPSYCH	-0.09 (0.01)	$<1x10^{-10}$	0.86
	PGC+iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.17
Spelling a 9 (NB)	PGC	-0.04 (0.01)	$1.4x10^{-3}$	0.18
	iPSYCH	-0.09 (0.01)	$<1x10^{-10}$	0.79
	PGC+iPSYCH	-0.10 (0.01)	$<1x10^{-10}$	1.06
PhonAware 7 (AAT)	PGC	-0.02 (0.01)	0.22	0.03
	iPSYCH	-0.08 (0.01)	$3.8x10^{-9}$	0.60
	PGC+iPSYCH	-0.09 (0.01)	$<1x10^{-10}$	0.86
Listening c 8 (WOLD)	PGC	-0.02 (0.01)	0.13	0.04
	iPSYCH	-0.06 (0.01)	$7.9x10^{-6}$	0.37
	PGC+iPSYCH	-0.08 (0.01)	$1.7x10^{-8}$	0.59
Non-word repetition 8 (CNRep)	PGC	-0.01 (0.01)	0.70	0.003
	iPSYCH	-0.05 (0.01)	$2.4x10^{-4}$	0.25
	PGC+iPSYCH	-0.06 (0.01)	$7.6x10^{-6}$	0.37
VIQ 8 (WISC-III)	PGC	-0.03 (0.01)	0.02	0.10
	iPSYCH	-0.11 (0.01)	$<1x10^{-10}$	1.27
	PGC+iPSYCH	-0.13 (0.01)	$<1x10^{-10}$	1.60

ADHD SNPs were selected from GWAS summary statistics based on a *P*-value threshold of 0.1, and alleles were aligned such that the effect allele increased ADHD risk. Children with ADHD were excluded from the ALSPAC sample based on the Development and Wellbeing Assessment[12]. LRAs were regressed on ADHD-PGS using ordinary least square regression. Effects were considered as significant if they passed the experiment-wide significance threshold ($P<0.007$). Abbreviations: LRAs, literacy- and language-related abilities; ADHD, Attention-Deficit/Hyperactivity Disorder; R^2, OLS-regression R^2; a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading

Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; PGC, Psychiatric Genomics Consortium; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research.

Table S7: Genetic correlations of literacy- and language-related measures with educational attainment

LRAs	r_g (SE)	*P*
Reading a/c 7 (WORD)	0.65 (0.08)	$<1x10^{-10}$
Reading a 9 (NBO)	0.57 (0.11)	$2.0x10^{-7}$
Reading s 9 (NARA II)	0.77 (0.12)	$<1x10^{-10}$
Reading a 9 (NARA II)	0.64 (0.08)	$<1x10^{-10}$
Reading s 13 (TOWRE)	0.80 (0.22)	$3.0x10^{-4}$
NW reading a 9 (NBO)	0.61 (0.14)	$2.1x10^{-5}$
NW reading s 13 (TOWRE)	0.89 (0.31)	$3.9x10^{-3}$
Spelling a 7 (NB)	0.57 (0.08)	$<1x10^{-10}$
Spelling a 9 (NBO)	0.69 (0.12)	$1.8x10^{-8}$
PhonAware 7 (AAT)	0.56 (0.09)	$8.6x10^{-10}$
Listening c 8 (WOLD)	0.62 (0.12)	$4.6x10^{-7}$
Non-word repetition 8 (CNRep)	0.68 (0.25)	$5.6x10^{-3}$
VIQ 8 (WISC-III)	0.82 (0.10)	$<1x10^{-10}$

Genetic correlations were estimated with unconstrained LD-score correlation analyses[34]. Genetic correlations were considered as significant if they passed the experiment-wide significance threshold ($P<0.007$). Abbreviations: LRAs, literacy- and language-related abilities; r_g, genetic correlation; a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III.

Table S8: Selection of ADHD and EA instruments

LRAs	ADHD-associated instruments		EA-associated instruments	
	Conservative (P_{thr}<5x10^{-8})	Subthreshold P_{thr}<0.0015)	Conservative (P_{thr}<5x10^{-8})	Subthreshold P_{thr}<0.0015)
Reading a/c 7 (WORD)	15	2,690	99	4,611
Reading a 9 (NBO)	15	2,690	99	4,611
Reading s 9 (NARA II)	15	2,689	99	4,608
Reading a 9 (NARA II)	15	2,689	99	4,608
Reading s 13 (TOWRE)	15	2,688	99	4,608
NW reading a 9 (NBO)	15	2,690	99	4,612
NW reading s 13 (TOWRE)	15	2,688	99	4,609
Spelling a 7 (NB)	15	2,691	99	4,613
Spelling a 9 (NB)	15	2,690	99	4,611
PhonAware 7 (AAT)	15	2,689	99	4,612
Listening c 8 (WOLD)	15	2,689	99	4,613
Non-word repetition 8 (CNRep)	15	2,688	99	4,613
VIQ 8 (WISC-III)	15	2,688	99	4,612

ADHD and EA instruments were selected based on ADHD (PGC+iPSYCH) and EA (SSGAC) GWAS summary statistics respectively. Conservative instruments passed the genome-wide significance level (P<5x10^{-8}), whereas the subthreshold set, containing typically defined instruments, was based on a more lenient P-value threshold (P<0.0015). Abbreviations: LRAs, literacy- and language-related abilities; ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; P_{thr}, P-value threshold; a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; PGC, Psychiatric Genetics Consortium; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research

Table S9: Multivariable regression analysis of polygenic associations between ADHD and literacy- and language-related abilities (standardised)

LRAs	ADHD (β_{ADHD}) (ADHD-specific effects independent of EA)						EA (β_{EA}) (ADHD genetic effects shared with EA)[1]					
	Conservative instruments (P_{thr}<5x10^{-8})			Subthreshold instruments (P_{thr}<0.0015)			Conservative instruments (P_{thr}<5x10^{-8})			Subthreshold instruments (P_{thr}<0.0015)		
	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}
Reading a/c 7 (WORD)	-0.21 (0.23)	0.39	-	-0.06 (0.01)	8.3x10^{-6}	-	0.12 (0.27)	0.66	-	-0.15 (0.02)	1.5x10^{-9}	-
Reading a 9 (NBO)	-0.46 (0.16)	0.01	-	-0.05 (0.01)	1.4x10^{-4}	-	-0.27 (0.19)	0.18	-	-0.11 (0.02)	7.0x10^{-6}	-
Reading s 9 (NARA II)	-0.44 (0.23)	0.07	-	-0.06 (0.01)	3.4x10^{-5}	-	-0.30 (0.27)	0.29	-	-0.12 (0.03	2.2x10^{-6}	-
Reading a 9 (NARA II)	-0.51 (0.22)	0.04	-	-0.05 (0.01)	3.8x10^{-4}	-	-0.27 (0.26)	0.33	-	-0.15 (0.03)	3.4x10^{-9}	-
Reading s 13 (TOWRE)	-0.63 (0.25)	0.02	-	-0.07 (0.02)	2.1x10^{-5}	-	-0.48 (0.29)	0.12	-	-0.11 (0.03)	2.3x10^{-4}	-
NW reading a 9 (NBO)	-0.56 (0.18)	0.01	-	-0.04 (0.01)	8.5x10^{-4}	-	-0.26 (0.21)	0.25	-	-0.11 (0.02)	4.6x10^{-6}	-
NW reading s 13 (TOWRE)	-0.32 (0.30)	0.31	-	-0.06 (0.02)	1.4x10^{-4}	-	-0.28 (0.35)	0.45	-	-0.10 (0.03)	2.4x10^{-4}	-
Spelling a 7 (NB)	0.004 (0.21)	0.98	-	-0.08 (0.01)	9.7x10^{-9}	-	0.34 (0.25)	0.21	-	-0.09 (0.02)	1.5x10^{-4}	-
Spelling a 9 (NB)	-0.34 (0.14)	0.03	-	-0.07 (0.01)	6.4x10^{-7}	-	-0.27 (0.17)	0.14	-	-0.11 (0.02)	1.3x10^{-5}	-
PhonAware 7 (AAT)	-0.25 (0.26)	0.34	-	-0.04 (0.01)	0.002	-	0.20 (0.31)	0.53	-	-0.16 (0.02)	<1x10^{-10}	-
Listening c 8 (WOLD)	-0.05 (0.13)	0.74	-	-0.03 (0.01)	0.02	-	0.33 (0.16)	0.05	-	-0.13 (0.03)	3.9x10^{-7}	-
Non-word repetition 8 (CNRep)	-0.20 (0.25)	0.42	-	-0.03 (0.01)	0.02	-	0.08 (0.29)	0.79	-	-0.11 (0.03)	1.2x10^{-5}	-
VIQ 8 (WISC-III)	-0.30 (0.27)	0.29	-	-0.06 (0.01)	5.2x10^{-5}	-	0.18 (0.32)	0.57	-	-0.17 (0.02)	<1x10^{-10}	-
Pooled reading	-0.55 (0.16)	7.2x10^{-4}	0.13	-0.06 (0.01)	1.3x10^{-6}	0.69	0.34 (0.19)	0.08	0.19	-0.12 (0.02)	5.0x10^{-8}	0.09
Pooled spelling	-0.22 (0.20)	0.27	0.01	-0.07 (0.01)	8.4x10^{-9}	0.31	0.01 (0.31)	0.97	3.0x10^{-4}	-0.09 (0.02)	1.3x10^{-5}	0.40
Pooled LRAs	-0.27 (0.13)	0.03	0.01	-0.05 (0.01)	1.4x10^{-6}	0.05	0.02 (0.16)	0.91	0.002	-0.12 (0.02)	<1x10^{-10}	0.002

Sets of conservative (P<5x10^{-8}) and subthreshold (P<0.0015) ADHD instruments were extracted from ADHD (PGC+iPSYCH), EA (SSGAC) and LRAs (ALSPAC) GWAS summary statistics. ADHD-specific effects independent of EA (β_{ADHD}) and ADHD effects shared with EA (β_{EA}) on LRAs were estimated with MVRs. To compare the magnitude of MVR estimates, analyses were conducted using standardised regression estimates (Supplementary Methods). β_{ADHD} estimates measure the change in LRA Z-score per Z-score in ADHD liability. β_{EA} estimates measure the change in LRA Z-scores per Z-score in missing school years. MVR estimates based on raw genetic effect estimates are provided in Table 3. Pooled estimates for reading, spelling and global LRA measures (Table 1) were obtained through random-effects meta-regression. Effects were considered significant if they passed the experiment-wide significance threshold (*P*<0.007). 1. ADHD genetic effects shared with EA as assessed through EA genetic effects of ADHD-associated variants. Abbreviations: LRAs, literacy- and language-related abilities; ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; P_{thr}, *P*-value threshold; P_{het}, Heterogeneity *P*-value; a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware; phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III; MVR, Multivariable regression

Table S10: Comparison of ADHD-specific MVR effects on literacy-related abilities versus other LRAs

ADHD instruments	Effect	Beta (SE)	*P*	$P_{res\ het}$
Conservative (P_{thr}<5x10^{-8})	Reading	-0.300 (0.082)	2x10^{-4}	0.054
	Other LRAs (Moderator)	0.150 (0.059)	0.011	
Subthreshold (P_{thr}<0.0015)	Reading	-0.024 (0.006)	1x10^{-4}	0.054
	Other LRAs (Moderator)	-0.003 (0.004)	0.427	
	Spelling	-0.036 (0.006)	2x10^{-9}	0.56
	Other LRAs (Moderator)	0.012 (0.004)	0.001	
	Reading+spelling	-0.035 (0.006)	5x10^{-8}	0.22
	Other LRAs (Moderator)	0.012 (0.005)	0.016	

ADHD-specific effect differences on LRAs were compared using contrasts within a random-effects meta-regression model, based on all LRAs studied, while accounting for phenotypic inter-correlations. Moderator effects were considered significant at a significance level of 0.05. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; LRAs, language- and literacy-related abilities; $P_{res\ het}$ - Evidence for residual effect heterogeneity; P_{thr}, *P*-value threshold

Table S11: Multivariable regression analysis of polygenic associations between EA and literacy- and language-related abilities (standardised)

LRAs	EA (β_{EA}) (EA-specific effects independent of ADHD)						ADHD (β_{ADHD}) (EA genetic effects shared with ADHD)[1]					
	Conservative instruments (P_{thr}<5x10^{-8})			Subthreshold instruments (P_{thr}<0.0015)			Conservative instruments (P_{thr}<5x10^{-8})			Subthreshold instruments (P_{thr}<0.0015)		
	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}	β (SE)	*P*	P_{het}
Reading a/c 7 (WORD)	-0.16 (0.03)	4.4x10^{-6}	-	-0.11 (0.01)	<1x10^{-10}	-	-0.27 (0.14)	0.07	-	-0.16 (0.03)	6.9x10^{-9}	-
Reading a 9 (NBO)	-0.14 (0.03)	5.8x10^{-5}	-	-0.09 (0.01)	<1x10^{-10}	-	-0.24 (0.14)	0.09	-	-0.16 (0.03)	1.4x10^{-8}	-
Reading s 9 (NARA II)	-0.21 (0.04)	1.3x10^{-7}	-	-0.11 (0.01)	<1x10^{-10}	-	-0.13(0.16)	0.41	-	-0.15 (0.03)	2.8x10^{-7}	-
Reading a 9 (NARA II)	-0.18 (0.04)	4.0x10^{-6}	-	-0.11 (0.01)	<1x10^{-10}	-	-0.42 (0.16)	0.01	-	-0.17 (0.03)	8.9x10^{-8}	-
Reading s 13 (TOWRE)	-0.17 (0.04)	3.7x10^{-5}	-	-0.09 (0.01)	<1x10^{-10}	-	-0.09 (0.17)	0.59	-	-0.14 (0.03)	6.0x10^{-5}	-
NW reading a 9 (NBO)	-0.13 (0.03)	1.3x10^{-4}	-	-0.09 (0.01)	<1x10^{-10}	-	-0.38 (0.14)	0.01	-	-0.09 (0.03)	0.001	-
NW reading s 13 (TOWRE)	-0.13 (0.03)	1.3x10^{-4}	-	-0.09 (0.01)	<1x10^{-10}	-	-0.20 (0.14)	0.16	-	-0.15 (0.03)	7.9x10^{-6}	-
Spelling a 7 (NB)	-0.17 (0.04)	1.5x10^{-5}	-	-0.09 (0.01)	<1x10^{-10}	-	-0.26 (0.16)	0.10	-	-0.16 (0.03)	1.4x10^{-8}	-
Spelling a 9 (NB)	-0.13 (0.04)	3.8x10^{-4}	-	-0.10 (0.01)	<1x10^{-10}	-	-0.29(0.15)	0.06	-	-0.13 (0.03)	4.8x10^{-6}	-
PhonAware 7 (AAT)	-0.13 (0.03)	2.7x10^{-4}	-	-0.09 (0.01)	<1x10^{-10}	-	-0.25 (0.14)	0.09	-	-0.13 (0.03)	3.9x10^{-6}	-
Listening c 8 (WOLD)	-0.19 (0.03)	6.6x10^{-8}	-	-0.08 (0.01)	<1x10^{-10}	-	0.04 (0.14)	0.76	-	-0.08 (0.03)	0.005	-
Non-word repetition 8 (CNRep)	-0.13 (0.04)	8.4x10^{-4}	-	-0.08 (0.01)	<1x10^{-10}	-	0.05 (0.16)	0.73	-	-0.04 (0.03)	0.15	-
VIQ 8 (WISC-III)	-0.24 (0.03)	1.9x10^{-10}	-	-0.13 (0.01)	<1x10^{-10}	-	-0.06 (0.15)	0.71	-	-0.18 (0.03)	5.7x10^{-9}	-
Pooled reading	-0.14 (0.03)	5.0x10^{-7}	0.08	-0.10 (0.01)	<1x10^{-10}	<0.001	-0.23 (0.12)	0.06	0.12	-0.14 (0.03)	1.2x10^{-7}	0.07
Pooled spelling	-0.15 (0.04)	2.9x10^{-5}	0.16	-0.10 (0.01)	<1x10^{-10}	0.34	-0.28 (0.14)	0.06	0.81	-0.15 (0.03)	1.6x10^{-7}	0.15
Pooled LRAs	-0.15 (0.02)	1.2x10^{-10}	0.02	-0.09 (0.01)	<1x10^{-10}	<0.001	-0.06 (0.09)	0.50	0.18	-0.11 (0.02)	4.0x10^{-7}	<0.001

Sets of conservative ($P<5x10^{-8}$) and subthreshold ($P<0.0015$) EA instruments were extracted from EA (SSGAC), ADHD (PGC+iPSYCH), and LRAs (ALSPAC) GWAS summary statistics. EA-specific effects independent of ADHD (β_{EA}) and EA effects shared with ADHD (β_{ADHD}) on LRAs were estimated with MVRs. To compare the magnitude of β_{ADHD} and β_{EA}, MVR analyses were conducted using standardised regression estimates (Supplementary Methods). β_{EA} estimates measure the change in LRA Z-scores per Z-score in missing school years. β_{ADHD} estimates measure the change in LRA Z-score per Z-score in ADHD liability. Pooled estimates for reading, spelling and global LRA measures (Table 1) were obtained through random-effects meta-regression. Effects were considered significant if they passed the experiment-wide significance threshold ($P<0.007$). 1. EA genetic effects shared with ADHD as assessed through ADHD genetic effect estimates of EA-associated variants. Abbreviations: LRAs, literacy- and language-related abilities; ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; P_{thr}, *P*-value threshold; P_{het}, Heterogeneity *P*-value; a, accuracy; c, comprehension; s, speed; WORD, Wechsler Objective Reading Dimension; NBO, Nunes, Bryant and Olson (ALSPAC specific instrument); NARA II, The Neale Analysis of Reading Ability- Second Revised British Edition; TOWRE, Test Of Word Reading Efficiency; NW, non-word; NB, Nunes and Bryant (ALSPAC specific instrument); PhonAware, phonemic awareness; AAT, Auditory Analysis Test; WOLD, Wechsler Objective Language Dimensions; CNRep, Children's Test of Nonword Repetition; VIQ, verbal intelligence quotient; WISC-III, Wechsler Intelligence Scale for Children III.; MVR - Multivariable regression

Table S12: Association between polygenic ADHD risk and sample-dropout

Outcome	ADHD sample	OR (SE)	*P*	McFadden's pseudo R^2
Sample drop-out for reading accuracy and comprehension at age 7 (WORD)	PGC	-0.995 (0.005)	0.36	$8.6x10^{-5}$
	iPSYCH	1.029 (0.005)	$5.0x10^{-9}$	$3.5x10^{-3}$
	PGC+iPSYCH	1.028 (0.005)	$1.4x10^{-8}$	$3.3x10^{-3}$

A yes/no variable indicating sample drop-out for reading accuracy and comprehension at age 7 (WORD) was generated. ADHD SNPs were selected based on a *P*-value threshold of 0.1, and alleles were aligned such that the effect allele increased ADHD risk. Sample drop-out was regressed on polygenic ADHD risk scores using logistic regression. The experiment-wide significance threshold is *P*<0.007. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; WORD, Wechsler Objective Reading Dimension; PGC, Psychiatric Genomics Consortium; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research

Table S13: Multivariable regression analysis of polygenic associations between ADHD, sample drop-out and literacy- and language-related abilities

	ADHD (β_{ADHD}) (ADHD-specific effects independent of EA)			EA (β_{EA}) (EA genetic effects of ADHD-associated variants)[1]		
ADHD instruments	**OR (SE)**	**β (SE)**	***P***	**OR (SE)**	**β (SE)**	***P***
Conservative (P_{thr}<$5x10^{-8}$)	0.99 (0.29)	-0.01 (0.30)	0.96	50.0 (133.7)	3.91 (2.68)	0.17
Subthreshold (P_{thr}<0.0015)	1.05 (0.02)	0.05 (0.01)	$3.7x10^{-4}$	2.83 (0.54)	1.04 (0.19)	$7.3x10^{-8}$

Sets of conservative (P<$5x10^{-8}$) and subthreshold (P<0.0015) ADHD instruments were extracted from ADHD (PGC+iPSYCH), EA (SSGAC) and sample drop-out for reading accuracy and comprehension at age 7 (WORD) (ALSPAC) GWAS summary statistics. ADHD-specific effects independent of EA (β_{ADHD}) and ADHD effects shared with EA (β_{EA}) on LRAs were estimated with MVRs. β_{ADHD} estimates measure the change in liability to drop-out per log odds increase in ADHD liability. β_{EA} estimates measure the change in liability to drop-out per missing school year. The experiment-wide significance threshold is *P*<0.007. 1. ADHD genetic effects shared with EA as assessed through EA genetic effect estimates of ADHD-associated variants. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; P_{thr}, *P*-value threshold; MVR - Multivariable regression

Supplementary Figures

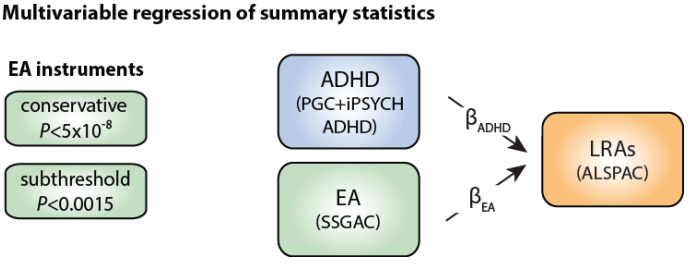


Figure S1: Multivariable regression analysis of polygenic associations between EA and literacy- and language-related abilities. Sets of conservative ($P<5x10^{-8}$) and subthreshold ($P<0.0015$) EA instruments were extracted from ADHD (PGC+iPSYCH), EA (SSGAC) and LRAs (ALSPAC) GWAS summary statistics. EA-specific effects independent of ADHD (β_{EA}) and EA effects shared with ADHD (β_{ADHD}) on LRAs were estimated with MVRs. EA effects shared with ADHD were assessed through ADHD genetic effect estimates of EA-associated variants. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; LRAs, literacy- and language-related abilities; MVR, Multivariable regression

Supplementary References

1. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
2. *WORD, Wechsler Objective Reading Dimensions Manual*. (Psychological Corporation, 1993).
3. Nunes, T., Bryant, P. & Olsson, J. Learning Morphological and Phonological Spelling Rules: An Intervention Study. *Sci. Stud. Read.* **7**, 289–307 (2003).
4. Schonell, F. J. & Goodacre, E. The Psychology and Teaching of Reading. Fifth Edition. (1974).
5. Neale, M. D. *Neale Analysis of Reading Ability: Second Revised British Edition*. (NFER-Nelson, 1997).
6. *TOWRE, Test of Word Reading Efficiency: Examiner's Manual*. (PRO-ED, 1999).
7. Rosner, J. & Simon, D. P. The Auditory Analysis Test: An Initial Report. *J. Learn. Disabil.* **4**, 384–392 (1971).
8. Rust, J. & Wechsler, D. *WOLD: Wechsler objective language dimensions.* (Psychological Corp. ; Artsberg Enterprises Ltd. [distributor, 1996).
9. Gathercole, S. E., Willis, C. S., Baddeley, A. D. & Emslie, H. The Children's Test of Nonword Repetition: a test of phonological working memory. *Mem. Hove Engl.* **2**, 103–127 (1994).
10. Wechsler, D., Golombok, S. & Rust, J. *WISC-III UK Wechsler Intelligence Scale for Children – Third Edition UK Manual*. (The Psychological Corporation, 1992).
11. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
12. Goodman, R., Ford, T., Richards, H., Gatward, R. & Meltzer, H. The Development and Well-Being Assessment: Description and Initial Validation of an Integrated Assessment of Child and Adolescent Psychopathology. *J. Child Psychol. Psychiatry* **41**, 645–655 (2000).
13. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. vol. fifth (American Psychiatric Association, 2013).
14. Wigginton, J. E., Cutler, D. J. & Abecasis, G. R. A Note on Exact Tests of Hardy-Weinberg Equilibrium. *Am. J. Hum. Genet.* **76**, 887–893 (2005).
15. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
16. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
17. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
18. Durbin, R. Efficient haplotype matching and storage using the positional Burrows-Wheeler transform (PBWT). *Bioinforma. Oxf. Engl.* **30**, 1266–1272 (2014).
19. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2007).
20. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
21. *International Statistical Classification of Diseases and Related Health Problems*. vol. 10th revision (World Health Organization, 2010).
22. Mors, O., Perto, G. P. & Mortensen, P. B. The Danish Psychiatric Central Research Register. *Scand. J. Public Health* **39**, 54–57 (2011).
23. Demontis, D. *et al.* Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv* 145581 (2017) doi:10.1101/145581.
24. Pedersen, C. B. *et al.* The iPSYCH2012 case–cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol. Psychiatry* **23**, 6–14 (2018).
25. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
26. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
27. Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340**, 1467–1471 (2013).
28. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–42 (2016).

29. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
30. Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J. & Rohde, L. A. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am. J. Psychiatry* **164**, 942–948 (2007).
31. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
32. International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
33. Wray, N. R. *et al.* Research Review: Polygenic methods and their application to psychiatric traits. *J. Child Psychol. Psychiatry* **55**, 1068–1087 (2014).
34. Howie, B. N., Donnelly, P. & Marchini, J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLOS Genet.* **5**, e1000529 (2009).
35. Palla, L. & Dudbridge, F. A Fast Method that Uses Polygenic Scores to Estimate the Variance Explained by Genome-wide Marker Panels and the Proportion of Variants Affecting a Trait. *Am. J. Hum. Genet.* **97**, 250–259 (2015).
36. Burgess, S. *et al.* Dissecting Causal Pathways Using Mendelian Randomization with Summarized Genetic Data: Application to Age at Menarche and Risk of Breast Cancer. *Genetics* genetics.300191.2017 (2017) doi:10.1534/genetics.117.300191.
37. Lajeunesse, M. J. Meta-Analysis and the Comparative Phylogenetic Method. *Am. Nat.* **174**, 369–381 (2009).
38. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 1–48 (2010).
39. Pooled Variance. in *The Concise Encyclopedia of Statistics* 427–428 (Springer New York, 2008). doi:10.1007/978-0-387-32833-1_323.
40. Lee, S. H., Wray, N. R., Goddard, M. E. & Visscher, P. M. Estimating Missing Heritability for Disease from Genome-wide Association Studies. *Am. J. Hum. Genet.* **88**, 294–305 (2011).
41. Martin, J. *et al.* Association of Genetic Risk for Schizophrenia With Nonparticipation Over Time in a Population-Based Cohort Study. *Am. J. Epidemiol.* kww009 (2016) doi:10.1093/aje/kww009.
42. Taylor, A. *et al.* The molecular genetics of participation in the Avon Longitudinal Study of Parents and Children. *bioRxiv* 206698 (2017) doi:10.1101/206698.
43. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).

Chapter 7

Shared risk alleles with discordant polygenic effects: disentangling the genetic overlap between ASD and ADHD

This chapter has been submitted and a previous version is available as preprint:
Verhoef, E., Grove, J., Shapland, C.Y., Demontis, D., Burgess, S., Rai, D., Børglum, A.D., St Pourcain, B. Shared risk alleles with discordant polygenic effects: disentangling the genetic overlap between ASD and ADHD (under review, doi: doi.org/10.1101/580365).

Abstract

Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD) are complex co-occurring neurodevelopmental conditions. Their genetic architectures reveal striking similarities but also differences, including strong, discordant polygenetic associations with educational attainment (EA). To study genetic mechanisms that present as ASD-related positive and ADHD-related negative genetic correlations with EA, we carried out multivariable regression analyses using genome-wide summary statistics based on 10,610 to 766,345 individuals. Our results showed that EA-related genetic variation is regionally shared across ASD and ADHD architectures, involving the same risk alleles at the same markers. However, the polygenic association profile with EA, across these shared markers, was discordant for ASD versus ADHD risk, indicating independent genetic effects. Thus, at the single-variant level, our results suggest either biological pleiotropy or co-localisation of different risk variants. At the polygenic level, they are consistent with local negative genetic covariance that may contribute to the total genome-wide correlation between ASD and ADHD.

Key words: ASD, ADHD, educational attainment, pleiotropy, shared genetic variation

7.1. Main

Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD) are genetically complex childhood-onset neurodevelopmental disorders[1,2] that often co-occur[3]. Approximately 15–25% of individuals with ADHD show ASD symptoms, and ~40–70% of individuals with ASD have a comorbid ADHD symptomatology[3], although knowledge of shared aetiological mechanisms is scarce.

Like many other complex psychiatric disorders, ASD and ADHD are highly polygenic, and the majority of genetic influences can be attributed to common genetic variation[4]. There is increasing evidence from twin and genome-wide association studies (GWAS)[5,6] suggesting genetic links between ASD and ADHD symptoms, both throughout population variation[7–13] and at the clinical level[14]. The largest and most recent cross-disorder GWAS analysis reported at least hundred loci (as tagged by single variants) with pleiotropic effects on more than one disorder, including ASD and ADHD[4]. A model of single-nucleotide polymorphism (SNP)-based genetic correlations among multiple psychiatric disorders, using exploratory factor analyses and genomic structural equation models, showed that both ASD and ADHD are part of the same cluster of early-onset neurodevelopmental disorders[4]. The existence of genetic links between these disorders is further strengthened by the familial co-aggregation of both clinical disorders in large register-based studies[15] and the identification of shared copy number variations, suggesting similar biological pathways[16].

Estimates of genetic correlations between ASD and ADHD diagnosis range between 0.36(95%-confidence interval (CI): 0.26-0.46)[17] in molecular studies to 0.87(95%-CI: 0.77-1.0)[18] in twin analyses[19]. Evidence for genetic links between ASD and ADHD symptom co-occurrence can even be stronger in population-based samples[8]. However, when both, clinical ASD and clinical ADHD, are investigated with respect to a third genetically complex trait, also differences in the genetic architecture become apparent. Each disorder, when predicted with GWAS variants, reveals an opposite genetic correlation with cognitive functioning and educational attainment (EA). While increased polygenic ADHD risk has been linked to lower cognitive abilities and EA[20,21], increased polygenic ASD risk has been associated with higher cognitive functionality and EA[17,20,22]. This discordant association pattern is most discernible for measures of years-of-schooling and college-completion[17,21]. Observational research in ADHD strongly confirms the associations with lower school performance and educational outcomes[23]. Reports of academic achievement in ASD are more variable[24], although high-functioning individuals can obtain higher-order qualifications, despite disadvantages in the labour market[25].

The mechanisms underlying the discordant polygenic association pattern with EA are not yet known and may involve different biological effects, including pleiotropy. Following Solofiev and colleagues[26], we define biological pleiotropy as processes where

the same gene has a direct biological influence on more than one phenotype. In contrast, spurious pleiotropy involves multiple sources of bias that cause a false association between a gene and multiple phenotypes[26]. Different causal risk variants in high linkage disequilibrium (LD) with the same marker are described as co-localising variants[26]. We do not consider mechanisms of mediated pleiotropy, i.e. an indirect association between a genetic variant and a further phenotype that arises due to causal associations between phenotypes[26], as this mechanism would imply concordant associations between EA and both ASD and ADHD risk. An overview of hypothetical non-pleiotropic and pleiotropic mechanisms is shown in Figure 1.

First (scenario I), the set of underlying causal variants linking ASD to EA is different from the set of underlying causal variants linking ADHD to EA. Here, causal alleles are tagged by different GWAS markers, residing either within low LD regions within the same gene locus (biological pleiotropy, Figure 1a) or at different loci (no pleiotropy, Figure 1b). Consequently, they could give rise to different polygenic associations with EA. Second (scenario II), association with EA is introduced because of ascertainment bias during the recruitment of ASD and ADHD cases (Figure 1c). In the US, the prevalence of ASD has been associated with higher parental socio-economic status[27]. In a large population-based Swedish study, also an association between lower SES and higher ASD risk[28] has been observed. In contrast, children in low SES families are consistently more likely to receive a diagnosis of ADHD than children in high SES families[29]. Third (scenario III), different causal ASD and ADHD risk alleles are tagged by different GWAS marker alleles, either at the same variant within the same gene locus (biological pleiotropy, Figure 1d), at different causal variants within the same locus (biological pleiotropy, co-localising variants, Figure 1e) or at different loci (spurious pleiotropy, co-localising variants , Figure 1f). The most recent cross-disorder GWAS analysis identified several loci with opposite directional allelic effects at the 10^{-6} *P*-value threshold. None of these loci were shared between ASD and ADHD[4], though these effects may become more prevalent when applying less stringent GWAS marker selection criteria. Fourth (scenario IV), different causal risk alleles for ASD and ADHD are tagged by the same GWAS marker allele, given LD[26], either at the same locus (biological pleiotropy, co-localising variants, Figure 1g) or at different loci (spurious pleiotropy, co-localising variants, Figure 1h). The vast majority of trait-associated loci across the genome overlaps with loci from multiple traits and each physical location can contain multiple groups of variants with independent genetic effects[30,31]. Hence, the distribution of genetic effects across the same GWAS marker alleles could differ for ASD and ADHD summary statistics and shape polygenic associations with a third phenotype, such as EA. Finally (scenario V), identical causal risk alleles are shared between ASD and ADHD at the same variant and locus, and exert, because of pleiotropy, different genetic effects leading to different polygenic associations with EA (Figure 1i). Note that in the absence of pleiotropy, identical ASD and

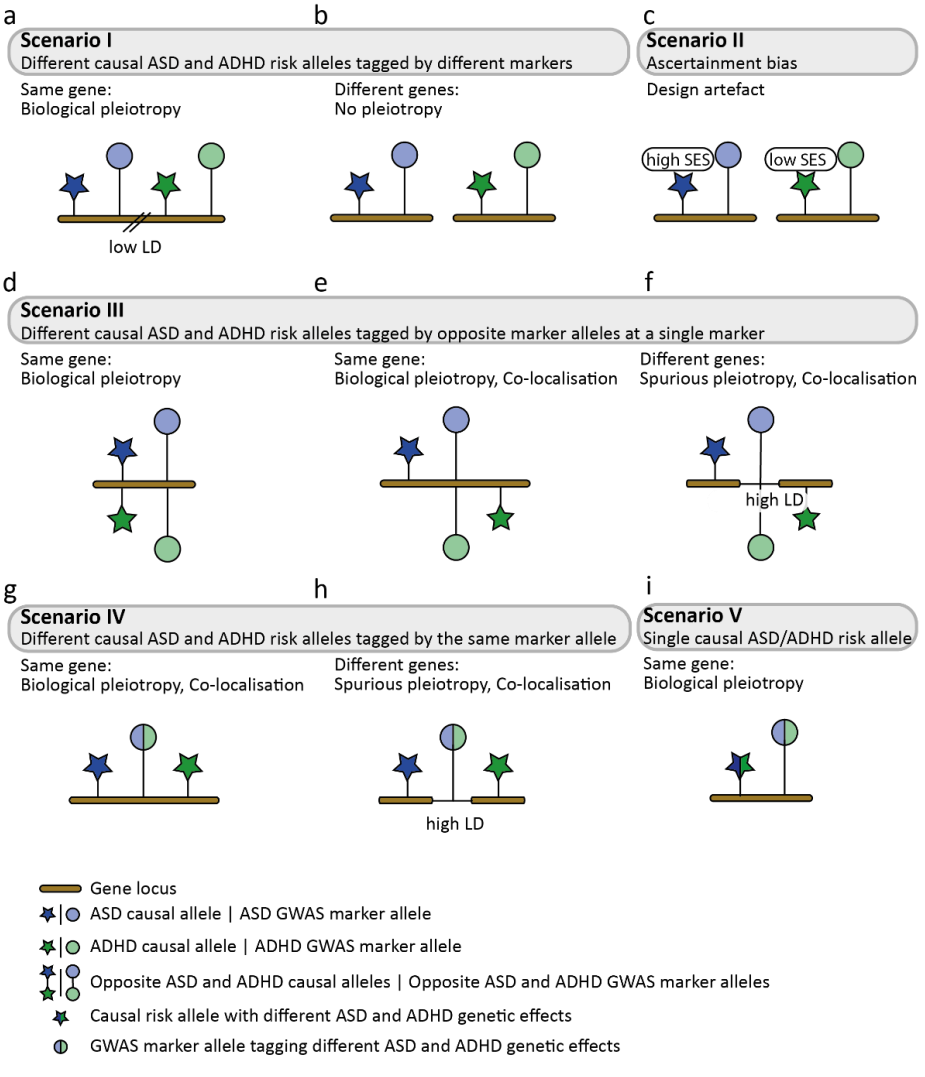


Figure 1: Non-pleiotropic and pleiotropic mechanisms. Discordant associations with EA for ASD versus ADHD risk may involve different non-pleiotropic or pleiotropic mechanisms **(a-i)**. Scenario I: Different causal ASD and ADHD risk variants are tagged by different markers, **(a)** at the same gene locus in regions with low linkage disequilibrium (LD) or **(b)** at different loci. **(c)** Scenario II: Ascertainment bias during the recruitment of cases leads to an artificial association of ASD with higher socio-economic status (SES) and ADHD with lower SES (non-testable). Scenario III: Different causal ASD and ADHD risk alleles are captured by opposite alleles at a single marker tagging **(d)** a single risk variant, **(e)** different risk variants at the same gene or **(f)** different risk variants at different genes in regions of high LD. Scenario IV: Different causal ASD and ADHD risk alleles are captured by the same marker allele tagging **(g)** different risk variants at the same gene or **(h)** different risk variants at different genes in regions of high LD. **(i)** Scenario V: A single causal risk allele is tagged by the same marker allele and exerts different ASD and ADHD genetic effects due to biological pleiotropy. Within each subfigure, one or more observed marker allele is shown in linkage disequilibrium with one or more causal risk allele for ASD and ADHD risk. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder, EA, educational attainment.

ADHD risk alleles captured by the same marker would lead to a concordant and not discordant polygenic association with EA.

In this work, we (1) study evidence for genetic mechanisms presenting as discordant polygenic association pattern between EA and both ASD and ADHD risk, (2) characterise ASD and ADHD risk variants that contribute to the association with EA and (3) assess the specificity of identified genetic mechanisms by investigating genetic overlap with EA for other disorders. We finally integrate our findings with knowledge of genetic correlations between ASD and ADHD.

To evaluate evidence for genetic mechanisms that may result in a discordant association pattern with EA (Figure 1), we model polygenic relationships through individual genetic variants, controlling both for the position of the studied SNPs and the direction of the genetic effects at the single-marker level. The simultaneous investigation of ASD and ADHD risk alleles facilitates the identification of multiple independent polygenic associations with EA and their direction of effect, as encoded by the same GWAS markers. Furthermore, we can estimate the extent of variance inflation due to correlations among independent variables (multicollinearity).

To this end, we studied SNP estimates from existing GWAS summary statistics for EA, ASD and ADHD (Table 1) with bidirectional multivariable regression (MVR) analyses (Figure 2), using a multivariate methodology borrowed from a causal modelling approach[32]. Without making causal inferences, as we allow for biological pleiotropy, this method can estimate polygenic associations conditional on each other, while controlling for potential bias that may arise when adjusting for heritable covariates[33]. The selection of variant sets follows guidelines established for polygenic scoring methods[34], but without generating accumulated allele risk scores[34]. Thus, we describe here, due to the polygenic context, genetic associations only. We assess evidence in support of genetic mechanisms as outlined in Figure 1, except evidence for ascertainment bias (scenario II).

Table 1: Sample description

Source	Phenotype	Consortium	GWAS	Imputation reference panel	N
Clinical sample	ASD	iPSYCH	ASD(iPSYCH)[17]	1000 Genomes phase 3	35,740 (13,076 cases)
			ASD(iPSYCH,woADHD)[17]	1000 Genomes phase 3	32,985 (10,321 cases)
		PGC	ASD(PGC)[35]	1000 Genomes phase 1 (v3)	10,610 (5,305 cases)
	ADHD	iPSYCH	ADHD[21]	1000 Genomes phase 3	37,076 (14,584 cases)
Population sample	Years-of-schooling	SSGAC	EA[49]	1000 Genomes phase 3[#]	766,345

All individuals were of European descent. [#.] Predominantly 1000 genomes phase 3, see Lee et al.[49]. Abbreviations: ASD, Autism Spectrum Disorder; ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; PGC, Psychiatric Genomics Consortium; SSGAC, Social Science Genetic Consortium; woADHD; without ADHD.

7.2. Results

Analyses of genetic mechanisms underlying discordant polygenic association patterns with educational attainment

Applying a MVR analysis framework, we investigate polygenic associations with EA by studying evidence for ASD-specific (Figure 2a) and ADHD-specific association effects (Figure 2b), in addition to genetic association effects that are shared, by genetic marker position, across both conditions (termed here forth cross-disorder genetic effects). Here, we study evidence for cross-disorder associations bi-directionally, as shown in Figure 2, using two MVR model designs:

(1) ASD-MVR models (Figure 2a) estimate ASD-specific associations with EA ($\hat{\theta}_{ASD}$), based on individual ASD SNP estimates ($\hat{\beta}_{ASD}$), and ADHD cross-disorder associations ($\hat{\theta}_{\maltese ADHD}$) with EA, based on individual ADHD SNP estimates ($\hat{\beta}_{ADHD}$), as encoded by the same set of independent variants that were selected because of their subthreshold association with ASD risk (Gi).

(2) ADHD-MVR models (Figure 2b) estimate ADHD-specific associations with EA ($\hat{\theta}_{ADHD}$), using individual ADHD SNP estimates ($\hat{\beta}_{ADHD}$), and ASD cross-disorder associations with EA ($\hat{\theta}_{\maltese ASD}$), using individual ASD SNP estimates ($\hat{\beta}_{ASD}$), as tagged by the same set of independent variants that were selected because of their subthreshold association with ADHD risk (Gj).

These complementary models, and modifications thereof, were implemented into multiple stages of the study design (Supplementary Figure 1). Using GWAS summary statistics with independent cases from large consortia (Table 1), we apply a weighted regression framework (Methods, Formulae 1-4), analogous to Mendelian Randomization (MR) approaches[32], but with markers as selected for polygenic scoring[34] to assess concurrent polygenic ASD and ADHD risk associations with EA.

During the discovery stage, we analysed 11 ASD-MVRs and 11 ADHD-MVRs using a series of variant sets selected at different *P*-value selection thresholds ($5x10^{-8}<P_{thr}<0.5$, Supplementary Figure 1a-b) of which findings at $P_{thr}<0.0015$ and $P_{thr}<0.05$ are shown in Figure 2c. These analyses provided evidence for ASD-specific, ADHD-specific and cross-disorder associations with EA, across multiple *P*-value thresholds (Supplementary Table 6-7). For example, for ASD-MVR at $P_{thr}<0.0015$ (N_{SNPs}=1,973, Figure 2a,c,d, Supplementary Table 8), we observed a positive ASD-specific association with EA, with an 0.009 increase in years-of-schooling per log-odds in ASD liability (ASD-MVR $\hat{\theta}_{ASD}$=0.009(SE=0.003), *P*=0.002).

Conditionally, the same ASD-related risk alleles captured a negative association between ADHD and EA with a 0.029 decrease in years-of-schooling per log-odds in ADHD liability (ASD-MVR $\hat{\theta}_{\maltese ADHD}$=-0.029(SE=0.004), $P<1x10^{-10}$). Thus, ADHD cross-disorder associations showed an opposite direction of effect compared to ASD-specific

associations (Figure 2c,d), even though they were modelled with respect to same, here ASD-related, risk alleles.

An analogous approach with ADHD-MVRs (Supplementary Figure 1b) revealed a complementary association profile (Supplementary Table 7). There was a negative ADHD-specific association between ADHD risk and EA. Conditionally, ASD cross-disorder associations with EA were positive, thus discordant, even though they were modelled with the same ADHD-related risk alleles (Figure 2c,e). For ADHD-MVR at P_{thr}<0.0015 (N_{SNPs}=2,717, Figure 2b,c,e, Supplementary Table 8), this corresponds to an 0.012 decrease in years-of-schooling per log-odds in ADHD liability (ADHD-MVR $\hat{\theta}_{ADHD}$=-0.012(SE=0.003), P=4x10^{-5}), and an increase in 0.022 years-of-schooling per log-odds in ASD liability (ADHD-MVR $\hat{\theta}_{ASD}$=0.022(SE=0.003), P<1x10^{-10}). Increasing the number of variants in ASD-MVRs and ADHD-MVRs using more relaxed selection thresholds (e.g. P_{thr}<0.05) boosted the statistical power (Figure 2c, Supplementary Table 6-8).

Compared to univariable regression (UVR) models (modelling genetic associations between a single disorder and EA only), the simultaneous estimation of ASD and ADHD effects on EA using MVR improved the model fit (Supplementary Table 8). Multivariable models explained up to 3% more variation in genetically predictable EA compared to univariable models, with modest evidence for multi-collinearity (Supplementary Table 8, Variance inflation factor (VIF)≤1.2). Hence, the concurrent identification of disorder-specific and cross-disorder MVR effects suggests that different causal ASD and ADHD risk effects are encoded at the same genetic markers (scenario III, IV or V) and may, consequently, result in different genetic associations with EA (Figure 1d-i).

If genetic markers are shared between ASD and ADHD genetic architectures, different causal ASD and ADHD effects can be encoded with respect to the same (scenario IV and V, Figure 1g-i) or opposite alleles (scenario III, Figure 1d-f) at a single GWAS marker. Within a next step, we therefore restricted discovery ASD and ADHD variant sets to markers carrying the same risk-increasing allele for both disorders (termed here forth concordant variants; ~80% of the initial sets, Supplementary Figure 1c,d) and studied the robustness of discovery MVR signals. These follow-up analyses with ASD and ADHD concordant variants confirmed the discordant association patterns with EA, with little evidence for attenuation (Supplementary Table 9). The corresponding bivariate relationships between SNP estimates for ASD, ADHD and EA are displayed in Supplementary Figure 2 and illustrate both the positive correlation of ASD ($\hat{\beta}_{ASD}$) and ADHD ($\hat{\beta}_{ADHD}$) risk allele effects and their opposite association with EA. Thus, MVR analyses with concordant variants demonstrate that discordant genetic associations with EA can arise independent of the allelic alignment to ASD or ADHD risk. These findings are inconsistent with scenario III (Figure 1d-f) and suggest that different ASD and ADHD causal alleles are tagged by the same GWAS marker allele (scenario IV, Figure g-h) or that identical causal risk alleles exert different ASD and ADHD genetic effects due to biological

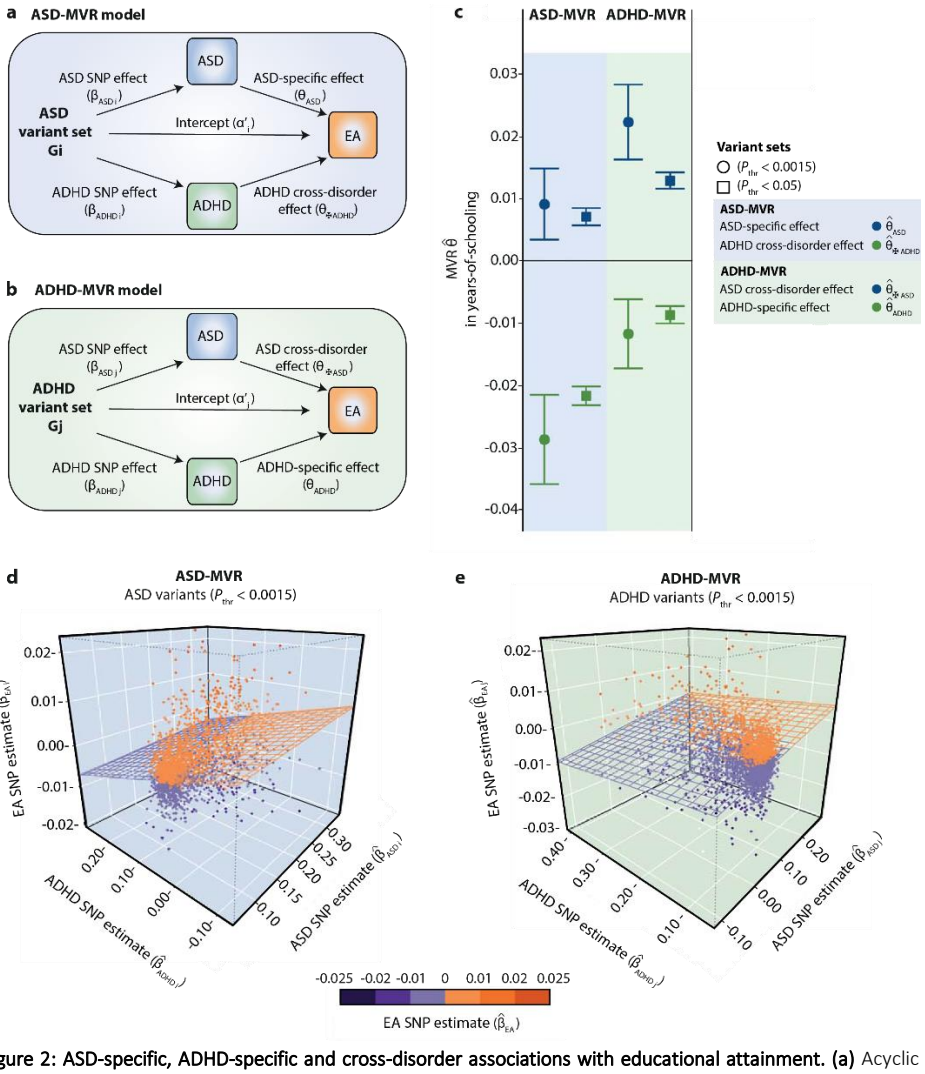


Figure 2: ASD-specific, ADHD-specific and cross-disorder associations with educational attainment. (a) Acyclic graph illustrating multivariable regression (MVR) for a set of ASD variants Gi (ASD-MVR), two independent variables (ASD and ADHD risk) and the dependent variable EA. The genetic association effect of Gi on ASD and ADHD risk is β_{ASDi} and β_{ADHDi} respectively. The genetic association effect of ASD risk on EA is the ASD-specific effect θ_{ASD}. The genetic association effect of ADHD risk on EA is the ADHD cross-disorder effect $\theta_{\maltese ADHD}$. The intercept α'_{ASD} represents the direct effect of ASD variants Gi on EA that are neither captured by θ_{ASD} nor $\theta_{\maltese ADHD}$. **(b)** Analogous acyclic graph illustrating MVR for a set of ADHD variants Gj (ADHD-MVR), two independent variables (ASD and ADHD risk) and the dependent variable EA. The genetic association effect of Gj on ASD and ADHD risk is β_{ASDj} and β_{ADHDj} respectively. The genetic association effect of ASD risk on EA is the ASD cross-disorder effect $\theta_{\maltese ASD}$. The genetic association effect of ADHD risk on EA is the ADHD-specific effect θ_{ADHD}. The intercept α'_j represents the direct effect of ADHD variants G_j on EA that are neither captured by θ_{ADHD} nor $\theta_{\maltese ASD}$. **(c)** Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and ADHD cross-disorder effect $\hat{\theta}_{\maltese ADHD}$ as fitted with ASD-MVR (a) and estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effects $\hat{\theta}_{\maltese ASD}$ as fitted with ADHD-MVR (b). Sets of independent ASD (Gi) and ADHD (Gj) genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively and are shown for two *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH, woADHD),

ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. Bars represent 95% confidence intervals. All estimated MVR effects $\hat{\theta}$ passed the multiple testing threshold of P<0.0023. **(d)** 3D scatter plot of ASD SNP estimates ($\hat{\beta}_{ASD}$(lnOR) , x-axis), ADHD SNP estimates ($\hat{\beta}_{ADHD}$(lnOR), y-axis) and EA SNP estimates ($\hat{\beta}_{EA}$, z-axis) for ASD-related variants Gi(P_{thr}<0.0015), as analysed with ASD-MVR (c). The regression plane reflects the estimated ASD-specific ($\hat{\theta}_{ASD}$) and ADHD cross-disorder ($\hat{\theta}_{\pm ADHD}$) effects. **(e)** 3D scatter plot of ASD SNP estimates ($\hat{\beta}_{ASD}$(lnOR), x-axis), ADHD SNP estimates ($\hat{\beta}_{ADHD}$(lnOR), y-axis) and EA SNP estimates ($\hat{\beta}_{EA}$, z-axis) for ADHD-related variants Gj (P_{thr}<0.0015), as analysed with ADHD-MVR (c). The regression plane reflects the ADHD-specific ($\hat{\theta}_{ADHD}$) and ASD cross-disorder ($\hat{\theta}_{\pm ASD}$) effects. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; EA, educational attainment; MVR, multivariable regression; P_{thr}, P-value threshold.

pleiotropy (scenario V, Figure 1i). Using these fully comparable models, we could furthermore confirm the increase in strength and size of MVR disorder-specific and cross-disorder effects compared to the respective UVR effects (Supplementary Table 9). This suggests not only the presence of independent causal ASD and ADHD effects across the same GWAS marker alleles, but a destratification of ASD and ADHD risk effects as estimated with ASD(iPSYCH,woADHD) and ADHD(iPSYCH).

To replicate MVR findings, we replaced ASD SNP estimates ($\hat{\beta}_{ASD}$) from ASD(iPSYCH,woADHD) with SNP estimates derived from the independent ASD(PGC) sample (Supplementary Figure 1e-f). We repeated ASD-MVRs and ADHD-MVRs with the same sets of selected variants as studied in the discovery MVR analyses (Supplementary Figure 1a-b) and confirmed the discordant genetic association pattern with EA at the relaxed P-value threshold (P_{thr}<0.05; Supplementary Table 10). Here, ADHD subthreshold associated risk alleles carrying ASD SNP estimates show positive associations with EA (ADHD-MVR at P_{thr}<0.05: $\hat{\theta}_{\pm ASD}$=0.003(SE=4x10^{-4}), P<1x10^{-10}), despite zero genetic correlations between ADHD(iPSYCH) and ASD(PGC)(Supplementary Table 3). As a validation step, we confirmed the positive association with EA for ASD risk as captured by ASD(PGC) SNP estimates (ASD-MVR at P_{thr}<0.05: $\hat{\theta}_{ASD}$=0.005(SE=0.001), P<1x10^{-10}). At the more stringent threshold (P_{thr}<0.0015), only ASD-specific effects passed the multiple testing threshold (ASD-MVR at P_{thr}<0.0015: $\hat{\theta}_{ASD}$=0.01(SE=0.003), P<1x10^{-10}). This is consistent with the limited power of ASD(PGC)[35] and the lesser pool of aligned risk alleles as part of the harmonisation of ASD(iPSYCH,woADHD), ADHD(iPSYCH) and ASD(PGC) SNP estimates. Risk allele effect concordance in ASD(iPSYCH,woADHD) versus ASD(PGC) ranged only between 49% to 52%, similar to the risk allele distributions in ADHD(iPSYCH) versus ASD(PGC)(51% to 52%; Supplementary Table 10). The study of fully independent SNP estimates for ASD, ADHD and EA also increased the stability of disorder-specific and cross-disorder effects in UVRs compared to MVRs, with little evidence for destratification (Supplementary Table 10).

We finally investigated whether MVR findings for EA extend to general intelligence and reading (Supplementary Figure 1g-h), using summary statistics from intelligence(CTG) and reading. Discordant cross-disorder association patterns with intelligence were confirmed for both ASD and ADHD risk (Supplementary Table 11),

although ASD-specific effects at P_{thr}<0.0015 were only detectable as trend (ASD-MVR at P_{thr}<0.0015: $\hat{\theta}_{ASD}$=0.008(SE=0.004), P=0.05; $\hat{\theta}_{\text{Ŧ}ADHD}$=-0.033(SE=0.005), P=5x10^{-10}; ADHD-MVR at P_{thr}<0.0015: $\hat{\theta}_{ADHD}$=-0.012(SE=0.004), P=0.002; $\hat{\theta}_{\text{Ŧ}ASD}$=0.016(SE=0.004), P=2x10^{-4}). Discordant cross-disorder association patterns with reading were observed for ASD (Gi) and ADHD (Gj) variant sets selected at P_{thr}<0.05, with cross-disorder effects identified in ADHD-MVR only detectable as trend (Supplementary Table 12). These association effects agree with corresponding Linkage Disequilibrium Score (LDSC) genetic correlations, although there was no evidence for genetic links between reading and iPSYCH ASD samples (Supplementary Table 5).

Thus, our findings, validated by the use of different GWAS summary statistics, suggest that discordant genetic association patterns with EA are encoded at the same marker alleles for both ASD and ADHD risk (scenario IV or V).

Identification of cross-disorder loci

To identify and characterise variants that drive the observed discordant association patterns with EA, we systematically assessed the overlap between ASD (Gi) and ADHD (Gj) variant sets, based on iPSYCH samples. We created subsets of markers, which were associated with both ASD and ADHD risk and fulfilled joint ASD and ADHD selection criteria (Supplementary Figure 1i-j). Starting with an ASD variant set (Gi) selected at P_{thr}<0.0015 (N_{SNPs}≤1,973), we filtered markers for joint association with ADHD risk across multiple ADHD P-value thresholds (0.0015≤P_{thr}<0.5), creating six further subsets (Gij, Figure 3a). These sets captured 4.2% (ASD P_{thr}<0.0015 and ADHD P_{thr}<0.0015) to 46.3% (ASD P_{thr}<0.0015 and ADHD P_{thr}<0.5) of the discovery ASD variant set (Gi). Vice versa, applying similar selection criteria, we created six further ADHD variant subsets (Gji) based on the joint association with ASD risk (Figure 3b). These sets represented 3.1% (ASD P_{thr}<0.0015 and ADHD P_{thr}<0.0015) to 37.5% (ASD P_{thr}<0.5 and ADHD P_{thr}<0.0015) of the discovery ADHD variant set (Gj).

Fitting additional MVRs with these 12 jointly selected ASD and ADHD risk variants increased the size of observed genetic association effects with EA up to 5 times, compared to the effects identified with risk alleles selected for one disorder alone (Figure 3c, Supplementary Table 13-14). For example, selecting markers at P_{thr}<0.0015 for both ASD and ADHD risk (N_{SNPs}=83, 4.2% of the discovery ASD variant set (Gi), Figure 3c, Supplementary Table 13), ASD-MVR identified ASD-specific effects of 0.15(SE=0.025)(P=1x10^{-7}) and ADHD cross-disorder effects of -0.15(SE=0.025)(P=1x10^{-7}) years-of-schooling, per log-odds in ASD and ADHD liability respectively. In comparison, using the discovery ASD variant set (Gi) at P_{thr}<0.0015 (N_{SNPs}=1,973, Figure 3c, Supplementary Table 13), ASD-MVRs estimated ASD-specific effects of 0.009(SE=0.003)(P=0.002) and ADHD cross-disorder effects of -0.029(0.004)(P<1x10^{-10}) years-of-schooling, per log-odds in ASD and ADHD liability respectively. Findings for ADHD-MVR were highly similar (Figure 3c, Supplementary Table 14).

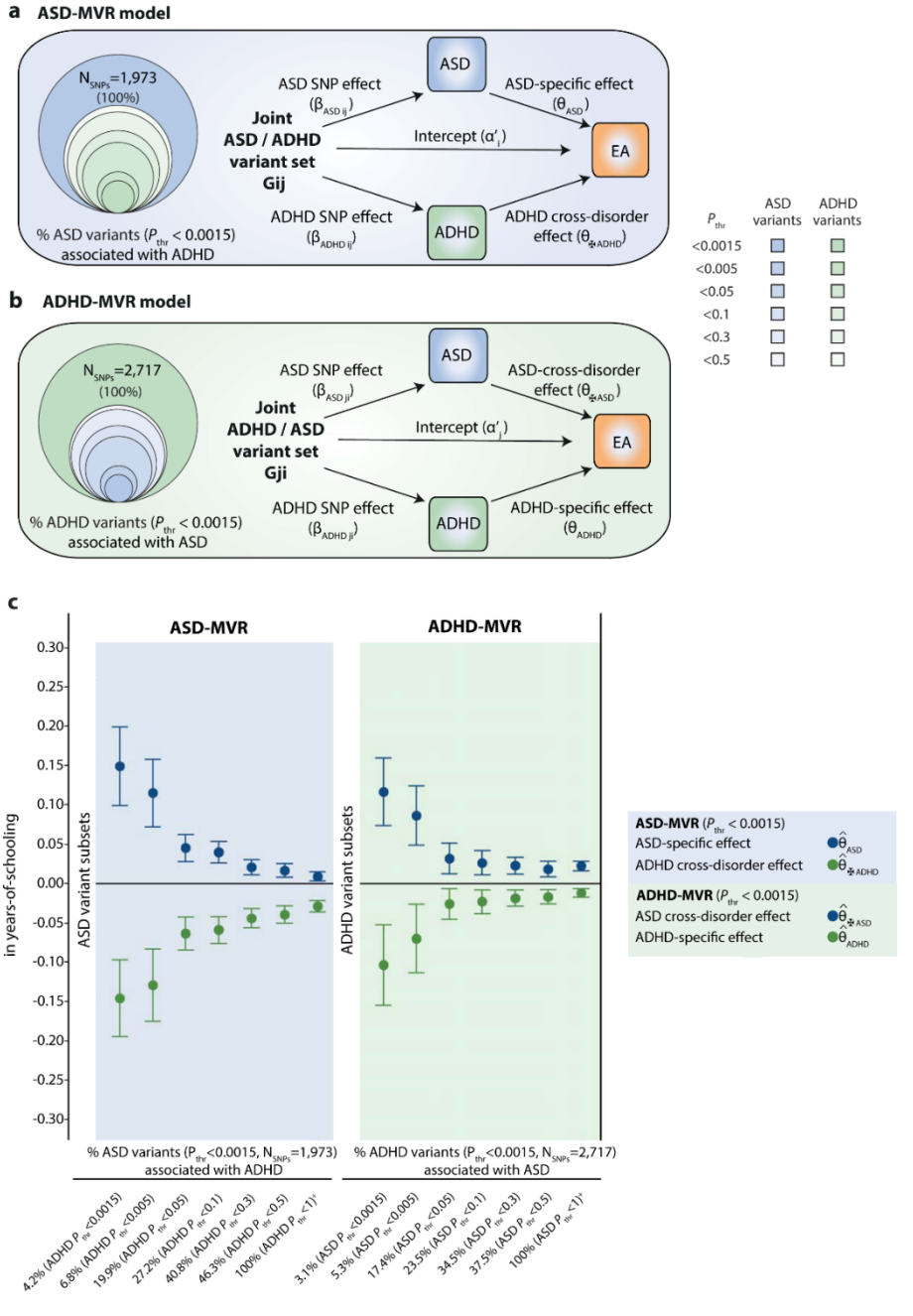


Figure 3: Identification of cross-disorder loci. (a) Acyclic graph illustrating ASD-MVR for subsets of ASD variants fulfilling joint ASD and ADHD selection criteria (Gij), two independent variables (ASD and ADHD risk) and the dependent variable EA, as shown in Figure 2a. The percentage of ASD variants (P_{thr}<0.0015) also associated with ADHD across multiple *P*-value selection thresholds (P_{thr}: 0.0015; 0.005; 0.05; 0.1; 0.3; 0.5) is represented through concentric circles. **(b)** Analogous acyclic graph illustrating ADHD-MVR for subsets of ADHD variants fulfilling joint ASD and ADHD selection criteria (Gji), two independent variables (ASD and ADHD risk) and the

dependent variable EA, as shown in Figure 2b. The percentage of ADHD variants (P_{thr}<0.0015) also associated with ASD across multiple *P*-value selection thresholds (6 thresholds: P_{thr}: 0.0015; 0.005; 0.05; 0.1; 0.3; 0.5) is represented through concentric circles. **(c)** Estimated ASD-specific effects $\hat{\theta}_{ASD}$ and ADHD cross-disorder effects $\hat{\theta}_{\pm ADHD}$ as fitted with ASD-MVR (a) and estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effects $\hat{\theta}_{\pm ASD}$ as fitted with ADHD-MVR (b) using SNP subsets Gij and Gji fulfilling joint ASD and ADHD selection criteria. Variant sets Gij and Gji are shown in % of the full (+) ASD (Gi) and ADHD (Gj) variant set respectively and consist of independent genetic variants from ASD(iPSYCH,woADHD) and ADHD(iPSYCH) GWAS statistics. SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH,woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. Bars represent 95% confidence intervals. All MVR effects passed the multiple testing threshold of P<0.0023, except ADHD-specific effects for ADHD-MVR (ADHD P_{thr} < 0.0015; ASD P_{thr} < 0.05, 0.1), which were present as trend (*P*=0.01). Bars represent 95% confidence intervals. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; MVR, multivariable regression; P_{thr}, P-value threshold.

The largest association effects with EA were identified with the most stringently selected ASD and ADHD risk variants, meeting a joint selection threshold of P_{thr}<0.0015 (Gij and Gji, Figure 3c). Irrespective of whether the selection process was started with a discovery ASD (Gi) or ADHD (Gj) variant set, the jointly selected variant sets (Gij and Gji) comprised the same 83 loci, based on identical or tagged proxy SNPs (LD-r^2=0.6, 500 kb window), of which 99% carried the same risk-increasing allele for both disorders (Supplementary Table 15). This combination of risk alleles and effects (selected at P_{thr}<0.0015 for both disorders) is unlikely to arise due to chance, as shown by permutations (Supplementary Table 16, empirical *P*<3x10^{-4}), and suggests locus specificity. The 83 variants mapped to at least 52 genes (RefSeq genes, Build37), including multiple regulatory RNAs such as microRNAs and lnc RNAs (Supplementary Table 15). A gene-set enrichment analysis using FUMA[36] software found enrichment for microRNA targets when screened against brain-expressed genes according to the BrainSpan samples[37]. This includes MIR19A/19B targets (False Discovery Rate (FDR)-adjusted *P*-value= 7.3x10^{-4}) at genes such as *CACNAC1* and *ERBB4*, as well as MIR9 targets (FDR-adjusted *P*-value= 2.8x10^{-2})(Supplementary Figure 3), although there was little evidence for a developmental specificity in expression patterns (data not shown).

Cross-disorder genetic associations for other disorders

Lastly, to assess the specificity of discordant cross-disorder effects with EA we also studied SNP estimates for other neuropsychiatric disorders (Supplementary Figure 1k-l). For this, we assigned SNP estimates derived from GWAS summary statistics for Major Depressive disorder (MDD), schizophrenia (SCZ) and Bipolar disorder (BD; Supplementary Table 1) to ASD (Gi) and ADHD (Gj) risk alleles, as defined for the discovery MVR analyses (P_{thr}<0.0015 and P_{thr}<0.05, Supplementary Figure 4a-b).

We identified further evidence for cross-disorder associations with EA, which was strengthened when studying variants at P_{thr}<0.05 (Supplementary Figure 4,

Supplementary Table 17-18) and associations were consistent with LDSC genetic correlations (Supplementary Table 3-4). Note that genetic association effects with EA for the studied adult-onset disorders might be inflated due to sample overlap between EA and MDD, SCZ and BD summary statistics (Methods). Still, cross-disorder effects with EA were weaker and smaller for adult-onset disorders compared to ASD and ADHD cross-disorder effects with EA in the discovery analyses (Figure 2c, Supplementary Figure 4c), and evidence for SCZ cross-disorder effects in ADHD-MVR did not pass the multiple-testing threshold. However, effects were comparable, in magnitude, to ASD cross-disorder effects in the follow-up analyses using ASD(PGC) SNP estimates (P_{thr}<0.05, Supplementary Table 10). Furthermore, unidirectionally, discordant association patterns with EA were detected for ASD in combination with MDD risk (ASD-MVR at P_{thr}<0.05: $\hat{\theta}_{\text{MDD}}$=-0.012, SE=0.001, P<1x10^{-10}, Supplementary Table 17), and for ADHD in combination with BD risk (ADHD-MVR at P_{thr}<0.05: $\hat{\theta}_{\text{BD}}$=0.008, SE=0.001, P<1x10^{-10}, Supplementary Table 18). Harmonised SNP effect concordance rates across disorders ranged between 51%-68%.

7.3. Discussion

Using a multivariate analysis approach, we investigated genetic mechanisms embedded in ASD and ADHD genetic architectures that present as discordant polygenic association pattern with EA. We found strong evidence that EA-related genetic variation is shared across ASD and ADHD architectures, consistent with pleiotropy or co-localisation of genetic effects at the same subthreshold ASD- or ADHD-risk associated marker alleles.

Genetic mechanisms for discordant polygenic association patterns with EA

ASD and ADHD genetic risk effects encoded at subthreshold GWAS marker alleles, selected for association with either disorder, were found to predict discordant, and thus fully independent, association patterns with EA. These patterns remained robustly detectable even when GWAS markers with alleles conferring opposite directional ASD and ADHD risk effects were excluded. Despite positive genetic correlations between ASD and ADHD risk effects at the single-variant level, combinations of the same risk-increasing alleles, carrying either ASD or ADHD risk effects, showed different associations with EA, and predicted either ASD-related positive or ADHD-related negative associations. The pattern of ASD- and ADHD-specific associations with EA, in combination with discordant genetic cross-disorder links, was (i) reciprocally detectable using either ASD- or ADHD-associated variant sets as selected for polygenic scoring approaches[36], (ii) replicated at P_{thr}<0.05 using ASD(PGC) summary statistics,(iii) consistent with the

previously reported genetic overlap between EA, ASD and ADHD[17,21] and (iv) independent of the harmonisation of GWAS marker alleles according to ASD or ADHD risk.

At the single-variant level, our findings are consistent with mechanisms assuming different ASD and ADHD causal risk alleles that are tagged by the same marker allele, either at the same locus (biological pleiotropy, co-localisation, scenario IV, Figure 1g) or at different loci (spurious pleiotropy, co-localisation)(scenario IV, Figure 1h) and mechanisms that involve identical ASD and ADHD risk alleles and biological pleiotropy, including special cases such as GxE[38] (scenario V, Figure 1i). Our findings are at odds with mechanisms proposing different ASD and ADHD risk alleles that are encoded at independent GWAS markers (scenario I) and mechanisms implicating opposite allelic effects at the same genetic marker (scenario III). Despite positive genetic correlations between harmonised ASD and ADHD risk effects, our findings can also not be explained by identical risk alleles, e.g. through mediated pleiotropy, which would result in concordant associations with EA.

At the level of polygenic inheritance, our findings are consistent with reports of local genetic covariance, which is predominantly, but not exclusively, positive[39]. For example, partitioning genetic covariance of high-density lipoprotein (HDL) and low-density lipoprotein LDL identified at least 11 loci with local genetic covariance, including negative covariance[39]. Here, the presence of discordant genetic cross-disorder associations with EA across the same risk alleles suggests local negative genetic covariance patterns that implicate not few but thousands of genetic risk variants associated with EA. This implies either wide-spread pleiotropy or co-localisation across ASD and ADHD genetic architectures, as well as extensive regional similarity. Given co-localisation, the underlying causal ASD and ADHD risk alleles can be fully independent of each other, despite high genetic effect correlations between genetic markers, due to patterns of high LD[39].

Biological characterisation of risk variants

Against the shared polygenic background involving several thousands of subthreshold ASD- and/or ADHD-risk associated variants, a small fraction (<5%, N=83) of SNPs passed a joint ASD and ADHD risk variant selection threshold (P_{thr}<0.0015). MVRs with these SNPs predicted larger association effects with EA, with non-overlapping 95%-CIs, compared to those observed in the discovery MVR analyses. Mapping the 83 loci to at least 52 genes, we found an enrichment of miRNA targets as well as several miRNA and lncRNA loci. Identified genes with miRNA targets encode, for example, biological signalling proteins such as the calcium voltage-gated channel subunit alpha1C (*CACNA1C*) and the tyrosinkinase ERBB4, which been previously associated with both ASD and ADHD as well as other disorders[40–42]. miRNAs are key regulators of many biological processes and often exert post-transcriptional gene silencing that can also be

influenced by environmental signals[43,44]. Their functionality is consistent with multiple regulatory sites in close genomic proximity (scenario IV, Figure 1g) or different regulations of the same site (scenario V, Figure 1i), but not with spurious pleiotropy due to functionally unrelated causal genetic variants in high LD (scenario IV, Figure 1h). Thus, our results provide support for a recently proposed class of genetic influences for psychiatric illness which does not confer broad liability to disorder, but is thought to shape the phenotype expression through direct and interactive genetic effects or environmental factors[4]. As construed by the omnigenic model[45], such 'peripheral' genetic influences, acting through trans effects, could control shared ADHD/ASD 'core' variation, although the candidacy of miRNAs functionality underlying pleiotropic effects requires replication in larger ASD and ADHD analyses.

Genetic inter-correlations between ASD and ADHD

Despite zero genetic correlations between ADHD(iPSYCH) and ASD(PGC), we observed strong evidence for cross-disorder polygenetic associations conveyed by the same GWAS marker alleles. Thus, it is possible that, beside a lack of power, the absence of genome-wide genetic correlation between ASD(PGC) and ADHD(iPSYCH) may involve a near symmetric distribution of positive and negative local genetic covariances and thus a cancellation of signals[39].

To illustrate this hypothesis, we re-formulated LDSC genetic correlations and SNP-heritability (SNP-h^2) estimates for ASD, ADHD and EA algebraically, as implied by a saturated trivariate structural equation model (Cholesky model, Supplementary Figure 5). This entails an assumption-free description of trait inter-relationships according to genetic theory (Formulae 8-13) and does not represent a statistical model, given the lack of raw data.

The simplified multi-factorial decomposition of genetic interrelationships between EA, ADHD(iPSYCH) and either ASD(iPSYCH) (Figure 4a) or ASD(PGC) (Figure 4b) describes two sources of shared genetic variation between ASD and ADHD. The first genetic factor (A1) captures shared variation between EA, ASD and ADHD and predicts negative genetic covariance between ASD and ADHD, consistent with MVR findings in this study. The second genetic factor (A2) describes, independently of A1, positive genetic covariance between ASD and ADHD, reflecting known positive or null genetic correlations between disorders[17,46]. The third genetic factor (A3) encodes ASD-specific variation, while alternative definitions of the model can also allow for ADHD-specific influences (Supplementary Figure 6).

According to such a model, the observed net genetic covariance between ASD and ADHD reflects the sum of negative and positive covariance contributions from the independent genetic factors A1 and A2. Consequently, the genome-wide genetic correlations between ASD and ADHD might be reduced, as hypothesised for ASD(iPSYCH)

and ADHD(iPSYCH) (Figure 4a), or completely abolished, as hypothesised for ASD(PGC) and ADHD(iPSYCH) (Figure 4b).

To confirm the plausibility of such a model, we conducted proof-of-principle simulations with model parameters derived from observed genetic correlations between ASD(PGC), ADHD(iPSYCH) and EA (Figure 4b). Assuming 6,000 unrelated individuals per trait and, for simplicity, larger SNP-h^2 for ASD, ADHD and EA (which does not affect

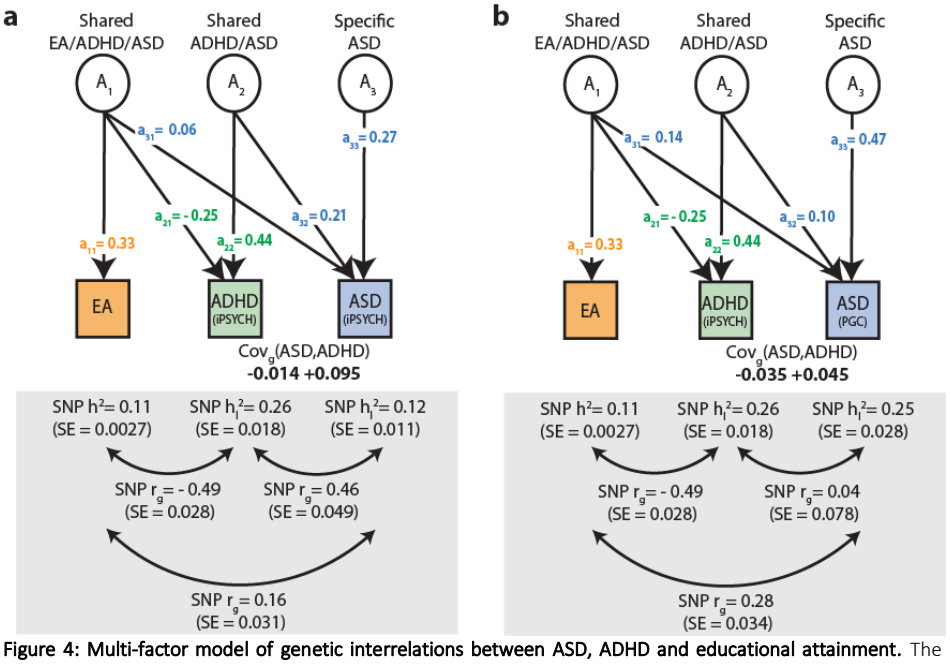


Figure 4: Multi-factor model of genetic interrelations between ASD, ADHD and educational attainment. The model predicts two sources of shared genetic influences between ASD and ADHD, as captured by common variants within an infinitely large population. The first genetic factor (A1, shared EA/ADHD/ASD) refers to shared genetic variation between EA, ADHD and ASD. It allows for negative genetic covariance between ASD and ADHD. The second genetic factor (A2, shared ADHD/ASD) acts independently of A1, explaining positive genetic covariance between ASD and ADHD. Additional ASD-specific effects are captured by a third factor (A3). Each factor loading ("a") for the Cholesky decomposition of a trivariate trait is described in the Methods. **(a)** Multi-factor model consistent with ASD(iPSYCH), ADHD(iPSYCH) and EA(SSGAC) summary statistics. **(b)** Multi-factor model consistent with ASD(PGC), ADHD(iPSYCH) and EA(SSGAC) summary statistics and supported by simulations (Supplementary Table 18). Factor loadings ("a") were derived from LDSC SNP-h^2 and genetic correlations according to theory. Phenotypic measures are represented by squares, while latent genetic factors are represented by circles. Single-headed arrows denote genetic factor loadings ("a"), double-headed arrows genetic correlations ("r_g"). Residual influences and unit variances for latent variables were omitted. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; EA, educational attainment; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; PGC, Psychiatric Genomics Consortium; SNP h^2, SNP heritability observed scale; SNP h_l^2, SNP heritability liability scale; SNP r_g, SNP genetic correlation, cov_g, genetic covariance

genetic correlations but reduces the computational burden), we could re-capture zero genetic correlations between ASD(PGC) and ADHD(iPSYCH) (Supplementary Table 19).

Thus, ASD and ADHD genetic architectures might be interlinked through a combination of positive and negative covariances, acting locally, that influence detectable genome-wide correlations between ASD and ADHD and may, potentially, also affect GWAS signals, when both ASD and ADHD patients are jointly analysed within cross-disorder investigations.

Cross-disorder genetic associations for other disorders

Genetic cross-disorder association effects with EA were strongest and largest when ADHD SNP estimates were assigned to ASD-risk associated alleles and, vice versa, ASD SNP estimates were assigned to ADHD-risk associated alleles, as studied in iPSYCH samples. However, cross-disorder effects are unlikely to be limited to ASD and ADHD genetic architectures. Despite the preliminary character of our analyses, we showed that the same selected ASD- and ADHD-related marker alleles also captured polygenetic relationships with EA for adult-onset disorders, such as SCZ, MDD and BD, consistent with the wide-spread pleiotropy among neuropsychiatric conditions[4]. Several of these cross-disorder effects were discordant compared to ASD or ADHD risk effects. Thus, local negative covariance patterns may also shape the genetic overlap between other psychiatric conditions.

Strengths and limitations

Adopting a statistical framework developed for Mendelian Randomization analyses[32], this study disentangled ASD and ADHD single-variant effects at the same GWAS marker allele and identified independent polygenic associations with EA for ASD and ADHD risk. Evidence for discordant ADHD and ASD cross-disorder association profiles with EA was replicated using two independent ASD collections at a variant selection threshold (P_{thr}<0.05), which is commonly applied in polygenic scoring analyses[47]. This suggests that our findings are unlikely to be affected by diagnostic classification systems for clinical ASD, routes of patient ascertainment or association analysis designs. Furthermore, the use of fully independent samples, ASD(PGC), ADHD(iPSYCH) and EA(SSGAC), increased the robustness of ASD- and ADHD-specific findings in UVRs and MVRs and fully eliminated any variance inflation (VIF=1, Supplementary Table 10). However, the SNP effect concordance between ASD(PGC) and ASD(iPSYCH, woADHD) was only ~50%, consistent with the low power of ASD(PGC), but also some genetic heterogeneity between samples, possibly due to different routes of patient ascertainment[17,21,41]. This may have, consequently, decreased the power of follow-up analyses in this study. In comparison, the SNP effect concordance between risk

effects in ASD(iPSYCH,woADHD) and ADHD(iPSYCH) approached ~80%, suggesting similar genetic architectures, as captured by GWAS markers. The modest variance inflation (VIF≤1.4) between ASD(iPSYCH,woADHD) and ADHD(iPSYCH) might be the consequence of shared controls between different iPSYCH samples, as included in other large cross-disorder-studies[4]. This could, possibly, result in negative confounding in UVRs. However, we showed that independent polygenic risk effects can be successfully disentangled and de-stratified using MVR. In addition, there is a possibility that our findings are affected by ascertainment bias (scenario II, Figure 1c). Finally, ASD and ADHD symptom heterogeneity may shape the genetic overlap between neurodevelopmental disorders, EA and cognition-related traits[17,48] and future studies with access to this information are warranted to fully understand the underlying complex multivariate inter-correlations.

7.4. Conclusions

Our findings show that EA-related polygenic variation is shared across ASD and ADHD genetic architectures and that combinations of the same risk alleles, through mechanisms consistent with pleiotropy or co-localisation, can encode ASD-related positive and ADHD-related negative associations with EA, without involving further loci. Our results imply local negative genetic covariance between ASD and ADHD risk that may contribute to the total detectable genome-wide correlation between both disorders.

7.5. Methods

Data sets

Genome-wide SNP information on EA, intelligence, reading and neuropsychiatric disorders was obtained from GWAS summary statistics[17,20,21,49–51]. These aggregated results are briefly summarised here and described in detail in Table 1 and Supplementary Table 1.

EA, intelligence and reading: GWAS summary statistics on years-of-schooling (excluding 23andMe) were obtained from the Social Science Genetic Association Consortium (SSGAC, https://www.thessgac.org/, Table 1)[49]. EA was coded according to the International Standard Classification of Education (1997) scale[49] and analysed as a quantitative variable defined as an individual's years-of-schooling. Participants were >30 years of age at the time of assessment and of European ancestry. The meta-analysis consisted primarily of population-based cohorts, but also included family-based and case-control samples. 55.2% of participants were female. For most cohorts, genome-wide data were imputed to a 1000 genomes project version 3 reference template, as previously described[49].

GWAS summary statistics on intelligence[20] were retrieved from the Complex Trait Genetics (CTG) lab (https://ctg.cncr.nl/software/summary_statistics, Supplementary Table 1). Participating cohorts were primarily population-based. Each cohort assessed intelligence with different instruments that were re-defined to index a common latent factor of general intelligence[20]. Participants had a wide age range (from 5 to 98 years), 51.2% were female and all of them were of European descent. Genome-wide data were predominantly imputed to the Haplotype Reference Consortium (HRC) reference panel, as previously described[20].

Genome-wide association summary statistics on reading were derived by conducting a fixed-effect meta-analysis of reading abilities (N=13,027, Supplementary Table 1) as assessed in the Avon Longitudinal Study of Parents and Children[52,53] (ALSPAC, N=4,247), the 1958 Birth Cohort[54] (1958BC, N=4,638) and Philadelphia Neurodevelopmental Cohort[55,56] (PNC, N=4,142). Within ALSPAC, word reading speed was assessed at 13 years using the Test of Word Reading Efficiency[57] (TOWRE). Within the PNC, reading accuracy was assessed in participants between 8 and 22 years of age using the reading items of the Wide Range Achievement Test[58] and within the 1958BC, reading comprehension was assessed at 11 years of age using a study-specific reading comprehension test designed to parallel the Watts-Vernon test of reading ability. Here, the child was required to choose from a selection of five words the word that appropriately completed the sentence. There were 35 questions in total and the reliability coefficient of this test is 0.82. Reading scores were adjusted for sex, age, age^2, the first two principal components and study-specific covariates such as batch, if applicable. For each cohort, genome-wide genotyping data were imputed against the HRC r1.1 reference panel[59] and association tests were performed using SNPTEST[60] (version 2.5.2). Finally, a fixed-effect meta-analysis across all three cohorts was performed using METAL[61] (N_{max}=13,027).

ASD and ADHD: GWAS summary statistics for ASD and ADHD were accessed through the Danish Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH, http://ipsych.au.dk/) using samples from the Danish Neonatal Screening Biobank hosted by Statens Serum Institute[17,21,62] (ASD(iPSYCH), ADHD(iPSYCH), Table 1). iPSYCH adopts a case-control design (26.6% female ASD-cases[17], 21.6% female ADHD-cases[21]) with shared controls (~49% female)[17,21], all of European ancestry with age ranges spanning infancy to adulthood[17,21]. For MVR analyses, ASD samples were restricted to ASD-cases without (wo) an additional ADHD diagnosis (ASD(iPSYCH,woADHD), Table 1) to avoid overlap with ADHD(iPSYCH). However, ADHD-cases may have an additional ASD diagnosis. Information on ADHD cases without ASD was not available.

ASD cases and ADHD cases were diagnosed according to ICD-10[63] and identified using the Danish Psychiatric Central Research Register[64]. Registry-based ASD diagnoses were validated previously[17,21]. Controls were randomly selected from the same nationwide birth cohort and did not have a diagnosis of ASD or ADHD or moderate-

severe mental retardation (F71-F79)[17,21,62]. The median age at first diagnosis of ASD was 10 years. Genotyping was performed using the Illumina Infinium PsychArray BeadChip and genotypes were imputed to a 1000 Genomes template (Phase3, release 02-05-2013). Genotyping, quality control, imputation and genetic association analysis were carried out using the Ricopili pipeline with standard PGC settings[17,21].

Independent ASD GWAS summary statistics were obtained from the Psychiatric Genomics Consortium (PGC, www.med.unc.edu/pgc/). They were based on a case-control/pseudo-control design and all individuals were ≥3 years of age and of European ancestry (ASD(PGC), Table 1). Information on the male-female ratio was not available[41]. A consensus ASD diagnosis was made using research standard diagnoses and expert clinical consensus diagnoses. The majority of ASD-cases (94.1%) also had a clinical diagnosis based on the Autism Diagnostic Interview-Revised[65] or the Autism Diagnostic Observation Schedule[66]. Genome-wide data were imputed to a 1000 Genomes reference template (Phase1 v3). Note that the sample size for ADHD(iPSYCH) is about three times larger than for ASD(PGC).

Other psychiatric disorders: To assess the specificity of MVR association profiles, we also investigated GWAS summary statistics for MDD[50], SCZ[51] and BD[51]. Cases were identified based on international consensus criteria. For MDD, cases were identified based on a lifetime diagnosis of MDD, established using DSM-III, DSM-IV, ICD-9 and/or ICD-10 criteria or self-report[50]. For SCZ, the majority of cases were diagnosed using DSM-III, DSM-III-R, DSM-IV, ICD-10, and SCID criteria[47,51]. BD cases were diagnosed according to DSM-III, DSM-IV-TR, DSM-IV, SCID, ICD-10 or RDC criteria[51,67]. For all three data sets, genotype imputation was performed using the IMPUTE2/SHAPEIT pipeline against the 1000 Genomes Project (v3) template. Summary data were obtained from the PGC (www.med.unc.edu/pgc/, Supplementary Table 1), all based on participants of European ancestry.

Sample overlap: Independent GWAS summary statistics were available for EA[49], intelligence[20], reading, ASD(PGC)[35] and ADHD(iPSYCH)[21] (Table 1). ASD(PGC), ADHD(iPSYCH) and ASD(iPSYCH,woADHD)[17] GWAS statistics have independent case samples, and case-overlap with MDD, SCZ or BD cases is at most 2%, similar to recent cross-disorder analyses[4] (Table 1, Supplementary Table 1). Similarly, iPSYCH controls were shared across reported ASD(iPSYCH,woADHD), ADHD(iPSYCH), MDD, SCZ and BD summary statistics, as in recent cross-disorder analyses[4]. Also, UKBiobank and other large cohorts are shared across EA and MDD, SCZ or BD summary statistics (but not ASD and ADHD), and presented MVR studies for adult-onset disorders have thus exploratory character only.

SNP-heritability and genetic correlations

SNP-h^2, the proportion of phenotypic or liability variance tagged by SNPs on genotyping arrays, was estimated for EA, intelligence, reading and psychiatric disorders

using LDSC regression[67] (Supplementary Table 2). To estimate LDSC-h^2, genome-wide χ^2-statistics are regressed on the amount of genetic variation captured by each SNP[68], while the intercept of this regression minus one is an estimator of the mean contribution of confounding bias to the inflation in the mean χ^2-statistic[68]. SNP-h^2 was calculated on the liability scale for psychiatric disorder samples, assuming a population prevalence of 0.012 for ASD[17], 0.05 for ADHD[69], 0.162 for MDD[70], 0.007 for SCZ[71] and 0.006 for BD[72].

In extension, unconstrained LDSC correlation[72] analysis was applied to estimate bivariate genetic correlations (r_g) (Supplementary Table 3-5). This involves a regression of the product of test statistics on LD score and captures the extent of shared genetic influences between phenotypes assessed in different samples[73].

All analyses were performed with LDSC software[67,73] and based on the set of well-imputed HapMap3 SNPs and a European reference panel of LD scores[73].

Multivariable regression analysis

We adopted a bidirectional inverse-variance weighted regression framework, MVR, analogous to statistical models proposed for multivariable MR[32]. This approach was implemented here using GWAS summary statistics, often described as Egger regression[32]. MVR analyses do not infer causality as we allow for biological pleiotropy. We apply this method to simultaneously estimate genetic ASD and ADHD risk associations with EA. We control for collider bias that may arise when adjusting for heritable covariates[33] by studying relationships between genetically predicted phenotypes only.

Variant set identification: Multiple ASD (Gi) and ADHD (Gj) variant sets were created according to guidelines for polygenic scoring methods[34]. ASD-related variant sets Gi (with i = 1,...,I SNPs) and ADHD-related variant sets Gj (with j = 1,...,J SNPs) in this study were selected from ASD(iPSYCH,woADHD) and ADHD(iPSYCH) GWAS statistics respectively, using multiple *P*-value thresholds as described below (MVR study design). All variant sets were restricted to common (minor allele frequency>0.01), independent (LD-r^2<0.25 within ±500 kb) and well-imputed (imputation quality(INFO) >0.7) SNPs.

Estimation of ASD-specific, ADHD-specific and cross-disorder genetic associations with EA: For each ASD variant set Gi, an ASD-MVR was fitted as follows:

$$\hat{\beta}_{EA_i} = \theta_{0*} + \theta_{ASD}\hat{\beta}_{ASD_i} + \theta_{\maltese ADHD}\hat{\beta}_{ADHD_i} + \varepsilon_i \quad (1)$$

$$weights = se(\hat{\beta}_{EA_i})^{-2} \quad (2)$$

where $\hat{\beta}_{EA_i}$ (dependent variable) are SNP estimates for EA, $\hat{\beta}_{ASD_i}$ (independent variable) are SNP estimates for ASD and $\hat{\beta}_{ADHD_i}$ (independent variable) are SNP estimates for ADHD. θ_{0*} is the MVR regression intercept, θ_{ASD} is the MVR ASD-specific effect and $\theta_{\maltese ADHD}$ the MVR cross-disorder effect, weighted by the inverse variance of the dependent variable, consistent with the statistical framework of Egger regression-based MVR

analyses[32]. The intercept θ_{0*} is an estimate of α_i'[32], the direct pleiotropic influences between the analysed variants Gi and EA that are neither captured by θ_{ASD} nor $\theta_{\maltese ADHD}$.

Similarly, for each ADHD variant set Gj, an ADHD-MVR was fitted as follows:

$$\hat{\beta}_{EA_j} = \theta_{0\#} + \theta_{ADHD}\hat{\beta}_{ADHD_j} + \theta_{\maltese ASD}\hat{\beta}_{ASD_j} + \varepsilon_j \quad (3)$$

$$weights = se(\hat{\beta}_{EA_j})^{-2} \quad (4)$$

where $\hat{\beta}_{EA_j}$ (dependent variable) are SNP estimates for EA, $\hat{\beta}_{ADHD_j}$ (independent variable) are SNP estimates for ADHD and $\hat{\beta}_{ASD_j}$ (independent variable) are the SNP estimates for ASD. $\theta_{0\#}$ is the MVR regression intercept, θ_{ADHD} the MVR ADHD-specific effect and $\theta_{\maltese ASD}$ the MVR cross-disorder effect. The intercept $\theta_{0\#}$ is an estimate of α_j'[32], the direct pleiotropic influences between the analysed variants G_j and EA that are neither captured by θ_{ADHD} nor $\theta_{\maltese ASD}$.

Cross-disorder effects were thus estimated bi-directionally: (1) $\theta_{\maltese ADHD}$, based on ADHD SNP estimates ($\hat{\beta}_{ADHD_i}$) for ASD variant set Gi and (2) $\theta_{\maltese ASD}$, based on ASD SNP estimates ($\hat{\beta}_{ASD_j}$) for ADHD variant set Gj.

Reported MVR effects ($\hat{\theta}$) present changes in years-of-schooling, either per increase in log-odds ASD or ADHD liability, pooled across the variant set. The overall MVR model fit was compared to UVR models (see below) using likelihood-ratio tests, as implemented in the R:stats library (Rv3.5.1). Collinearity between independent variables was assessed by the VIF(R:car library (Rv3.5.1)).

Multivariable MR Egger-related approaches with intercept terms, including MVR analyses applied in this study, are sensitive to the allelic alignment. It has been recommended to orient all variants with respect to the genetic association with the independent variable of primary interest[32]. Thus, SNP estimates were aligned to increase ASD risk in ASD-MVRs and ADHD risk in ADHD-MVRs. As follow-up analyses, we carried out MVR with subsets of variants that have the same risk-increasing allele for both ASD and ADHD (concordant variants, Supplementary Figure 1c-d). Consequently, by design, SNP estimates were aligned to increase both ASD and ADHD risk (see below).

Multivariable regression study design: MVR analyses were conducted in six different stages (Supplementary Figure 1): For discovery MVR analyses (1, Supplementary Figure 1a-b), SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH,woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics. 11 ASD-related Gi variant sets and 11 ADHD-related Gj variant sets were selected from ASD(iPSYCH,woADHD) and ADHD(iPSYCH) GWAS statistics respectively, using multiple *P*-value thresholds (P_{thr}, $5x10^{-8}$; $5x10^{-7}$; $5x10^{-6}$; $5x10^{-5}$; 0.0005; 0.0015; 0.005; 0.05; 0.1; 0.3; 0.5). For simplicity, MVR findings in the main manuscript are presented for two *P*-value thresholds only: P_{thr}<0.0015, consistent with conservative selection thresholds recommended for polygenic scoring approaches[34], and P_{thr}<0.05, a less stringent threshold that has been previously selected to study polygenic scores in complex psychiatric disorders[47], to increase the statistical power and precision of MVR estimates.

Discovery MVR analyses were repeated using concordant variant sets (2, Supplementary Figure 1c-d, *P*-value thresholds: P_{thr}<0.0015; P_{thr}<0.05).

To replicate MVR findings (3, Supplementary Figure 1e-f), ASD-MVR and ADHD-MVR were conducted using ASD SNP estimates ($\hat{\beta}_{ASD}$) from ASD(PGC), instead of ASD(iPSYCH,woADHD) (aligned to increase ASD risk as observed in ASD(PGC), using ASD and ADHD variant sets from the discovery analyses (*P*-value thresholds: P_{thr}<0.0015; P_{thr}<0.05).

As part of sensitivity analyses (4, Supplementary Figure 1g-h), ASD-MVR and ADHD-MVR models were fitted with intelligence and reading as dependent variables ($\hat{\beta}_{Intelligence}$ and $\hat{\beta}_{reading}$), instead of EA ($\hat{\beta}_{EA}$), using GWAS summary statistics from Intelligence(CTG) and reading respectively, and ASD and ADHD variant sets from the discovery analyses (*P*-value thresholds: P_{thr}<0.0015; P_{thr}<0.05).

To identify variants underlying the observed MVR cross-disorder associations with EA, we created variant sets meeting joint ASD and ADHD selection criteria (5, Supplementary Figure 1i-j). We assessed the proportion of overlapping independent SNPs associated with both ASD and ADHD risk using PLINK (https://www.cog-genomics.org/plink2; 500 kb and LD-r^2≥0.6). We started with the ASD and ADHD variant sets from the discovery analyses at P_{thr}<0.0015: ASD (Gi, N_{SNPs}=1,973) and ADHD (Gj, N_{SNPs}=2,717). For each variant set, we identified the SNPs that were also associated with the other disorder across a range of *P*-value thresholds (0.0015; 0.005; 0.05; 0.1; 0.3; 0.5). In total, this resulted in six subsets for ASD-related variants (Gij) and six subsets for ADHD-related variants (Gji).

To assess the specificity of the observed discordant cross-disorder associations with EA, we carried out MVR sensitivity analyses (6, Supplementary Figure 1k-l) modelling SNP estimates for MDD ($\hat{\beta}_{MDD}$), SCZ ($\hat{\beta}_{SCZ}$) or BD ($\hat{\beta}_{BD}$) for either ASD (Gi, ASD-MVR) or ADHD (Gj, ADHD-MVR) variant sets (P_{thr}<0.0015; P_{thr}<0.05) as used in the discovery MVR analyses. SNP estimates $\hat{\beta}_{ASD}$, $\hat{\beta}_{ADHD}$, $\hat{\beta}_{MDD}$, $\hat{\beta}_{SCZ}$, $\hat{\beta}_{BD}$ and $\hat{\beta}_{EA}$ were extracted from ASD(iPSYCH,woADHD), ADHD(iPSYCH), MDD(PGC), SCZ(PGC), BD(PGC) and EA(SSGAC) GWAS statistics respectively (Table 1, Supplementary Table 1).

Multiple testing correction: We applied the following conservative Bonferroni-corrected multiple testing thresholds for the MVR analysis stages described above: (1) discovery analyses with two MVR models across 11 variant sets (22 tests, $P_{Adjusted}$=0.0023, Supplementary Figure 1a-b) with (2) concordant variant analyses being nested within these discovery analyses (Supplementary Figure 1c-d); (3) follow-up analyses with independent ASD(PGC) SNP estimates with two MVR models across two variant sets (4 tests, $P_{Adjusted}$=0.0125, Supplementary Figure 1e-f); (4) follow-up analyses with intelligence(CTG) and reading SNP estimates with two MVR models across two variant sets (8 tests, $P_{Adjusted}$=0.0006, Supplementary Figure 1g-h); (5) screening of MVR effect sizes with variant sets meeting joint ASD and ADHD variant selection criteria: For each MVR model, variant sets selected at P_{thr}<0.0015 were successively restricted to SNPs that

are associated with both disorders, using six *P*-value thresholds nested within the set of discovery analyses ($P_{Adjusted}$=0.0023 as described above, Supplementary Figure 1i-j); (6) follow-up analyses with MDD(PGC), SCZ(PGC) and BD(PGC) SNP estimates with two MVR models across two variant sets (12 tests, $P_{Adjusted}$=0.0042, Supplementary Figure 1k-l).

Univariable regression models

To assess the robustness of MVR findings, weighted UVRs were included in the stages 1 to 4 and 6 of the MVR study design (Supplementary Figure 1a-h,k-l). UVRs included the same dependent variable but only one of the two independent variables described for MVR analyses (Formulae: 1-4), capturing either disorder-specific or cross-disorder effects. Additionally, we carried out a weighted UVR regressing ADHD SNP estimates ($\hat{\boldsymbol{\beta}}_{\text{ADHD}}$) on ASD SNP estimates ($\hat{\boldsymbol{\beta}}_{\text{ASD}}$) for concordant variants. UVR and MVRs model fit was compared using a likelihood-ratio test as implemented in the R:stats library (Rv3.5.1).

Gene-set enrichment analyses

SNPs were mapped to 52 RefSeq gene IDs (genome build 37) based on positional mapping using PLINK software (0 kb gene window) (https://www.cog-genomics.org/plink2), similar to the default options in MAGMA gene-enrichment software[73]. Gene-set enrichment (>5 overlapping genes, FDR-adjusted *P*-values) was conducted within BrainSpan samples[37] (https://www.brainspan.org/), based on 29 different ages of brain samples and 11 developmental stages of brain samples. Enrichment analysis was carried out with MAGMA[74], as implemented in FUMA[36] software (https://fuma.ctglab.nl/), by mapping genes to unique Ensembl IDs (v92).

Structural equation modelling

To summarise genetic interrelationships between EA, ASD and ADHD with a multi-factor model, we translated LDSC SNP-h^2 and r_g estimates (Supplementary Table 3-4) into hypothetical factor loadings consistent with structural equations for a saturated model. Specifically, we propose a multi-factorial structural equation model consisting of three continuous phenotypes (EA, ASD liability and ADHD liability Z-scores), three independent latent genetic factors and three independent latent residual influences. We assume that genetic factors give rise to genetic variances and covariances between EA, ASD and ADHD liability, while residual covariances are assumed to be absent. Phenotypic variances and covariances were described according to a Cholesky decomposition[75] (i.e. a saturated model), assuming an infinitely large population and a fully identified model (Supplementary Figure 5). A Cholesky model involves the decomposition of both the genetic variances and residual variances into as many latent factors as there are

observed variables. The expected phenotypic covariance matrix $\boldsymbol{\Sigma}$ for Z-standardised traits based on the factor model is

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Gamma}\boldsymbol{\Theta}\boldsymbol{\Gamma}' \tag{5}$$

where $\boldsymbol{\Lambda}$ is a lower triangular matrix of genetic factor loadings, $\boldsymbol{\Phi}$ is a diagonal matrix of latent genetic factor variances (standardised to unit variance) such that $\boldsymbol{\Phi}$ is an identity matrix ***I***. The residual variance can be decomposed into latent residual factors, where $\boldsymbol{\Gamma}$ is a lower triangular matrix of residual factor loadings and $\boldsymbol{\Theta}$ is a diagonal matrix of latent residual factor variances (standardised to unit variance) such that $\boldsymbol{\Theta}$ is an identity matrix ***I***. For example, for a trivariate model consisting of measures P_1, P_2 and P_3, assuming three genetic factors (A_1, A_2 and A_3) and three residual factors (E_1, E_2 and E_3), the expected phenotypic covariance matrix can be expressed as follows:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{p1}^2 & \sigma_{p12} & \sigma_{p13} \\ \sigma_{p12} & \sigma_{p2}^2 & \sigma_{p23} \\ \sigma_{p13} & \sigma_{p23} & \sigma_{p3}^2 \end{bmatrix} \tag{6}$$

with the relevant matrices

$$\boldsymbol{\Lambda} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \boldsymbol{\Phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{\Gamma} = \begin{bmatrix} e_{11} & 0 & 0 \\ e_{21} & e_{22} & 0 \\ e_{31} & e_{32} & e_{33} \end{bmatrix}, \boldsymbol{\Theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

where σ_{p1}^2, σ_{p2}^2 and σ_{p3}^2 represent the phenotypic variances, and σ_{p12}, σ_{p13} and σ_{p23} phenotypic covariances. We annotate the genetic factor loadings *a* (factor loadings) such that the first number indicates the direction of the effect (the variable to which the arrow points) and the second the origin of the effect.

The trivariate AE Cholesky decomposition of three standardised measures, as described above, can be visualised by means of a path diagram (Supplementary Figure 5) and the expected phenotypic variances and covariances can be expressed as follows:

$$\sigma_{p1}^2 = a_{11}^2 + e_{11}^2 = 1 \tag{8}$$

$$\sigma_{p2}^2 = (a_{21}^2 + a_{22}^2) + (e_{21}^2 + e_{22}^2) = 1 \tag{9}$$

$$\sigma_{p3}^2 = (a_{31}^2 + a_{32}^2 + a_{33}^2) + (e_{31}^2 + e_{32}^2 + e_{33}^2) = 1 \tag{10}$$

$$\sigma_{p12} = a_{11}a_{21} + e_{11}e_{21} \tag{11}$$

$$\sigma_{p13} = a_{11}a_{31} + e_{11}e_{31} \tag{12}$$

$$\sigma_{p23} = a_{31}a_{21} + a_{32}a_{22} + e_{31}e_{21} + e_{32}e_{22} \quad (13)$$

The variance of the latent genetic and residual factors has been standardised to unit variance and is not shown.

Estimated genetic variances and covariances can be used to derive genetic correlation estimates between two phenotypes measuring the extent to which two phenotypes 1 and 2 share genetic factors (ranging from -1 to 1):

$$r_g = \frac{\sigma_{g12}}{\sqrt{\sigma^2_{g1}\sigma^2_{g2}}} \quad (14)$$

where σ_{g12} is the genetic covariance between phenotypes 1 and 2, and σ^2_{g1}, σ^2_{g2} their genetic variances.

We derived (but did not fit) hypothetical factor loadings, based on LDSC SNP-h^2 estimates for EA and, on the liability scale, ASD and ADHD risk (Supplementary Table 2), as well as unconstrained LDSC genetic correlations (Supplementary Table 3-4) using EA, ASD(iPSYCH), ASD(PGC) and ADHD(iPSYCH) GWAS statistics (Table 1). We support the plausibility of such a model using simulations (Supplementary Table 19).

Multi-factor model data simulation

To evaluate the plausibility of the proposed multi-factorial model, we carried out data simulations (Supplementary Table 19). Assuming multivariate normality and unrelated individuals, we simulated three continuous interrelated measures P_1, P_2 and P_3 corresponding to EA and liability for ADHD and ASD respectively, according to a Cholesky model. This includes three genetic factors with their variances and covariances and three residual factors with their variances and their covariances. The genetic interrelationships between these three traits were informed by unconstrained LDSC genetic correlations between EA, ADHD and ASD (Supplementary Table 3-4) using EA, ADHD(iPSYCH) and ASD(PGC) summary statistics and structural equations described above (Formulae 8-14). Residual interrelationships were assumed to be absent as the cohorts are independent of each other. However, simulated SNP-h^2 estimates were increased, and sample sizes restricted to 6,000 individuals per trait with 20,000 SNPs per genetic factor, to ease the computational burden (72h, using 16 cores). Multivariate variances and covariances within the simulated data were modelled using genetic-relationship-matrix structural equation modelling (GSEM, R gsem library, v0.1.2)[76]. This method involves a multivariate analysis of genetic variance by combining whole-genome genotyping information in unrelated individuals with structural equation modelling techniques using a full information maximum likelihood approach. Simulated parameters and estimated parameters are shown in Supplementary Table 19.

Acknowledgements

This research was facilitated by the Social Science Genetic Association Consortium (SSGAC), Complex Trait Genetics (CTG) lab, Psychiatric Genomics Consortium (PGC) and iPSYCH-Broad-PGC ASD Consortium, by providing access to GWAS summary statistics. This publication is the work of the authors and EV and BSTP will serve as guarantors for the contents of this paper. EV and BSTP are supported by the Max Planck Society. BSTP is supported by the Simons Foundation (Award ID: 514787). The iPSYCH project is funded by the Lundbeck Foundation (grant no R102-A9118 and R155-2014-1724) and the universities and university hospitals of Aarhus and Copenhagen. ADB is also supported by the EU's Horizon 2020 programme (grant no 667302, CoCA). Data handling and analysis was supported by NIMH (1U01MH109514-01 to Michael O'Donovan and Anders D Børglum). High-performance computer capacity for handling and statistical analysis of iPSYCH data on the GenomeDK HPC facility was provided by the Centre for Integrative Sequencing, iSEQ, Aarhus University, Denmark (grant to Anders D Børglum) and Center for Genomics and Personalized Medicine, Aarhus, Denmark.

We would also like to thank ALSPAC, PNC and 1958BC. We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and they will serve as guarantors for the contents of this paper. GWAS data were generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). We thank dbGap for providing access to data from PNC (accession number 12043). Support for the collection of the data for PNC was provided by grant RC2MH089983 awarded to Raquel Gur and RC2MH089924 awarded to Hakon Hakonarson. Subjects were recruited and genotyped through the Center for Applied Genomics (CAG) at The Children's Hospital in Philadelphia (CHOP). Phenotypic data collection occurred at the CAG/CHOP and at the Brain Behavior Laboratory, University of Pennsylvania. Cognitive assessment was performed using the Penn Computerized Neurocognitive Battery (CNB), which was customized and shortened for a pediatric population using both classical psychometric and item response theory methods. The samples released were all previously genotyped using Illumina or Affymetrix SNP-array platform by the Center for Applied Genomics at The Children's Hospital of Philadelphia. This work also made use of data and samples generated by the 1958 Birth Cohort (NCDS), which is managed by the Centre for Longitudinal Studies at the UCL Institute of

Education, funded by the Economic and Social Research Council (grant number ES/M001660/1). Access to these resources was enabled via the Wellcome Trust & MRC: 58FORWARDS grant [108439/Z/15/Z] (The 1958 Birth Cohort: Fostering new Opportunities for Research via Wider Access to Reliable Data and Samples). Before 2015 biomedical resources were maintained under the Wellcome Trust and Medical Research Council 58READIE Project (grant numbers WT095219MA and G1001799). Genotyping was undertaken as part of the Wellcome Trust Case-Control Consortium (WTCCC) under Wellcome Trust award 076113, and a full list of the investigators who contributed to the generation of the data is available at www.wtccc.org.uk. Data governance was provided by the METADAC data access committee, funded by ESRC, Wellcome, and MRC. (2015-2018: Grant Number MR/N01104X/1 2018-2020: Grant Number ES/S008349/1). Finally, we thank Simon E. Fisher for helpful discussions of the manuscript.

References

1. Thapar, A. & Cooper, M. Attention deficit hyperactivity disorder. *The Lancet* **387**, 1240–1250 (2016).
2. Lord, C., Elsabbagh, M., Baird, G. & Veenstra-Vanderweele, J. Autism spectrum disorder. *The Lancet* **392**, 508–520 (2018).
3. Antshel, K. M., Zhang-James, Y., Wagner, K. E., Ledesma, A. & Faraone, S. V. An update on the comorbidity of ADHD and ASD: a focus on clinical management. *Expert Rev Neurother* **16**, 279–293 (2016).
4. Lee, P. H. *et al.* Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482.e11 (2019).
5. Rommelse, N. N. J., Franke, B., Geurts, H. M., Hartman, C. A. & Buitelaar, J. K. Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *Eur Child Adolesc Psychiatry* **19**, 281–295 (2010).
6. Leitner, Y. The co-occurrence of autism and attention deficit hyperactivity disorder in children - what do we know? *Front Hum Neurosci* **8**, 268 (2014).
7. Ronald, A., Simonoff, E., Kuntsi, J., Asherson, P. & Plomin, R. Evidence for overlapping genetic influences on autistic and ADHD behaviours in a community twin sample. *J Child Psychol Psychiatry* **49**, 535–542 (2008).
8. Stergiakouli, E. *et al.* Shared genetic influences between dimensional ASD and ADHD symptoms during child and adolescent development. *Molecular Autism* **8**, 18 (2017).
9. Ronald, A., Larsson, H., Anckarsäter, H. & Lichtenstein, P. Symptoms of autism and ADHD: A Swedish twin study examining their overlap. *Journal of Abnormal Psychology* **123**, 440–451 (2014).
10. Polderman, T. J. C., Hoekstra, R. A., Posthuma, D. & Larsson, H. The co-occurrence of autistic and ADHD dimensions in adults: an etiological study in 17 770 twins. *Transl Psychiatry* **4**, e435 (2014).
11. Taylor, M. J., Charman, T. & Ronald, A. Where are the strongest associations between autistic traits and traits of ADHD? evidence from a community-based twin study. *Eur Child Adolesc Psychiatry* **24**, 1129–1138 (2015).
12. Taylor, M. J. *et al.* Developmental associations between traits of autism spectrum disorder and attention deficit hyperactivity disorder: a genetically informative, longitudinal twin study. *Psychol Med* **43**, 1735–1746 (2013).
13. Martin, J., Hamshere, M. L., Stergiakouli, E., O'Donovan, M. C. & Thapar, A. Genetic Risk for Attention-Deficit/Hyperactivity Disorder Contributes to Neurodevelopmental Traits in the General Population. *Biol Psychiatry* **76**, 664–671 (2014).
14. Lichtenstein, P., Carlström, E., Råstam, M., Gillberg, C. & Anckarsäter, H. The Genetics of Autism Spectrum Disorders and Related Neuropsychiatric Disorders in Childhood. *AJP* **167**, 1357–1363 (2010).
15. Ghirardi, L. *et al.* The familial co-aggregation of ASD and ADHD: a register-based cohort study. *Mol. Psychiatry* **23**, 257–262 (2018).
16. Martin, J. *et al.* Biological overlap of attention-deficit/hyperactivity disorder and autism spectrum disorder: evidence from copy number variants. *J Am Acad Child Adolesc Psychiatry* **53**, 761-770.e26 (2014).
17. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics* **51**, 431 (2019).
18. Tick, B. *et al.* Autism Spectrum Disorders and Other Mental Health Problems: Exploring Etiological Overlaps and Phenotypic Causal Associations. *J Am Acad Child Adolesc Psychiatry* **55**, 106-113.e4 (2016).
19. Martin, J., Taylor, M. J. & Lichtenstein, P. Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychol Med* **48**, 1759–1774 (2018).
20. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics* **50**, 912–919 (2018).
21. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics* 1 (2018) doi:10.1038/s41588-018-0269-7.
22. Clarke, T.-K. *et al.* Common polygenic risk for autism spectrum disorder (ASD) is associated with cognitive ability in the general population. *Mol Psychiatry* **21**, 419–425 (2016).
23. Loe, I. M. & Feldman, H. M. Academic and Educational Outcomes of Children With ADHD. *J Pediatr Psychol* **32**, 643–654 (2007).

24. Keen, D., Webster, A. & Ridley, G. How well are children with autism spectrum disorder doing academically at school? An overview of the literature. *Autism* **20**, 276–294 (2016).
25. Frank, F. *et al.* Education and employment status of adults with autism spectrum disorders in Germany – a cross-sectional-survey. *BMC Psychiatry* **18**, 75 (2018).
26. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics* **14**, 483–495 (2013).
27. Durkin, M. S. *et al.* Socioeconomic Inequality in the Prevalence of Autism Spectrum Disorder: Evidence from a U.S. Cross-Sectional Study. *PLOS ONE* **5**, e11551 (2010).
28. Rai, D. *et al.* Parental Socioeconomic Status and Risk of Offspring Autism Spectrum Disorders in a Swedish Population-Based Study. *Journal of the American Academy of Child & Adolescent Psychiatry* **51**, 467-476.e6 (2012).
29. Russell, A. E., Ford, T., Williams, R. & Russell, G. The Association Between Socioeconomic Disadvantage and Attention Deficit/Hyperactivity Disorder (ADHD): A Systematic Review. *Child Psychiatry Hum Dev* **47**, 440–458 (2016).
30. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nature Genetics* **51**, 1339–1348 (2019).
31. Ke, X. Presence of Multiple Independent Effects in Risk Loci of Common Complex Human Diseases. *Am J Hum Genet* **91**, 185–192 (2012).
32. Rees, J. M. B., Wood, A. M. & Burgess, S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med* **36**, 4705–4718 (2017).
33. Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for Heritable Covariates Can Bias Effect Estimates in Genome-Wide Association Studies. *The American Journal of Human Genetics* **96**, 329–339 (2015).
34. Wray, N. R. *et al.* Research Review: Polygenic methods and their application to psychiatric traits. *J Child Psychol Psychiatr* **55**, 1068–1087 (2014).
35. Autism Spectrum Disorder Working Group of the Psychiatry Genomics Consortium. Dataset: PGC-ASD summary statistics from a meta-analysis of 5,305 ASD-diagnosed cases and 5,305 pseudocontrols of European descent (based on similarity to CEPH reference genotypes). (2015).
36. Watanabe, K., Taskesen, E., Bochoven, A. van & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications* **8**, 1826 (2017).
37. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).
38. Paaby, A. B. & Rockman, M. V. The many faces of pleiotropy. *Trends Genet* **29**, 66–73 (2013).
39. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *The American Journal of Human Genetics* **101**, 737–751 (2017).
40. Moon, A. L., Haan, N., Wilkinson, L. S., Thomas, K. L. & Hall, J. CACNA1C: Association With Psychiatric Disorders, Behavior, and Neurogenesis. *Schizophr Bull* **44**, 958–965 (2018).
41. Cross-Disorder Group of the Psychiatric Genomics Consortium]. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet* **381**, 1371–1379 (2013).
42. Mei, L. & Nave, K.-A. Neuregulin-ERBB signaling in nervous system development and neuropsychiatric diseases. *Neuron* **83**, 27–49 (2014).
43. Morris, K. V. & Mattick, J. S. The rise of regulatory RNA. *Nature Reviews Genetics* **15**, 423–437 (2014).
44. Liu, H. *et al.* Nuclear functions of mammalian MicroRNAs in gene regulation, immunity and cancer. *Molecular Cancer* **17**, 64 (2018).
45. Liu, X., Li, Y. I. & Pritchard, J. K. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell* **177**, 1022-1034.e6 (2019).
46. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45**, 984–994 (2013).
47. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–7 (2014).
48. Greven, C. U., Harlaar, N., Dale, P. S. & Plomin, R. Genetic Overlap between ADHD Symptoms and Reading is largely Driven by Inattentiveness rather than Hyperactivity-Impulsivity. *J Can Acad Child Adolesc Psychiatry* **20**, 6–14 (2011).

49. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* **50**, 1112–1121 (2018).
50. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics* **50**, 668 (2018).
51. Ruderfer, D. M. *et al.* Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell* **173**, 1705-1715.e16 (2018).
52. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* **42**, 97–110 (2013).
53. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111–127 (2013).
54. Power, C. & Elliott, J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol* **35**, 34–41 (2006).
55. Satterthwaite, T. D. *et al.* The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *Neuroimage* **124**, 1115–1119 (2016).
56. Glessner, J. T. et al. Strong synaptic transmission impact by copy number variations in schizophrenia. Proc. Natl. Acad. Sci. 107, 10584–10589 (2010)
57. *TOWRE, Test of Word Reading Efficiency: Examiner's Manual*. (PRO-ED, 1999).
58. Gur, R. C. *et al.* Neurocognitive growth charting in psychosis spectrum youths. *JAMA Psychiatry* **71**, 366–374 (2014).
59. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
60. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906–913 (2007).
61. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
62. Pedersen, C. B. *et al.* The iPSYCH2012 case–cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Molecular Psychiatry* **23**, 6–14 (2018).
63. *International Statistical Classification of Diseases and Related Health Problems*. vol. 10th revision (World Health Organization, 2010).
64. Mors, O., Perto, G. P. & Mortensen, P. B. The Danish Psychiatric Central Research Register. *Scand J Public Health* **39**, 54–57 (2011).
65. Lord, C., Rutter, M. & Le Couteur, A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* **24**, 659–685 (1994).
66. Lord, C. *et al.* The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord* **30**, 205–223 (2000).
67. Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* **51**, 793–803 (2019).
68. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* **47**, 291–5 (2015).
69. Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J. & Rohde, L. A. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am J Psychiatry* **164**, 942–948 (2007).
70. Kupfer, D. J., Frank, E. & Phillips, M. L. Major depressive disorder: new clinical, neurobiological, and treatment perspectives. *The Lancet* **379**, 1045–1055 (2012).
71. Owen, M. J., Sawa, A. & Mortensen, P. B. Schizophrenia. *The Lancet* **388**, 86–97 (2016).
72. Merikangas, K. R. *et al.* Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch. Gen. Psychiatry* **68**, 241–251 (2011).
73. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236–1241 (2015).
74. Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology* **11**, e1004219 (2015).
75. Neale, M., Boker, S., Xie, G. & Maes, H. H. M. *Mx: Statistical modeling*. (Department of Psychiatry, 2006).

76. St Pourcain, B. *et al.* Developmental changes within the genetic architecture of social communication behaviour: A multivariate study of genetic variance in unrelated individuals. *Biological Psychiatry* **83**, 598–606 (2017).

Supplementary Materials

Supplementary Methods

Web resources

Brainspan: https://www.brainspan.org/
CTG lab: https://ctg.cncr.nl/software/summary_statistics
FUMA: https://fuma.ctglab.nl/
iPSYCH: https://ipsych.au.dk/downloads/
LDSC: https://github.com/bulik/ldsc
PLINK: https://www.cog-genomics.org/plink2
PGC: https://www.med.unc.edu/pgc/
R: https://www.r-project.org/
SSGAC: https://www.thessgac.org/

Supplementary Tables

Supplementary Table 1: Sample description for MDD, SCZ, BD, general intelligence and reading

Source	GWAS	Consortium	Imputation reference panel	N
Clinical sample	MDD[1]	PGC	1000 Genomes multi-ancestry	173,005 (cases=59,851)
	SCZ[2]	PGC	1000 Genomes phase 3	65,967 (cases=33,426)
	BD[2]	PGC	1000 Genomes phase 3	41,653 (cases=20,129)
Population sample	Intelligence[3]	CTG lab	HRC#	279,930
	Reading	-	HRC	13,027

All individuals were of European descent. # Predominantly HRC, see Savage et al.[3] Abbreviations: BD, Bipolar Disorder; CTG, Complex Trait Genetics lab; HRC, Haplotype Reference Consortium; MDD, Major Depressive Disorder; PGC, Psychiatric Genomics Consortium; SCZ, Schizophrenia

Supplementary Table 2: SNP-heritability estimates

Phenotype	Sample	SNP-h^2 (SE)	λ_{GC}	Intercept (SE)
ASD	ASD(iPSYCH)	0.12(0.01)	1.15	1.01(0.01)
	ASD(iPSYCH, woADHD)	0.13(0.01)	1.13	1.01(0.01)
	ASD(PGC)	0.26(0.03)	1.05	0.97(0.01)
ADHD	ADHD(iPSYCH)	0.26(0.02)	1.23	1.03(0.01)
MDD	MDD(PGC)	0.095(0.01)	1.24	0.99(0.01)
SCZ	SCZ(PGC)	0.24(0.01)	1.50	1.05(0.01)
BD	BD(PGC)	0.18(0.01)	1.27	1.02(0.01)
Years-of-schooling	EA(SSGAC)	0.11(0.003)	2.10	1.03(0.01)
General intelligence	Intelligence(CTG)	0.18(0.006)	1.75	1.08(0.01)
Reading	Reading	0.19(0.04)	1.06	1.02(0.01)

SNP-heritability (SNP-h^2) was estimated with LDSC regression analysis[4]. SNP-h^2 estimates for EA, general intelligence and reading were calculated on the observed scale and for psychiatric disorders on a liability scale assuming a population prevalence of 0.012 (ASD), 0.05 (ADHD), 0.162 (MDD), 0.007 (SCZ) and 0.006 (BD). Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; BD, Bipolar Disorder; EA; educational attainment; CTG, Complex Trait Genetics lab; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; MDD, Major Depressive Disorder; PGC, Psychiatric Genomics Consortium; SCZ, Schizophrenia; SSGAC, Social Science Genetic Association Consortium; λ_{GC}, lambda GC; woADHD, without ADHD

Supplementary Table 3: Genetic correlations among psychiatric disorders

Sample 1	Sample 2	r_g (SE)	*P*
ASD(iPSYCH, woADHD)	ASD(PGC)	0.84(0.11)	$<10^{-10}$
	ADHD(iPSYCH)	0.30 (0.06)	$2x10^{-7}$
	MDD(PGC)	0.43 (0.05)	$<10^{-10}$
	SCZ(PGC)	0.21 (0.06)	10^{-4}
	BD(PGC)	0.18 (0.06)	0.001
ASD(PGC)	ADHD(iPSYCH)	0.04(0.08)	0.61
	MDD(PGC)	0.12(0.05)	0.017
	SCZ(PGC)	0.20(0.06)	0.001
	BD(PGC)	0.12(0.06)	0.063
ADHD(iPSYCH)	MDD(PGC)	0.55 (0.04)	$<10^{-10}$
	SCZ(PGC)	0.12 (0.04)	0.004
	BD(PGC)	0.12 (0.05)	0.007

Genetic correlations (r_g) among psychiatric disorder samples were estimated using summary statistics and unconstrained LD score correlation[5]. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; BD, Bipolar Disorder; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; MDD; Major Depressive Disorder; PGC, Psychiatric Genomics Consortium; r_g, genetic correlation; SCZ, Schizophrenia; woADHD, without ADHD

Supplementary Table 4: Genetic correlations of psychiatric disorders with educational attainment

Sample 1	Sample 2	r_g (SE)	*P*
EA(SSGAC)	ASD(iPSYCH)	0.16(0.03)	$4x10^{-7}$
	ASD(iPSYCH, woADHD)	0.23(0.03)	$<10^{-10}$
	ASD(PGC)	0.28(0.03)	$<10^{-10}$
	ADHD(iPSYCH)	-0.49(0.03)	$<10^{-10}$
	MDD(PGC)	-0.22 (0.03)	$<10^{-10}$
	SCZ(PGC)	0.07 (0.02)	0.003
	BD(PGC)	0.18 (0.02)	$<10^{-10}$

Genetic correlations (r_g) were estimated using summary statistics and unconstrained LD score correlation[5]. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; BD, Bipolar Disorder; EA; educational attainment; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; MDD, Major Depressive Disorder; PGC, Psychiatric Genomics Consortium; r_g, genetic correlation; SCZ, Schizophrenia; SSGAC, Social Science Genetic Association Consortium; woADHD, without ADHD.

Supplementary Table 5: Genetic correlations of ASD and ADHD with general intelligence and reading

Sample 1	Sample 2	r_g (SE)	*P*
Intelligence(CTG)	ASD(iPSYCH)	0.20(0.04)	10^{-8}
	ASD(iPSYCH, woADHD)	0.25(0.04)	$<10^{-10}$
	ASD(PGC)	0.19(0.04)	$2x10^{-7}$
	ADHD(iPSYCH)	-0.33(0.03)	$<10^{-10}$
Reading	ASD(iPSYCH)	0.11(0.11)	0.32
	ASD(iPSYCH, woADHD)	0.14(0.11)	0.20
	ASD(PGC)	0.32(0.11)	0.004
	ADHD(iPSYCH)	-0.45(0.10)	$8x10^{-6}$

Genetic correlations (r_g) were estimated using summary statistics and unconstrained LD score correlation[5]. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; CTG, Complex Trait Genetics lab; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; PGC, Psychiatric Genomics Consortium; rg, genetic correlation; woADHD, without ADHD;

Supplementary Table 6: Discovery ASD-MVR ($5x10^{-8}<P_{thr}<0.5$)

Variant selection			Cross-disorder	Intercept		ASD-specific effects		ADHD cross-disorder effects	
Gi	N_{SNPs}	P_{thr}		$\hat{\theta}_{0*}$(SE)	*P*	$\hat{\theta}_{ASD}$(SE)	*P*	$\hat{\theta}_{\mp ADHD}$(SE)	*P*
ASD (iPSYCH, woADHD)	3*	$5x10^{-8}$	ADHD (iPSYCH)	NA	NA	NA	NA	NA	NA
	5	$5x10^{-7}$		-0.010(0.007)	0.28	0.10(0.065)	0.26	-0.015(0.028)	0.65
	26	$5x10^{-6}$		0.004(0.003)	0.16	-0.003(0.030)	0.93	-0.022(0.024)	0.38
	121	$5x10^{-5}$		0.001(0.001)	0.38	0.010(0.013)	0.44	-0.018(0.014)	0.22
	816	0.0005		$2x10^{-4}(4x10^{-4})$	0.54	0.015(0.005)	0.002	-0.064(0.006)	10^{-8}
	1,973	0.0015		$0.001(2x10^{-4})$	0.008	0.009(0.003)	0.002	-0.029(0.004)	$<10^{-10}$
	5,399	0.005		$5x10^{-4}(10^{-4})$	$3x10^{-4}$	0.010(0.002)	$6x10^{-9}$	-0.028(0.002)	$<10^{-10}$
	35,921	0.05		$3x10^{-4}(4x10^{-5})$	$<10^{-10}$	0.007(0.001)	$<10^{-10}$	-0.022(0.001)	$<10^{-10}$
	62,589	0.1		$2x10^{-4}(3x10^{-5})$	$2x10^{-10}$	0.007(0.001)	$<10^{-10}$	-0.020(0.001)	$<10^{-10}$
	134,210	0.3		$1x\text{-}10^{-4}(2x10^{-5})$	$3x10^{-9}$	$0.007(4x10^{-4})$	$<10^{-10}$	$-0.018(4x10^{-4})$	$<10^{-10}$
	185,632	0.5		$10^{-4}(2x10^{-5})$	$<10^{-10}$	$0.007(4x10^{-4})$	$<10^{-10}$	$-0.017(3x10^{-4})$	$<10^{-10}$

Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and ADHD cross-disorder effect $\hat{\theta}_{\mp ADHD}$ on EA as fitted with ASD-MVR (Supplementary Figure 1a). Sets of independent ASD (Gi) variants were selected from ASD(iPSYCH, woADHD). SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. The multiple testing threshold is $P<0.0023$. *MVR analyses for ASD-related variants ($P_{thr}<5x10^{-8}$) did not converge. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; woADHD, without ADHD.

Supplementary Table 7: Discovery ADHD-MVR ($5x10^{-8}<P_{thr}<0.5$)

Variant selection			Cross-disorder	Intercept		ADHD-specific effects		ASD cross-disorder effects	
Gj	N_{SNPs}	P_{thr}		$\hat{\theta}_{0\#}$(SE)	*P*	$\hat{\theta}_{ADHD}$(SE)	*P*	$\hat{\theta}_{\#ASD}$(SE)	*P*
ADHD (iPSYCH)	10	$5x10^{-8}$	ASD (iPSYCH, woADHD)	-0.006(0.017)	0.76	-0.068(0.16)	0.68	0.12(0.075)	0.16
	22	$5x10^{-7}$		-0.002(0.006)	0.70	-0.055(0.060)	0.37	0.061(0.044)	0.19
	61	$5x10^{-6}$		-0.003(0.002)	0.13	-0.048(0.024)	0.054	0.096(0.026)	$5x10^{-4}$
	242	$5x10^{-5}$		-0.003(0.001)	0.001	-0.013(0.011)	0.24	0.025(0.012)	0.036
	1,202	0.0005		-0.002($3x10^{-4}$)	$2x10^{-6}$	-0.015(0.004)	0.001	0.023(0.005)	$2x10^{-6}$
	2,717	0.0015		-0.002($2x10^{-4}$)	$<10^{-10}$	-0.012(0.003)	$4x10^{-5}$	0.022(0.003)	$<10^{-10}$
	6,781	0.005		-0.001(10^{-4})	$<10^{-10}$	-0.009(0.002)	$5x10^{-8}$	0.018(0.002)	$<10^{-10}$
	41,334	0.05		-0.001($4x10^{-5}$)	$<10^{-10}$	-0.009(0.001)	$<10^{-10}$	0.013(0.001)	$<10^{-10}$
	71,015	0.1		-0.001($3x10^{-5}$)	$<10^{-10}$	-0.009(0.001)	$<10^{-10}$	0.012(0.001)	$<10^{-10}$
	164,083	0.3		$-3x10^{-4}$($2x10^{-5}$)	$<10^{-10}$	-0.010($4x10^{-4}$)	$<10^{-10}$	0.010($3x10^{-4}$)	$<10^{-10}$
	234,530	0.5		$-2x10^{-4}$(10^{-5})	$<10^{-10}$	-0.011($4x10^{-4}$)	$<10^{-10}$	0.009($3x10^{-4}$)	$<10^{-10}$

Estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effect $\hat{\theta}_{\#ASD}$ on EA as fitted with ADHD-MVR (Supplementary Figure 1b). Sets of independent ADHD (G_j) genetic variants were selected from ADHD(iPSYCH) GWAS statistics. SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. The multiple testing threshold is $P<0.0023$. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; woADHD, without ADHD

Supplementary Table 8: Discovery ASD-MVR and ADHD-MVR (P_{thr}<0.0015; P_{thr} <0.05)

ASD-MVR

Variant selection			Cross-disorder	Intercept		ASD-specific effects		ADHD cross-disorder effects		VIF	MVR to UVR model fit comparison	
Gi	N_{SNPs}	P_{thr}		θ_{0*}(SE)	*P*	θ_{ASD}(SE)	*P*	θ_{*ADHD}(SE)	*P*		ΔR^2 (%)	LRT-χ^2(df=1), *P*
ASD (iPSYCH, woADHD)	1,973	0.0015	ADHD (iPSYCH)	0.001(2x10^{-4})	0.008	-5x10^{-6}(0.003)	1.00	-		-		-
				0.001(2x10^{-4})	0.008	0.009(0.003)	0.002	-0.029(0.004)	<10^{-10}	1.19	3.0	167.06, <10^{-10}
	35,921	0.05		3x10^{-4}(4x10^{-5})	10^{-9}	0.001(0.001)	0.17	-		-		-
				3x10^{-4}(4x10^{-5})	<10^{-10}	0.007(0.001)	<10^{-10}	-0.022(0.001)	<10^{-10}	1.11	2.1	1533.8, <10^{-10}

ADHD-MVR

Variant selection			Cross-disorder	Intercept		ADHD-specific effects		ASD cross-disorder effects		VIF	MVR to UVR model fit comparison	
Gj	N_{SNPs}	P_{thr}		$\theta_{0\#}$(SE)	*P*	θ_{ADHD}(SE)	*P*	θ_{*ASD}(SE)	*P*		ΔR^2 (%)	LRT-χ^2(df=1), *P*
ADHD (iPSYCH)	2,717	0.0015	ASD (iPSYCH, woADHD)	-0.001(2x10^{-4})	<10^{-10}	-0.004(0.003)	0.16	-		-		-
				-0.002(2x10^{-4})	<10^{-10}	-0.012(0.003)	4x10^{-5}	0.022(0.003)	<10^{-10}	1.18	1.9	154.24, <10^{-10}
	41,334	0.05		-0.001(4x10^{-5})	<10^{-10}	-0.004(0.001)	<10^{-10}	-		-		-
				-0.001(4x10^{-5})	<10^{-10}	-0.009(0.001)	<10^{-10}	0.013(0.001)	<10^{-10}	1.11	0.9	744.05, <10^{-10}

Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and ADHD cross-disorder effect $\hat{\theta}_{*ADHD}$ on EA as fitted with ASD-MVR (Figure 2a, Supplementary Figure 1a). Estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effect $\hat{\theta}_{*ASD}$ on EA as fitted with ADHD-MVR (Figure 2b, Supplementary Figure 1b). Sets of independent ASD (G_i) and ADHD (G_j) genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively and are shown for two *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. The model fit of MVRs and UVRs was compared with likelihood-ratio tests (LRTs). The multiple testing threshold is *P*<0.0023. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; df, degrees of freedom; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; LRT, likelihood-ratio test; MVR, multivariable regression; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; VIF, variance inflation factor; woADHD, without ADHD; UVR, univariable regression; ΔR^2, difference in Regression R^2 between MVR and UVR

Supplementary Table 9: Follow-up ASD-MVR and ADHD-MVR (P_{thr}<0.0015; P_{thr} <0.05), concordant variants

ASD-MVR

Variant selection			Cross-disorder	Intercept		ASD-specific effects		ADHD cross-disorder effects		VIF	MVR to UVR model fit comparison			
											ASD-specific UVR		ADHD cross-disorder UVR	
Concordant variants Gi	N_{SNPs}	P_{thr}		θ_{0*}(SE)	P	θ_{ASD}(SE)	P	θ_{*ADHD}(SE)	P		ΔR^2(%)	LRT-χ^2(df=1), P	ΔR^2(%)	LRT-χ^2(df=1), P
ASD (iPSYCH, woADHD)	1,716	0.0015	ADHD (iPSYCH)	$4x10^{-4}(2x10^{-4})$	0.15	0.001(0.003)	0.76	-		-	-			
				$0.001(2x10^{-4})$	$4x10^{-10}$	-		-0.018(0.004)	$9x10^{-6}$	-	-			
				$4x10^{-4}(2x10^{-4})$	0.12	0.011(0.003)	0.001	-0.026(0.005)	$5x10^{-8}$	1.40	1.7	79.67, $4x10^{-8}$	0.6	27.06, 0.001
	28,086	0.05		$3x10^{-5}(5x10^{-5})$	0.49	0.001(0.001)	0.11	-		-	-			
				$5x10^{-4}(3x10^{-5})$	$<1x10^{-10}$			-0.014(0.001)	$<10^{-10}$	-	-			
				$10^{-4}(5x10^{-5})$	0.054	0.009(0.001)	$<10^{-10}$	-0.020(0.001)	$<10^{-10}$	1.36	1.1	579.37, $<10^{-10}$	0.4	203.75, $<10^{-10}$

ADHD-MVR

Variant selection			Cross-disorder	Intercept		ADHD-specific effects		ASD cross-disorder effects		VIF	MVR to UVR model fit comparison			
											ADHD-specific UVR		ASD cross-disorder UVR	
Concordant variants Gj	N_{SNPs}	P_{thr}		$\theta_{0\#}$(SE)	P	θ_{ADHD}(SE)	P	θ_{*ASD}(SE)	P		ΔR^2(%)	LRT-χ^2(df=1), P	ΔR^2(%)	LRT-χ^2(df=1), P
ADHD (iPSYCH)	2,382	0.0015	ASD (iPSYCH, woADHD)	$-0.001(2x10^{-4})$	$8x10^{-10}$	-0.004(0.003)	0.16	-		-	-			
				$-0.002(1x10^{-4})$	$<1x10^{-10}$	-		0.014(0.003)	$3x10^{-5}$	-	-			
				$-0.001(2x10^{-4})$	$2x10^{-10}$	-0.013(0.003)	$4x10^{-5}$	0.022(0.004)	10^{-8}	1.36	1.3	92.59, 10^{-8}	0.7	48.75, $4x10^{-5}$
	32,176	0.05		$-4x10^{-4}(5x10^{-5})$	$<10^{-10}$	-0.005(0.001)	$3x10^{-10}$	-		-	-			
				$-9x10^{-4}(3x10^{-5})$	$<1x10^{-10}$	-		$0.007(9x10^{-4})$	$<10^{-10}$	-	-			
				$-0.001(5x10^{-5})$	$<10^{-10}$	-0.011(0.001)	$<10^{-10}$	0.013(0.001)	$<10^{-10}$	1.36	0.5	339.16, $<10^{-10}$	0.5	290.13, $<10^{-10}$

Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and ADHD cross-disorder effect $\hat{\theta}_{\nexists ADHD}$ on EA as fitted with ASD-MVR (Supplementary Figure 1c). Estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effect $\hat{\theta}_{\nexists ASD}$ on EA as fitted with ADHD-MVR (Supplementary Figure 1d). Sets of independent ASD (G_i) and ADHD (G_j) genetic variants were selected at two *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05) from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively and restricted to SNPs with concordant allelic effects. SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. The model fit of MVRs and UVRs was compared with likelihood-ratio tests (LRTs). The multiple testing threshold is *P*<0.0023. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; df, degrees of freedom; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; LRT, likelihood-ratio test; MVR, multivariable regression; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; VIF, variance inflation factor; woADHD, without ADHD; UVR, univariable regression; ΔR^2, difference in Regression R^2 between UVR and MVR

Supplementary Table 10: Follow-up ASD-MVR and ADHD-MVR (P_{thr}<0.0015; P_{thr}<0.05), analyses with ASD(PGC) SNP estimates (independent variable)

ASD-MVR

Variant selection			Cross-disorder	Intercept		ASD-specific effects		ADHD cross-disorder effects		VIF	MVR to UVR model fit comparison	
Gi	N_{SNPs}	P_{thr}		θ_{0*}(SE)	*P*	θ_{ASD}(SE)	*P*	θ_{*ADHD}(SE)	*P*		ΔR^2(%)	LRT-χ^2(df=1), *P*
ASD (iPSYCH, ASDwoADHD) variants with ASD(PGC) estimates	1,886	0.0015	ADHD (iPSYCH)	$-6x10^{-5}(10^{-4})$	0.69	0.010(0.003)	0.004	-		-	-	
				$-5x10^{-5}(10^{-4})$	0.70	0.010(0.003)	0.003	-0.003(0.003)	0.31	1.00	0.05	2.98, 0.31
	33,845	0.05		$-10^{-5}(3x10^{-5})$	0.61	0.005(0.001)	$<10^{-10}$	-		-	-	
				$-9x10^{-6}(3x10^{-5})$	0.76	0.005(0.001)	$<10^{-10}$	-0.007(0.001)	$<10^{-10}$	1.00	0.41	285.63, $<10^{-10}$

ADHD-MVR

Variant selection			Cross-disorder	Intercept		ADHD-specific effects		ASD cross-disorder effects		VIF	MVR to UVR model fit comparison	
Gj	N_{SNPs}	P_{thr}		$\theta_{0\#}$(SE)	*P*	θ_{ADHD}(SE)	*P*	θ_{*ASD}(SE)	*P*		ΔR^2(%)	LRT-χ^2(df=1), *P*
ADHD (iPSYCH)	2,627	0.0015	ASD (PGC)	$-0.001(2x10^{-4})$	$<10^{-10}$	-0.004(0.003)	0.14	-		-	-	
				$-0.001(2x10^{-4})$	$<10^{-10}$	-0.004(0.003)	0.13	0.002(0.002)	0.35	1.00	0.03	2.62, 0.35
	38,656	0.05		$-0.001(4x10^{-5})$	$<10^{-10}$	-0.005(0.001)	$<10^{-10}$	-		-	-	
				$-0.001(4x10^{-5})$	$<10^{-10}$	-0.005(0.001)	$<10^{-10}$	$0.003(4x10^{-4})$	$<10^{-10}$	1.00	0.16	123.59, $<10^{-10}$

Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and ADHD cross-disorder effect $\hat{\theta}_{\nexists ADHD}$ on EA as fitted with ASD-MVR (Supplementary Figure 1e). Estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effect $\hat{\theta}_{\nexists ASD}$ on EA as fitted with ADHD-MVR (Supplementary Figure 1f). Sets of independent ASD (G_i) and ADHD (G_j) genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively and are shown for two *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(PGC), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. The model fit of MVRs and UVRs was compared with likelihood-ratio tests (LRTs). The multiple testing threshold is *P*<0.0125. Among the 1,973 SNPs selected at P_{thr}<0.0015 in ASD(iPSYCH, woADHD), 1,886 were available in ASD(PGC) and 49% had concordant effects. Among the 35,921 SNPs selected at P_{thr} <0.05 in ASD(iPSYCH, woADHD), 33,845 were available in ASD(PGC) and 52% had concordant effects. Among the 2,717 SNPs selected at threshold P_{thr}<0.0015 in ADHD (iPSYCH), 2,627 were available in ASD(PGC) and 51% had concordant effects. Among the 35,921 SNPs selected at threshold P_{thr} <0.05 in ADHD(iPSYCH), 41,334 were available in ASD(PGC) and 52% had concordant effects. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; df, degrees of freedom; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; LRT, likelihood-ratio test; MVR, multivariable regression; N_{SNPs}, number of SNPs; PGC, Psychiatric Genomics Consortium; P_{thr}, *P*-value threshold; VIF, variance inflation factor; UVR, univariable regression; ΔR^2, difference in Regression R^2 between MVR and UVR

Supplementary Table 11: Follow-up ASD-MVR and ADHD-MVR (P_{thr}<0.0015; P_{thr}<0.05), analyses with general intelligence SNP estimates (dependent variable)

ASD-MVR

Variant selection			Cross-disorder	Intercept		ASD-specific effects		ADHD cross-disorder effects		VIF	MVR to UVR model fit comparison	
Gi	N_{SNPs}	P_{thr}		θ_{0*}(SE)	*P*	θ_{ASD}(SE)	*P*	θ_{*ADHD}(SE)	*P*		ΔR^2 (%)	LRT-χ^2(df=1), *P*
ASD (iPSYCH, woADHD)	1,904	0.0015	ADHD (iPSYCH)	0.001(3x10^{-4})	0.01	-0.002(0.004)	0.59	-		-		-
				0.001(3x10^{-4})	0.008	0.008(0.004)	0.05	-0.033(0.005)	5x10^{-10}	1.18	2.02	79.80, 4x10^{-10}
	34,626	0.05		4x10^{-4}(7x10^{-5})	5x10^{-9}	5x10^{-5}(0.001)	0.96	-		-		-
				4x10^{-4}(7x10^{-5})	5x10^{-10}	0.006(0.001)	2x10^{-8}	-0.021(0.001)	<10^{-10}	1.11	0.91	509.13, <10^{-10}

ADHD-MVR

Variant selection			Cross-disorder	Intercept		ADHD-specific effects		ASD cross-disorder effects		VIF	MVR to UVR model fit comparison	
Gj	N_{SNPs}	P_{thr}		$\theta_{0\#}$(SE)	*P*	θ_{ADHD}(SE)	*P*	θ_{*ASD}(SE)	*P*		ΔR^2 (%)	LRT-χ^2(df=1), *P*
ADHD (iPSYCH)	2,614	0.0015	ASD (iPSYCH, woADHD)	-0.001(3x10^{-4})	0.001	-0.006(0.004)	0.09	-		-		-
				-0.001(3x10^{-4})	0.001	-0.012(0.004)	0.002	0.016(0.004)	2x10^{-4}	1.18	0.55	29.57, 10^{-4}
	39,767	0.05		-6x10^{-4}(6x10^{-5})	<10^{-10}	-0.003 (0.001)	0.003	-		-		-
				-0.001(6x10^{-5})	<10^{-10}	-0.007(0.001)	3x10^{-10}	0.012(0.001)	<10^{-10}	1.10	0.33	211.34, <10^{-10}

Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and ADHD cross-disorder effect $\hat{\theta}_{\nexists ADHD}$ on general intelligence as fitted with ASD-MVR (Supplementary Figure 1g). Estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effect $\hat{\theta}_{\nexists ASD}$ on general intelligence as fitted with ADHD-MVR (Supplementary Figure 1h). Sets of independent ASD (G_i) and ADHD (G_j) genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively and are shown for two *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{Intelligence}$) were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and intelligence (CTG lab) GWAS statistics respectively. All MVR effects are presented as standard deviation-related changes in general intelligence per increase in log-odds of ASD or ADHD liability. The model fit of MVRs and UVRs was compared with likelihood-ratio tests (LRTs). The multiple testing threshold is *P*<0.006. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; df, degrees of freedom; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; LRT, likelihood-ratio test; MVR, multivariable regression; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; woADHD, without ADHD; UVR, univariable regression; ΔR^2, difference in Regression R^2 between MVR and UVR

Supplementary Table 12: Follow-up ASD-MVR and ADHD-MVR (P_{thr}<0.0015; P_{thr}<0.05), analyses with reading SNP estimates (dependent variable)

ASD-MVR

Variant selection			Cross-disorder	Intercept		ASD-specific effects		ADHD cross-disorder effects		VIF	MVR to UVR model fit comparison	
Gi	N_{SNPs}	P_{thr}		θ_{0*}(SE)	*P*	θ_{ASD}(SE)	*P*	θ_{*ADHD}(SE)	*P*		ΔR^2 (%)	LRT-χ^2(df=1), *P*
ASD (iPSYCH, woADHD)	1,932	0.0015	ADHD (iPSYCH)	-0.002(0.001)	0.13	0.029(0.013)	0.027	-		-		-
				-0.002(0.001)	0.13	0.035(0.014)	0.014	-0.019(0.017)	0.27	1.17	0.06	1.32, 0.27
	34,556	0.05		$7x10^{-5}$($3x10^{-4}$)	0.79	0.005(0.004)	0.18	-		-		-
				$1x10^{-4}$($2x10^{-4}$)	0.70	0.016(0.004)	$2x10^{-4}$	-0.036(0.004)	$<1x10^{-10}$	1.10	0.19	70.36, $<1x10^{-10}$

ADHD-MVR

Variant selection			Cross-disorder	Intercept		ADHD-specific effects		ASD cross-disorder effects		VIF	MVR to UVR model fit comparison	
Gj	N_{SNPs}	P_{thr}		$\theta_{0\#}$(SE)	*P*	θ_{ADHD}(SE)	*P*	θ_{*ASD}(SE)	*P*		ΔR^2 (%)	LRT-χ^2(df=1), *P*
ADHD (iPSYCH)	2,638	0.0015	ASD (iPSYCH, woADHD)	-0.002(0.001)	0.03	-0.003(0.013)	0.81	-		-		-
				-0.002(0.001)	0.03	-0.005(0.014)	0.73	0.005(0.014)	0.73	1.17	0.004	0.13, 0.73
	39,336	0.05		$-0.001(2x10^{-4})$	0.03	-0.013(0.004)	0.001	-		-		-
				$-0.001(2x10^{-4})$	0.03	-0.016(0.004)	$9x10^{-5}$	0.008(0.004)	0.03	1.10	0.01	5.04, 0.029

Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and ADHD cross-disorder effect $\hat{\theta}_{\ddagger ADHD}$ on reading as fitted with ASD-MVR (Supplementary Figure 1g). Estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effect $\hat{\theta}_{\ddagger ASD}$ on reading as fitted with ADHD-MVR (Supplementary Figure 1h). Sets of independent ASD (G_i) and ADHD (G_j) genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively and are shown for two *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and reading ($\hat{\beta}_{reading}$) were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and reading GWAS statistics respectively. All MVR effects are presented as standard deviation-related changes in reading ability per increase in log-odds of ASD or ADHD liability. The model fit of MVRs and UVRs was compared with likelihood-ratio tests (LRTs). The multiple testing threshold is *P*<0.006. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; df, degrees of freedom; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; LRT, likelihood-ratio test; MVR, multivariable regression; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; woADHD, without ADHD; UVR, univariable regression; ΔR^2, difference in Regression R^2 between MVR and UVR

Supplementary Table 13: Follow-up ASD-MVR (P_{thr} ASD<0.0015 and 0.0015 <P_{thr} ADHD<0.5), analyses with variants meeting joint ASD and ADHD selection criteria

Variant sets (Gij)				Intercept		ASD-specific effects		ADHD cross-disorder effects	
P_{thr} ASD (iPSYCH, woADHD)	N_{SNPs}	% of discovery variant set	P_{thr} ADHD (iPSYCH)	$\hat{\theta}_{0*}$(SE)	*P*	$\hat{\theta}_{ASD}$(SE)	*P*	$\hat{\theta}_{\mp ADHD}$(SE)	*P*
0.0015	83	4.21	0.0015	-0.003(0.001)	0.010	0.15(0.025)	10^{-7}	-0.15(0.025)	10^{-7}
	134	6.79	0.005	-0.002(0.001)	0.15	0.12(0.022)	$6x10^{-7}$	-0.13(0.024)	$2x10^{-7}$
	393	19.92	0.05	-0.001(0.001)	0.28	0.045(0.009)	$4x10^{-7}$	-0.063(0.011)	$7x10^{-9}$
	536	27.17	0.1	$-4x10^{-4}(5x10^{-4})$	0.40	0.040(0.007)	$2x10^{-8}$	-0.059(0.009)	$<10^{-10}$
	805	40.80	0.3	$5x10^{-4}(4x10^{-4})$	0.21	0.021(0.005)	$4x10^{-5}$	-0.044(0.006)	$<10^{-10}$
	914	46.33	0.5	$0.001(3x10^{-4})$	0.12	0.017(0.004)	10^{-4}	-0.039(0.006)	$<10^{-10}$
	1973[+]	100	1	$0.001(2x10^{-4})$	0.008	0.009(0.003)	0.002	-0.029(0.004)	$<10^{-10}$

Estimated ASD-specific effects $\hat{\theta}_{ASD}$ and ADHD cross-disorder effects $\hat{\theta}_{\mp ADHD}$ on EA as fitted with ASD-MVR (Figure 3a, Supplementary Figure 1i) using SNP subsets (Gij) fulfilling joint ASD and ADHD selection criteria. Independent genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics. Among ASD variants of the discovery variant set ([+], Gi, P_{thr}<0.0015), we identified those SNPs, which were also associated with ADHD (6 thresholds: 0.0015 ≤ADHD P_{thr} <0.5). ADHD variants were identified as tagged, if a SNP within was within 500kb and LD-r^2≥0.6 of a SNP in the ASD discovery variant set. SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. The multiple testing threshold is *P*<0.0023. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; woADHD, without ADHD; [+]Discovery ASD variant set (P_{thr} ASD<0.0015)

Supplementary Table 14: Follow-up ADHD-MVR (P_{thr} ADHD<0.0015 and 0.0015<P_{thr}ASD<0.5), analyses with variants meeting joint ASD and ADHD selection criteria

Variant sets (Gji)				Intercept		ADHD-specific effects		ASD cross-disorder effects	
P_{thr} ADHD (iPSYCH)	N_{SNPs}	% of discovery variant set	P_{thr} ASD (iPSYCH, woADHD)	$\hat{\theta}_{0\#}$(SE)	*P*	$\hat{\theta}_{ADHD}$(SE)	*P*	$\hat{\theta}_{\mp ASD}$(SE)	*P*
0.0015	83	3.05	0.0015	-0.003(0.001)	0.038	-0.10(0.026)	$2x10^{-4}$	0.12(0.022)	$9x10^{-7}$
	145	5.34	0.005	-0.002(0.001)	0.050	-0.070(0.022)	0.002	0.086(0.019)	10^{-5}
	473	17.41	0.05	-0.001(0.001)	0.044	-0.026(0.010)	0.010	0.032(0.010)	0.001
	638	23.48	0.1	-0.001($4x10^{-4}$)	0.047	-0.023(0.008)	0.003	0.026(0.008)	0.001
	937	34.49	0.3	-0.001($3x10^{-4}$)	0.003	-0.018(0.005)	$5x10^{-4}$	0.023(0.006)	$5x10^{-5}$
	1020	37.54	0.5	-0.001($3x10^{-4}$)	0.005	-0.017(0.005)	$3x10^{-4}$	0.018(0.005)	$3x10^{-4}$
	2717[+]	100	1	-0.002($2x10^{-4}$)	$<10^{-10}$	-0.012(0.003)	$4x10^{-5}$	0.022(0.003)	$<10^{-10}$

Estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and ASD cross-disorder effects $\hat{\theta}_{\mp ASD}$ on EA as fitted with ADHD-MVR (Figure 3b, Supplementary Figure 1j) using SNP subsets (Gji) fulfilling joint ASD and ADHD selection criteria. Independent genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics. Among ADHD variants of the discovery variant set ([+], Gj, P_{thr}<0.0015), we identified those SNPs, which were also associated with ASD (6 thresholds: 0.0015≤ ASD P_{thr} <0.5). ASD variants were identified as tagged, if a SNP was within 500kb and LD-r^2≥0.6 of a SNP in the ADHD discovery variant set. SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or ADHD liability. The multiple testing threshold is *P*<0.0023. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; woADHD, without ADHD; [+]Discovery ADHD variant set (P_{thr} ADHD<0.0015)

Supplementary Table 15: SNP estimates for variants meeting joint ASD and ADHD selection criteria (P_{thr} ADHD<0.0015 and P_{thr} ASD<0.0015)

ASD-MVR: ASD variants tagged by ADHD variants							ADHD-MVR: ADHD variants tagged by ASD variants							Mapped gene(s)
SNP	A1	A2	ASD(iPSYCH, woADHD)		ADHD(iPSYCH)		SNP	A1	A2	ASD(iPSYCH, woADHD)		ADHD(iPSYCH)		
			β_{ASD} (SE)	*P*	β_{ADHD} (SE)	*P*				β_{ASD} (SE)	*P*	β_{ADHD} (SE)	*P*	
rs10478063	T	G	0.06(0.02)	$7x10^{-4}$	0.06(0.02)	$8x10^{-4}$	rs6869021	T	C	0.06(0.02)	10^{-3}	0.07(0.02)	$7x10^{-5}$	*STARD4-AS1*
rs10503223	C	G	0.10(0.03)	$4x10^{-4}$	0.11(0.03)	$4x10^{-4}$	rs10503223	C	G	0.10(0.03)	$4x10^{-4}$	0.11(0.03)	$4x10^{-4}$	*CSMD1*
rs1052607	G	A	0.11(0.03)	$7x10^{-4}$	0.12(0.03)	$4x10^{-5}$	rs77881576	A	G	0.11(0.03)	$8x10^{-4}$	0.13(0.03)	$2x10^{-5}$	*MAST2, LOC110117498-PIK3R, PIK3R3*
rs10764449	C	T	0.07(0.02)	$3x10^{-4}$	0.06(0.02)	10^{-3}	rs10764449	C	T	0.07(0.02)	$3x10^{-4}$	0.06(0.02)	10^{-3}	*KIAA1217*
rs10932543	A	G	0.09(0.02)	$4x10^{-6}$	0.07(0.02)	10^{-4}	rs10932543	A	G	0.09(0.02)	$4x10^{-6}$	0.07(0.02)	10^{-4}	*VWC2L*
rs11124404	G	T	0.21(0.06)	$6x10^{-4}$	0.20(0.05)	$3x10^{-4}$	rs12328829	T	C	0.18(0.06)	$2x10^{-3}$	0.21(0.05)	$4x10^{-5}$	
rs112635299	T	G	0.17(0.05)	$7x10^{-4}$	0.15(0.05)	10^{-3}	rs112635299	T	G	0.17(0.05)	$7x10^{-4}$	0.15(0.05)	10^{-3}	
rs1133878	A	G	0.06(0.02)	10^{-3}	0.04(0.02)	10^{-2}	rs2623245	T	A	0.05(0.02)	$4x10^{-3}$	0.05(0.02)	10^{-3}	*SCG3, TMOD3*
rs114068758	T	C	0.12(0.04)	10^{-3}	0.09(0.03)	$4x10^{-3}$	rs150844750	C	T	0.10(0.04)	$6x10^{-3}$	0.12(0.03)	$4x10^{-4}$	*EDAR, MIR4435-1, MIR4435-2, MIR4771-1, MIR4771-2, RANBP2*
rs114457163	C	T	0.30(0.07)	$4x10^{-6}$	0.21(0.06)	$5x10^{-4}$	rs114457163	C	T	0.30(0.07)	$4x10^{-6}$	0.21(0.06)	$5x10^{-4}$	
rs11656656	A	G	0.07(0.02)	10^{-4}	0.06(0.02)	$3x10^{-4}$	rs8074682	G	A	0.07(0.02)	10^{-4}	0.06(0.02)	$2x10^{-4}$	*MGAT5B*
rs116687082	T	C	0.16(0.05)	$4x10^{-4}$	0.16(0.04)	$8x10^{-5}$	rs116687082	T	C	0.16(0.05)	$4x10^{-4}$	0.16(0.04)	$8x10^{-5}$	
rs118058985	T	C	0.18(0.04)	$3x10^{-6}$	0.14(0.03)	$5x10^{-5}$	rs118058985	T	C	0.18(0.04)	$3x10^{-6}$	0.14(0.03)	$5x10^{-5}$	
rs1198850	C	T	0.08(0.02)	$3x10^{-4}$	0.07(0.02)	10^{-3}	rs1198850	C	T	0.08(0.02)	$3x10^{-4}$	0.07(0.02)	10^{-3}	*ATP6V1C2*
rs12142993	G	A	0.22(0.06)	$8x10^{-5}$	0.15(0.05)	10^{-3}	rs12142993	G	A	0.22(0.06)	$8x10^{-5}$	0.15(0.05)	10^{-3}	*FAM72C, FMO2, SRGAP2D*
rs12482859	A	T	0.10(0.03)	10^{-3}	0.09(0.03)	$2x10^{-3}$	rs12482860	C	T	0.09(0.03)	$5x10^{-3}$	0.09(0.03)	10^{-3}	
rs12900962	C	T	0.08(0.03)	10^{-3}	0.07(0.02)	$3x10^{-3}$	rs3794565	A	G	0.08(0.03)	$2x10^{-3}$	0.08(0.02)	10^{-3}	*MYO5A*
rs13207266	A	C	0.06(0.02)	$5x10^{-4}$	0.05(0.02)	$2x10^{-3}$	rs9406227	G	T	0.06(0.02)	10^{-3}	0.05(0.02)	10^{-3}	
rs13338042	C	A	0.18(0.04)	$4x10^{-5}$	0.13(0.04)	$2x10^{-3}$	rs28422934	G	C	0.18(0.04)	$5x10^{-5}$	0.13(0.04)	$8x10^{-4}$	
rs139028896	G	C	0.12(0.04)	10^{-3}	0.11(0.03)	$7x10^{-4}$	rs41300124	C	T	0.11(0.04)	$4x10^{-3}$	0.12(0.03)	$5x10^{-4}$	*CELA3A*
rs146374277	C	T	0.25(0.08)	10^{-3}	0.23(0.07)	$7x10^{-4}$	rs146374277	C	T	0.25(0.08)	10^{-3}	0.23(0.07)	$7x10^{-4}$	
rs148532090	T	C	0.29(0.07)	$3x10^{-5}$	0.28(0.06)	10^{-5}	rs148532090	T	C	0.29(0.07)	$3x10^{-5}$	0.28(0.06)	10^{-5}	*IP6K2*
rs1554380	A	G	0.10(0.03)	$2x10^{-4}$	0.08(0.03)	$2x10^{-3}$	rs17012945	G	A	0.10(0.03)	10^{-3}	0.09(0.03)	$4x10^{-4}$	
rs17477236	C	A	0.08(0.02)	10^{-4}	0.06(0.02)	10^{-3}	rs17477236	C	A	0.08(0.02)	10^{-4}	0.06(0.02)	10^{-3}	*VAV3*
rs184032510	C	T	0.20(0.06)	$6x10^{-4}$	0.14(0.05)	$9x10^{-3}$	rs141735644	T	C	0.21(0.06)	10^{-3}	0.21(0.06)	$4x10^{-4}$	
rs1985169	A	G	0.11(0.02)	$2x10^{-6}$	0.09(0.02)	$7x10^{-6}$	rs112337001	T	C	0.10(0.02)	10^{-6}	0.08(0.02)	$7x10^{-6}$	

rs202823	A	G	0.16(0.04)	$6x10^{-5}$	0.13(0.04)	$2x10^{-4}$	rs17664036	C	T	0.14(0.04)	10^{-4}	0.12(0.03)	10^{-4}	*XRN2, KIZ*
rs2228048	T	C	0.21(0.06)	10^{-3}	0.18(0.06)	$2x10^{-3}$	rs3773648	C	G	0.20(0.07)	$2x10^{-3}$	0.19(0.06)	10^{-3}	*TGFBR2*
rs2239113	G	A	0.07(0.02)	$9x10^{-4}$	0.07(0.02)	$2x10^{-4}$	rs4765959	A	T	0.06(0.02)	$2x10^{-3}$	0.07(0.02)	$9x10^{-5}$	*CACNA1C*
rs2328878	G	A	0.06(0.02)	10^{-3}	0.05(0.02)	$3x10^{-3}$	rs2744296	A	G	0.04(0.02)	10^{-2}	0.05(0.02)	$6x10^{-4}$	*CARMIL1*
rs2391769	G	A	0.09(0.02)	$5x10^{-7}$	0.09(0.02)	$8x10^{-8}$	rs2391769	G	A	0.09(0.02)	$5x10^{-7}$	0.09(0.02)	$8x10^{-8}$	
rs2635182	T	C	0.07(0.02)	$6x10^{-5}$	0.05(0.02)	$2x10^{-3}$	rs250274	T	G	0.06(0.02)	$2x10^{-4}$	0.06(0.02)	$2x10^{-4}$	
rs33823	T	C	0.07(0.02)	$3x10^{-5}$	0.04(0.02)	$6x10^{-3}$	rs57457691	C	T	0.05(0.02)	10^{-2}	0.05(0.02)	10^{-3}	*PEPD*
rs35357088	C	T	0.13(0.03)	$2x10^{-4}$	0.10(0.03)	10^{-3}	rs77529534	C	T	0.13(0.04)	$9x10^{-4}$	0.13(0.03)	$2x10^{-4}$	*TEK*
rs36092443	A	G	0.15(0.04)	$4x10^{-4}$	0.15(0.04)	$9x10^{-5}$	rs36092443	A	G	0.15(0.04)	$4x10^{-4}$	0.15(0.04)	$9x10^{-5}$	
rs3767602	G	A	0.13(0.04)	$4x10^{-4}$	0.11(0.03)	$8x10^{-4}$	rs11810580	C	T	0.11(0.04)	$3x10^{-3}$	0.11(0.03)	$7x10^{-4}$	*TMIGD3*
rs416223	C	A	0.06(0.02)	$5x10^{-4}$	0.07(0.02)	$2x10^{-6}$	rs325506	C	G	0.05(0.02)	$3x10^{-3}$	0.08(0.02)	$5x10^{-7}$	
rs4322805	A	G	0.06(0.02)	$2x10^{-4}$	0.06(0.02)	$3x10^{-4}$	rs4322805	A	G	0.06(0.02)	$2x10^{-4}$	0.06(0.02)	$3x10^{-4}$	
rs4609618	C	A	0.08(0.02)	$9x10^{-6}$	0.05(0.02)	$7x10^{-4}$	rs4609618	C	A	0.08(0.02)	$9x10^{-6}$	0.05(0.02)	$7x10^{-4}$	
rs4679557	A	G	0.06(0.02)	$8x10^{-4}$	0.05(0.02)	$9x10^{-4}$	rs6787365	A	G	0.06(0.02)	10^{-3}	0.06(0.02)	$5x10^{-4}$	
rs4916723	C	A	0.07(0.02)	10^{-4}	0.09(0.02)	$5x10^{-9}$	rs4916723	C	A	0.07(0.02)	10^{-4}	0.09(0.02)	$5x10^{-9}$	*LINC00461*
rs4952312	A	G	0.10(0.03)	$5x10^{-4}$	0.08(0.03)	$2x10^{-3}$	rs77384847	C	T	0.09(0.03)	$8x10^{-4}$	0.08(0.03)	10^{-3}	*LINC00486*
rs4981706	A	T	0.07(0.02)	$6x10^{-4}$	0.07(0.02)	$5x10^{-4}$	rs4981706	A	T	0.07(0.02)	$6x10^{-4}$	0.07(0.02)	$5x10^{-4}$	
rs501371	G	A	0.18(0.05)	10^{-4}	0.14(0.04)	$7x10^{-4}$	rs1277718	C	G	0.18(0.05)	10^{-4}	0.14(0.04)	$7x10^{-4}$	
rs529507	G	A	0.14(0.02)	$3x10^{-8}$	0.07(0.02)	$2x10^{-3}$	rs478324	A	C	0.13(0.02)	10^{-7}	0.07(0.02)	10^{-3}	*NTM*
rs58689390	C	T	0.07(0.02)	$8x10^{-4}$	0.04(0.02)	$2x10^{-2}$	rs10416388	G	A	0.07(0.02)	10^{-3}	0.06(0.02)	10^{-3}	
rs59884341	C	T	0.22(0.06)	$5x10^{-4}$	0.22(0.06)	$2x10^{-4}$	rs59884341	C	T	0.22(0.06)	$5x10^{-4}$	0.22(0.06)	$2x10^{-4}$	
rs6029251	G	A	0.06(0.02)	$8x10^{-4}$	0.07(0.02)	$8x10^{-5}$	rs735031	G	A	0.06(0.02)	$8x10^{-4}$	0.07(0.02)	$8x10^{-5}$	
rs60798171	G	T	0.07(0.02)	$4x10^{-4}$	0.07(0.02)	$6x10^{-5}$	rs60798171	G	T	0.07(0.02)	$4x10^{-4}$	0.07(0.02)	$6x10^{-5}$	
rs62192813	G	T	0.18(0.06)	10^{-3}	0.18(0.05)	$4x10^{-4}$	rs62192813	G	T	0.18(0.06)	10^{-3}	0.18(0.05)	$4x10^{-4}$	
rs638311	A	T	0.07(0.02)	10^{-4}	0.05(0.02)	$7x10^{-4}$	rs10802808	T	C	0.07(0.02)	$2x10^{-4}$	0.06(0.02)	$6x10^{-5}$	*CHRM3*
rs6421157	G	A	0.07(0.02)	$2x10^{-5}$	0.05(0.02)	$3x10^{-4}$	rs10454999	G	A	0.07(0.02)	$8x10^{-5}$	0.06(0.02)	$9x10^{-5}$	
rs6422311	G	A	0.07(0.02)	10^{-4}	0.07(0.02)	$2x10^{-5}$	rs6858688	A	T	0.07(0.02)	$2x10^{-4}$	0.07(0.02)	$2x10^{-5}$	
rs6435661	C	T	0.07(0.02)	$2x10^{-4}$	0.06(0.02)	10^{-3}	rs6435661	C	T	0.07(0.02)	$2x10^{-4}$	0.06(0.02)	10^{-3}	*ERBB4*
rs6467754	G	A	0.08(0.02)	$4x10^{-5}$	0.05(0.02)	10^{-3}	rs4732306	C	T	0.07(0.02)	$8x10^{-5}$	0.06(0.02)	$8x10^{-4}$	
rs6584356	A	C	0.15(0.04)	$2x10^{-4}$	0.13(0.04)	$5x10^{-4}$	rs35662118	T	C	0.15(0.04)	$2x10^{-4}$	0.13(0.04)	$4x10^{-4}$	*PKD2L1*
rs6701243	A	C	0.08(0.02)	$2x10^{-5}$	0.06(0.02)	$5x10^{-4}$	rs35518820	C	T	0.07(0.02)	$2x10^{-4}$	0.06(0.02)	$4x10^{-4}$	
rs6750890	T	G	0.10(0.02)	$3x10^{-5}$	0.07(0.02)	$8x10^{-4}$	rs6750890	T	G	0.10(0.02)	$3x10^{-5}$	0.07(0.02)	$8x10^{-4}$	*LOC100507443*
rs71429234	G	A	0.07(0.02)	10^{-3}	0.06(0.02)	$2x10^{-3}$	rs13006868	G	A	0.07(0.02)	$2x10^{-3}$	0.06(0.02)	10^{-3}	
rs72660658	C	T	0.06(0.02)	10^{-3}	0.06(0.02)	$2x10^{-4}$	rs10488866	T	C	0.06(0.02)	10^{-3}	0.06(0.02)	$2x10^{-4}$	*CXXC4-AS1*
rs72762883	T	G	0.11(0.03)	$3x10^{-4}$	0.08(0.03)	$5x10^{-3}$	rs72762892	A	G	0.10(0.03)	$7x10^{-4}$	0.08(0.03)	$8x10^{-4}$	*KIF26B*
rs73086963	T	C	0.25(0.06)	10^{-4}	0.17(0.06)	$6x10^{-3}$	rs41308228	A	T	0.24(0.07)	$3x10^{-4}$	0.21(0.06)	$7x10^{-4}$	*CACNA2D2*

rs73088112	C	T	0.27(0.07)	10^{-4}	0.23(0.06)	$4x10^{-4}$	rs73074869	A	G	0.27(0.07)	10^{-4}	0.24(0.06)	$2x10^{-4}$	*RHOA, BSN*
rs73116288	G	T	0.07(0.02)	10^{-4}	0.06(0.02)	$7x10^{-4}$	rs62254854	C	T	0.07(0.02)	$3x10^{-4}$	0.06(0.02)	$2x10^{-4}$	
rs7319835	T	C	0.09(0.03)	10^{-3}	0.09(0.03)	$4x10^{-4}$	rs61965320	A	G	0.09(0.03)	$2x10^{-3}$	0.09(0.03)	$4x10^{-4}$	
rs73214716	T	C	0.11(0.04)	10^{-3}	0.11(0.03)	$4x10^{-4}$	rs73214716	T	C	0.11(0.04)	10^{-3}	0.11(0.03)	$4x10^{-4}$	
rs74992427	C	T	0.20(0.06)	10^{-3}	0.17(0.06)	$2x10^{-3}$	rs77709635	T	C	0.16(0.06)	$5x10^{-3}$	0.20(0.05)	$9x10^{-5}$	*ZNF704*
rs75263467	A	G	0.16(0.04)	$3x10^{-4}$	0.21(0.04)	$5x10^{-7}$	rs75263467	A	G	0.16(0.04)	$3x10^{-4}$	0.21(0.04)	$5x10^{-7}$	
rs75874195	A	G	0.15(0.04)	$9x10^{-4}$	0.10(0.04)	10^{-2}	rs78924818	C	G	0.17(0.06)	$3x10^{-3}$	0.18(0.05)	$3x10^{-4}$	
rs7625233	G	A	0.06(0.02)	$4x10^{-4}$	0.06(0.02)	10^{-4}	rs7629352	G	A	0.05(0.02)	$5x10^{-3}$	0.07(0.02)	$4x10^{-5}$	
rs76530346	G	A	0.10(0.03)	10^{-3}	0.09(0.03)	10^{-3}	rs78826784	G	A	0.09(0.03)	$2x10^{-3}$	0.09(0.03)	$4x10^{-4}$	*AUH, LINC00484*
rs7781266	C	A	0.07(0.02)	10^{-3}	-0.05(0.02)	10^{-2}	rs17167170	G	A	-0.06(0.02)	$8x10^{-3}$	0.06(0.02)	10^{-3}	*EXOC4, LOC101928861*
rs77966298	A	G	0.09(0.02)	$2x10^{-4}$	0.09(0.02)	$5x10^{-5}$	rs6737620	C	T	0.08(0.02)	$6x10^{-4}$	0.09(0.02)	10^{-5}	
rs78491961	T	C	0.21(0.06)	$3x10^{-4}$	0.22(0.06)	$3x10^{-4}$	rs78491961	T	C	0.21(0.06)	$3x10^{-4}$	0.22(0.06)	$3x10^{-4}$	
rs78737628	C	A	0.09(0.02)	$4x10^{-4}$	0.08(0.02)	$5x10^{-4}$	rs78737628	C	A	0.09(0.02)	$4x10^{-4}$	0.08(0.02)	$5x10^{-4}$	
rs80013344	A	G	0.10(0.03)	$3x10^{-4}$	0.09(0.02)	$3x10^{-4}$	rs4947694	C	A	0.10(0.03)	$3x10^{-4}$	0.11(0.03)	$2x10^{-5}$	
rs80088989	C	A	0.15(0.05)	10^{-3}	0.17(0.05)	$2x10^{-4}$	rs80088989	C	A	0.15(0.05)	10^{-3}	0.17(0.05)	$2x10^{-4}$	*TEK*
rs80229434	A	G	0.22(0.07)	10^{-3}	0.19(0.06)	$3x10^{-3}$	rs111620031	C	G	0.14(0.07)	$3x10^{-2}$	0.20(0.06)	10^{-3}	*MIR3680-1, MIR3680-2*
rs927603	T	C	0.06(0.02)	$3x10^{-4}$	0.04(0.02)	$6x10^{-3}$	rs8016878	A	G	0.05(0.02)	$5x10^{-3}$	0.06(0.02)	10^{-4}	
rs978216	C	T	0.06(0.02)	$8x10^{-4}$	0.06(0.02)	$2x10^{-5}$	rs978216	C	T	0.06(0.02)	$8x10^{-4}$	0.06(0.02)	$2x10^{-5}$	
rs9816530	T	G	0.08(0.02)	$6x10^{-4}$	0.06(0.02)	$4x10^{-3}$	rs6441815	T	C	0.07(0.02)	10^{-3}	0.06(0.02)	10^{-3}	
rs9855048	G	A	0.12(0.03)	10^{-4}	0.10(0.03)	$4x10^{-4}$	rs9855048	G	A	0.12(0.03)	10^{-4}	0.10(0.03)	$4x10^{-4}$	
rs9878955	G	A	0.06(0.02)	10^{-3}	0.04(0.02)	10^{-2}	rs7616968	G	A	0.05(0.02)	10^{-2}	0.06(0.02)	$5x10^{-4}$	*OSTN*

ASD(iPSYCH, woADHD) and ADHD(iPSYCH) SNP estimates for ASD variants ($\hat{\beta}_{ASD}$, P_{thr} <0.0015) that were also represented within the ADHD variant set (P_{thr} <0.0015, 500kb and LD-r^2≥0.6). Likewise, SNP estimates for ADHD variants ($\hat{\beta}_{ADHD}$, P_{thr} <0.0015) that were also represented within the ASD variant set (P_{thr} <0.0015, 500kb and LD-r^2≥0.6) are shown. ASD and ADHD variants that tag each other are represented on the same row. All SNP estimates are aligned according to A1. Mapped genes were identified using a 0kb window range. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; P_{thr}, *P*-value threshold; woADHD, without ADHD

Supplementary Table 16: Permutation analysis

MVR	Variant selection			Empirical *P*-value (SE)	
	P_{thr} ASD (iPSYCH, woADHD)	P_{thr} ADHD (iPSYCH)	N_{SNPs}	Specific effects	Cross-disorder effects
ASD-MVR	0.0015	0.0015	83	$P(\hat{\theta}_{ASD})$ $<10^{-4}(<10^{-4})$	$P(\hat{\theta}_{\text{✠}ADHD})$ $7x10^{-4}(3x10^{-4})$
ADHD-MVR	0.0015	0.0015	83	$P(\hat{\theta}_{ADHD})$ $9x10^{-4}(3x10^{-4})$	$P(\hat{\theta}_{\text{✠}ASD})$ $2x10^{-4}(10^{-4})$

83 SNPs were randomly selected from either ASD variants (P_{thr}<0.0015) or ADHD variants (P_{thr}<0.0015). Corresponding SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were subsequently extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics, respectively. 10,000 MVRs were performed and the number of times a permuted MVR effect was at least as significant as an observed MVR effect counted. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; N_{SNPs}, number of SNPs; P_{thr}, *P*-value threshold; woADHD, without ADHD

Supplementary Table 17: Follow-up ASD-MVR (P_{thr}<0.0015; P_{thr}<0.05), with MDD, SCZ and BD SNP estimates (independent variable)

Variant selection			Cross-disorder	Intercept		ASD-specific effect		Cross-disorder effect		VIF	MVR to UVR model fit comparison	
Gi	N_{SNPs}	P_{thr}		θ_0*(SE)	P	θ_{ASD}(SE)	P	$\theta_{✠}$*(SE)	P		ΔR^2 (%)	LRT-χ^2(df), P
ASD (iPSYCH, woADHD)	2,644	0.0015	MDD (PGC)	$0.001(10^{-4})$	$3x10^{-5}$	$-0.001(0.002)$	0.45	-		-		-
				$0.001(2x10^{-4})$	$3x10^{-5}$	$-0.001(0.002)$	0.55	$\theta_{✠*MDD}$=-0.004(0.006)	0.46	1.04	0.02	1.37, 0.45
	50,904	0.05		$3x10^{-4}$ $(4x10^{-5})$	$<10^{-10}$	$7x10^{-4}(5x10^{-4})$	0.13	-		-		-
				$3x10^{-4}(4x10^{-5})$	$<10^{-10}$	$0.001(5x10^{-4})$	0.002	$\theta_{✠*MDD}$=-0.012(0.001)	$<10^{-10}$	1.02	0.29	264.96, $<10^{-10}$
	1,786	0.0015	SCZ (PGC)	$0.001(3x10^{-4})$	0.01	$-0.001(0.003)$	0.78	-		-		-
				$0.001(3x10^{-4})$	0.02	$-0.001(0.003)$	0.83	$\theta_{✠*SCZ}$=0.016(0.004)	$3x10^{-4}$	1.00	0.71	37.92, $3x10^{-4}$
	31,026	0.05		$2x10^{-4}(5x10^{-5}$	$6x10^{-5}$	$0.002(8x10^{-4})$	0.01	-		-		-
				$2x10^{-4}(5x10^{-5})$	10^{-4}	0.002(0.001)	0.02	$\theta_{✠*SCZ}$ =0.006(0.001)	$3x10^{-9}$	1.00	0.11	73.31, $3x10^{-9}$
	1,859	0.0015	BD (PGC)	$0.001(2x10^{-4})$	0.003	$-0.001(0.003)$	0.61	-		-		-
				$0.001(2x10^{-4})$	0.004	$-0.001(0.003)$	0.61	$\theta_{✠*BD}$=0.019(0.004)	$6x10^{-7}$	1.00	1.33	71.40, $6x10^{-7}$
	32,367	0.05		$2x10^{-4}(5x10^{-5})$	$2x10^{-6}$	$0.002(7x10^{-4})$	0.04	-		-		-
				$2x10^{-4}(5x10^{-5})$	$7x10^{-6}$	0.002(0.001)	0.033	$\theta_{✠*BD}$ =0.009(0.001)	$<10^{-10}$	1.00	0.41	272.24, $<10^{-10}$

Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and MDD ($\hat{\theta}_{✠*MDD}$), SCZ ($\hat{\theta}_{✠*SCZ}$) or BD ($\hat{\theta}_{✠*BD}$) cross-disorder effects on EA as fitted with ASD-MVR (Supplementary Figure 1k, Supplementary Figure 4a). Sets of independent ASD (Gi) genetic variants were selected from ASD(iPSYCH, woADHD) and are shown for two P-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). SNP estimates $\hat{\beta}_{ASD}$, $\hat{\beta}_{MDD}$, $\hat{\beta}_{SCZ}$, $\hat{\beta}_{BD}$ and $\hat{\beta}_{EA}$ were extracted from ASD(iPSYCH, woADHD), MDD(PGC), SCZ(PGC), BD(PGC) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD or MDD, SCZ or BD liability. The model fit of MVRs and UVRs was compared with likelihood-ratio tests (LRTs). The multiple testing threshold is P<0.0042. Abbreviations: ASD, Autism Spectrum Disorder; BD, Bipolar Disorder; df, degrees of freedom; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; LRT, likelihood-ratio test; MDD; Major Depressive Disorder; MVR, multivariable regression; PGC, Psychiatric Genomics Consortium; SCZ, Schizophrenia; P_{thr}, P-value threshold; woADHD; without ADHD; UVR, univariable regression; ΔR^2, difference in Regression R^2 between MVR and UVR

Supplementary Table 18: Follow-up ADHD-MVR (P_{thr}<0.0015; P_{thr}<0.05), with MDD, SCZ and BD SNP estimates (independent variable)

Variant selection			Cross-disorder	Intercept		ADHD-specific effect		Cross-disorder effect		VIF	MVR to UVR model fit comparison	
Gj	N_{SNPs}	P_{thr}		θ_0(SE)	P	θ_{ADHD}(SE)	P	$\theta_{\#}$(SE)	P		ΔR^2 (%)	LRT-χ^2(df), P
ADHD (iPSYCH)	2,567	0.0015	MDD (PGC)	-0.001(2x10^{-4})	<10^{-10}	-0.005(0.003)	0.11	-		-		-
				-0.001(2x10^{-4})	4x10^{-10}	-0.002(0.003)	0.56	$\theta_{\#MDD}$=0.033(0.007)	5x10^{-7}	1.04	0.98	76.87, 4x10^{-7}
	38,093	0.05		-0.001(4x10^{-5})	<10^{-10}	-0.005(0.001)	10^{-10}	-		-		-
				-0.001(4x10^{-5})	<10^{-10}	-0.004(0.001)	5x10^{-7}	$\theta_{\#MDD}$=0.013(0.001)	<10^{-10}	1.02	0.23	181.10, <10^{-10}
	2,278	0.0015	SCZ (PGC)	-0.001(3x10^{-4})	7x10^{-7}	-0.007(0.004)	0.056	-		-		-
				-0.001(3x10^{-4})	9x10^{-7}	-0.007(0.004)	0.059	$\theta_{\#SCZ}$=-0.006(0.004)	0.14	1.00	0.09	6.91, 0.14
	31,395	0.05		-0.001(5x10^{-5})	<10^{-10}	-0.007(0.001)	<10^{-10}	-		-		-
				-0.001(5x10^{-5})	<10^{-10}	-0.007(0.001)	<10^{-10}	$\theta_{\#SCZ}$=0.003(0.001)	0.007	1.00	0.02	16.21, 0.007
	2,340	0.0015	BD (PGC)	-0.001(2x10^{-4})	10^{-7}	-0.007(0.003)	0.040	-		-		-
				-0.001(2x10^{-4})	8x10^{-8}	-0.007(0.003)	0.042	$\theta_{\#BD}$=0.007(0.004)	0.046	1.00	0.17	12.56, 0.046
	32,788	0.05		-0.001(5x10^{-5})	<10^{-10}	-0.007(0.001)	<10^{-10}	-		-		-
				-0.001(5x10^{-5})	<10^{-10}	-0.007(0.001)	<10^{-10}	$\theta_{\#BD}$=0.008(0.001)	<10^{-10}	1.00	0.29	205.72, <10^{-10}

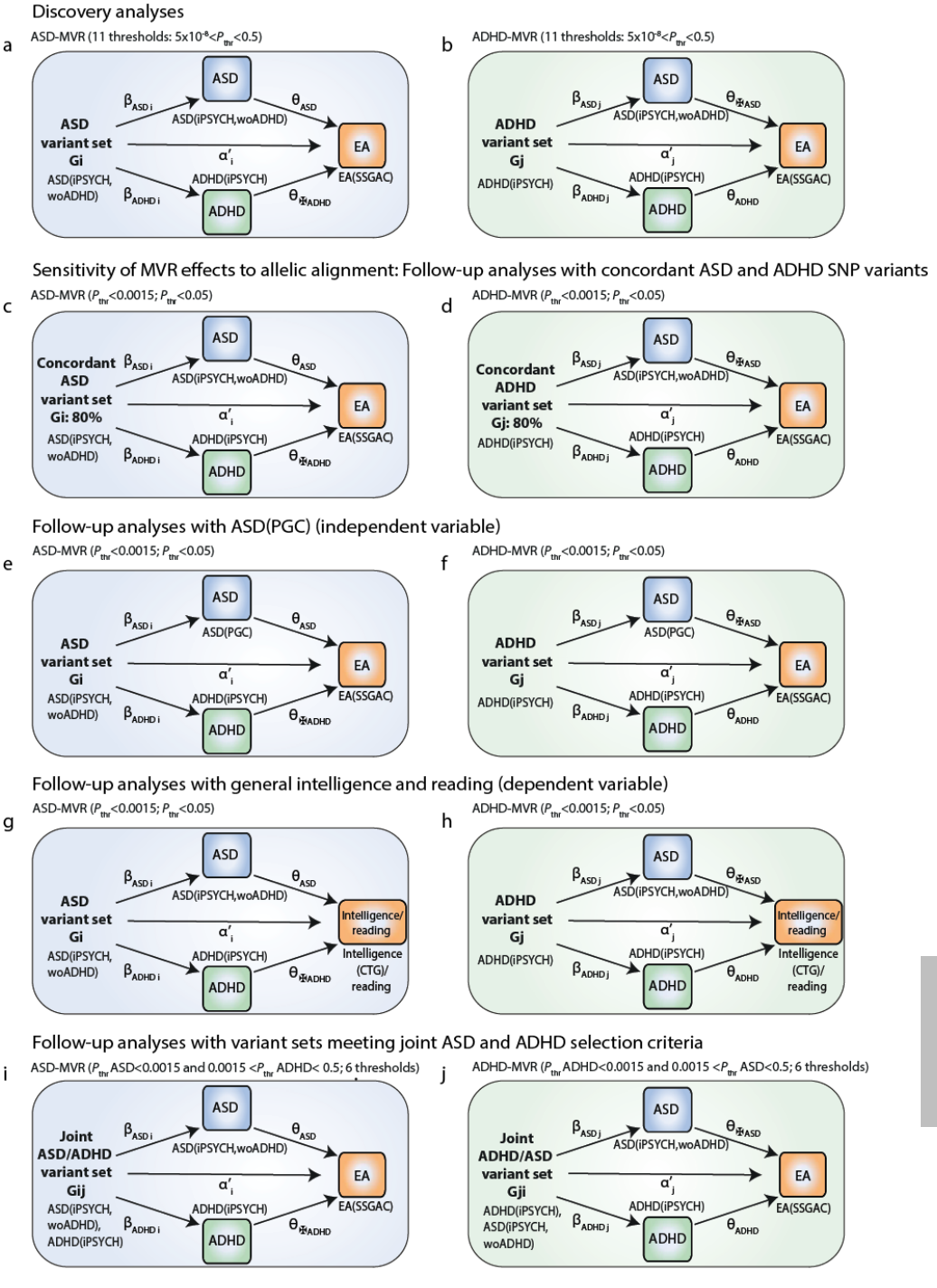
Estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and MDD ($\hat{\theta}_{\#MDD}$), SCZ ($\hat{\theta}_{\#SCZ}$) or BD ($\hat{\theta}_{\#BD}$) on EA as fitted with ADHD-MVR (Supplementary Figure 1l, Supplementary Figure 4b). Sets of independent ADHD (Gj) genetic variants were selected from ADHD(iPSYCH) and are shown for two *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). SNP estimates $\hat{\beta}_{ADHD}$, $\hat{\beta}_{MDD}$, $\hat{\beta}_{SCZ}$, $\hat{\beta}_{BD}$ and $\hat{\beta}_{EA}$ were extracted from ADHD(iPSYCH), MDD(PGC), SCZ(PGC), BD(PGC) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ADHD or MDD, SCZ or BD liability. The model fit of MVRs and UVRs was compared with likelihood-ratio tests (LRTs). The multiple testing threshold is *P*<0.0042. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; BD, Bipolar Disorder; df, degrees of freedom; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; LRT, likelihood-ratio test; MDD; Major Depressive Disorder; MVR, multivariable regression; PGC, Psychiatric Genomics Consortium; SCZ, Schizophrenia; P_{thr}, *P*-value threshold; woADHD; without ADHD; UVR, univariable regression; ΔR^2, difference in Regression R^2 between MVR and UVR

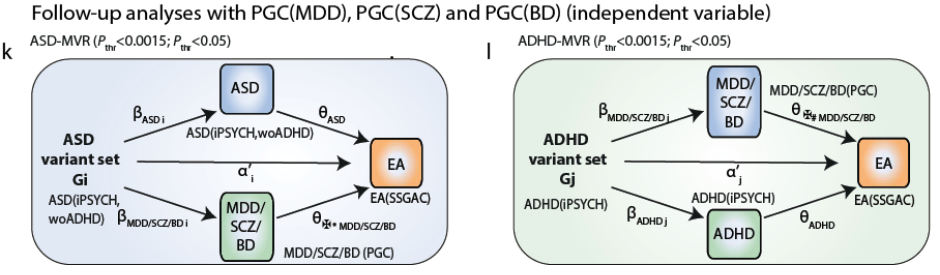
Supplementary Table 19: Simulation of a hypothetical multifactorial model of EA, ASD and ADHD interrelationships

Parameter		Simulated parameter	Estimated parameter (SE)
SNP-h^2	EA	0.25	0.34 (0.06)
	ADHD liability	0.50	0.42 (0.06)
	ASD liability	0.50	0.40 (0.06)
e^2	EA	0.75	0.66 (0.06)
	ADHD liability	0.50	0.58 (0.06)
	ASD liability	0.50	0.60 (0.06)
Genetic factor loadings	a_{11}	0.50	-0.58 (0.05)
	a_{21}	-0.35	0.30 (0.07)
	a_{31}	0.20	-0.14 (0.07)
	a_{22}	0.62	-0.58 (0.05)
	a_{32}	0.14	-0.11 (0.08)
	a_{33}	0.66	-0.61 (0.05)
Residual factor loadings	e_{11}	0.87	-0.81 (0.04)
	e_{21}	0.00	0.01 (0.05)
	e_{31}	0.00	-0.03 (0.05)
	e_{22}	0.71	0.76 (0.04)
	e_{32}	0.00	0.00 (0.05)
	e_{33}	0.71	-0.77 (0.04)
r_g	EA, ADHD	-0.49	-0.46 (0.11)
	EA, ASD	0.28	0.23 (0.11)
	ADHD, ASD	0.04	0.04 (0.10)

We simulated three continuous interrelated measures corresponding to EA, ADHD liability and ASD liability (N=6000 each), informed by unconstrained LDSC genetic correlations (r_g) using GWAS statistics for EA, ADHD(iPSYCH) and ASD(PGC). Residual correlations were assumed to be absent, given the independence of the respective GWAS statistics. Note that simulated SNP-h^2 estimates were increased, compared to the observed values, to reduce the computational burden. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; EA, educational attainment; SNP-h^2, Single Nucleotide Polymorphism heritability; e^2, residual variance; r_g, genetic correlation

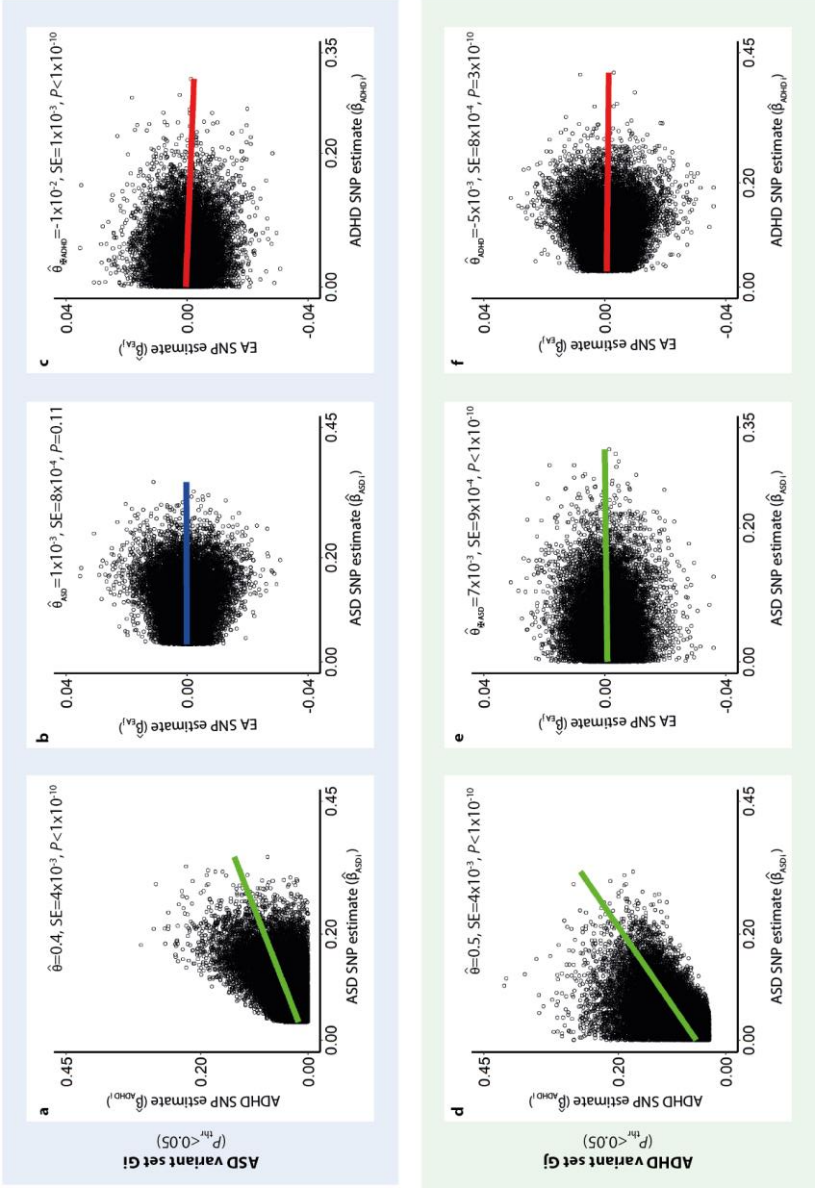
Supplementary Figures





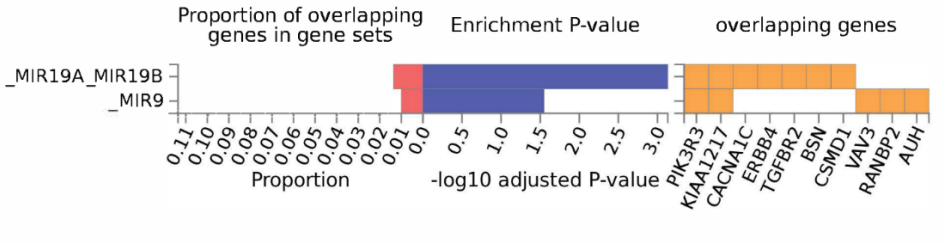
Supplementary Figure 1: MVR study design

Acyclic graphs illustrating the multivariable regression (MVR) design and selected GWAS statistics. ASD-MVR **(a,c,e,g,i,k)** involve a set of ASD variants (Gi), two independent variables (ASD risk and risk for another disorder such as ADHD, MDD, SCZ or BD) and one dependent variable (EA, general intelligence or reading). The genetic association effect of Gi on ASD is $\beta_{ASDi.}$ The genetic association effect of Gi on ADHD, MDD, SCZ or BD is β_{ADHDi}, β_{MDDi}, β_{SCZi} or β_{BDi} respectively. The genetic association effect of ASD risk on the dependent variable is the ASD-specific effect $\theta_{ASD.}$ For ASD-MVR studying ADHD risk **(a,c,e,g,i)**, the genetic association effect of ADHD risk on the dependent variable is the cross-disorder effect $\theta_{\#ADHD}$. For ASD-MVR studying MDD, SCZ or BD risk **(k)**, the genetic association effects of MDD, SCZ or BD on the dependent variable are the cross-disorder effects $\theta_{\#*MDD,}$ $\theta_{\#*SCZ}$ or $\theta_{\#*BD}$ respectively. The intercept α'_{ASD} represents the direct effect of ASD variants Gi on the dependent variable that are neither captured by θ_{ASD} nor any $\theta_{\#}$. ADHD-MVR **(b,d,f,h,j,l)** involve a set of ADHD variants (Gj), two independent variables (ADHD risk and risk for another disorder such as ASD, MDD, SCZ or BD) and one dependent variable (EA, general intelligence or reading). The genetic association effect of Gj on ADHD is $\beta_{ADHDj.}$ The genetic association effect of Gj on ASD, MDD, SCZ or BD is β_{ASDj}, β_{MDDj}, β_{SCZj} or β_{BDj} respectively. The genetic association effect of ADHD risk on the dependent variable is the ADHD-specific effect $\theta_{ADHD.}$ For ADHD-MVR studying ASD risk **(b,d,f,h,j)**, the genetic association effect of ASD risk on the dependent variable is the cross-disorder effect $\theta_{\#ASD}$. For ADHD-MVR studying MDD, SCZ or BD risk **(l)**, the genetic association effect of MDD, SCZ or BD on the dependent variable are the cross-disorder effects $\theta_{\#MDD,}$ $\theta_{\#SCZ}$ or $\theta_{\#BD}$ respectively. The intercept α'_{ADHD} represents the direct effect of ADHD variants Gj on the dependent variable that are neither captured by θ_{ADHD} nor any $\theta_{\#}$. For MVR models **(a,b,c,d,i,j)**, SNP estimates $\hat{\beta}_{ASD,}$ $\hat{\beta}_{ADHD}$ and $\hat{\beta}_{EA}$ were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. For MVR models (**e,f**), SNP estimates $\hat{\beta}_{ASD,}$ $\hat{\beta}_{ADHD}$ and $\hat{\beta}_{EA}$ were extracted from ASD(PGC), ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. For MVR models **(g,h)**, SNP estimates $\hat{\beta}_{ASD,}$ $\hat{\beta}_{ADHD.}$ $\hat{\beta}_{Intelligence}$ and $\hat{\beta}_{reading}$ were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH), Intelligence(CTG) and reading GWAS statistics respectively. For MVR models **(k,l)**, SNP estimates $\hat{\beta}_{ASD,}$ $\hat{\beta}_{ADHD,}$ $\hat{\beta}_{MDD,}$ $\hat{\beta}_{SCZ,}$ $\hat{\beta}_{BD}$ and $\hat{\beta}_{EA}$ were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH), MDD(PGC), SCZ(PGC), BD(PGC) and EA(SSGAC) GWAS statistics respectively. Independent ASD(Gi), ADHD(Gj) and joint ASD/ADHD (G_{ij} and G_{ji}) genetic variant sets were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively. Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; BD, Bipolar Disorder; EA, Educational attainment, Intelligence, General intelligence, GWAS, genome-wide association study, iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; MDD; Major Depressive Disorder; MVR, multivariable regression; P_{thr}, *P*-value threshold; PGC, Psychiatric Genomics Consortium; SCZ, Schizophrenia; SSGAC, Social Science Genetic Consortium; woADHD; without ADHD


a
b
c
d
e
f
ASD variant set Gi
(P_{thr}<0.05)
ADHD variant set Gj
(P_{thr}<0.05)
ADHD SNP estimate ($\hat{\beta}_{ADHD,i}$)
ASD SNP estimate ($\hat{\beta}_{ASD,i}$)
EA SNP estimate ($\hat{\beta}_{EA,i}$)
ADHD SNP estimate ($\hat{\beta}_{ADHD,i}$)
$\hat{\theta}$=0.4, SE=4x10^{-3}, P<1x10^{-10}
$\hat{\theta}_{ASD}$=1x10^{-3}, SE=8x10^{-4}, P=0.11
$\hat{\theta}_{ADHD}$=-1x10^{-2}, SE=1x10^{-3}, P<1x10^{-10}
$\hat{\theta}$=0.5, SE=4x10^{-3}, P<1x10^{-10}
$\hat{\theta}_{ASD}$=7x10^{-3}, SE=9x10^{-4}, P<1x10^{-10}
$\hat{\theta}_{ADHD}$=-5x10^{-3}, SE=8x10^{-4}, P=3x10^{-10}
ASD SNP estimate ($\hat{\beta}_{ASD,i}$)
ADHD SNP estimate ($\hat{\beta}_{ADHD,i}$)
0.45
0.20
0.00
-0.04
0.04
0.35

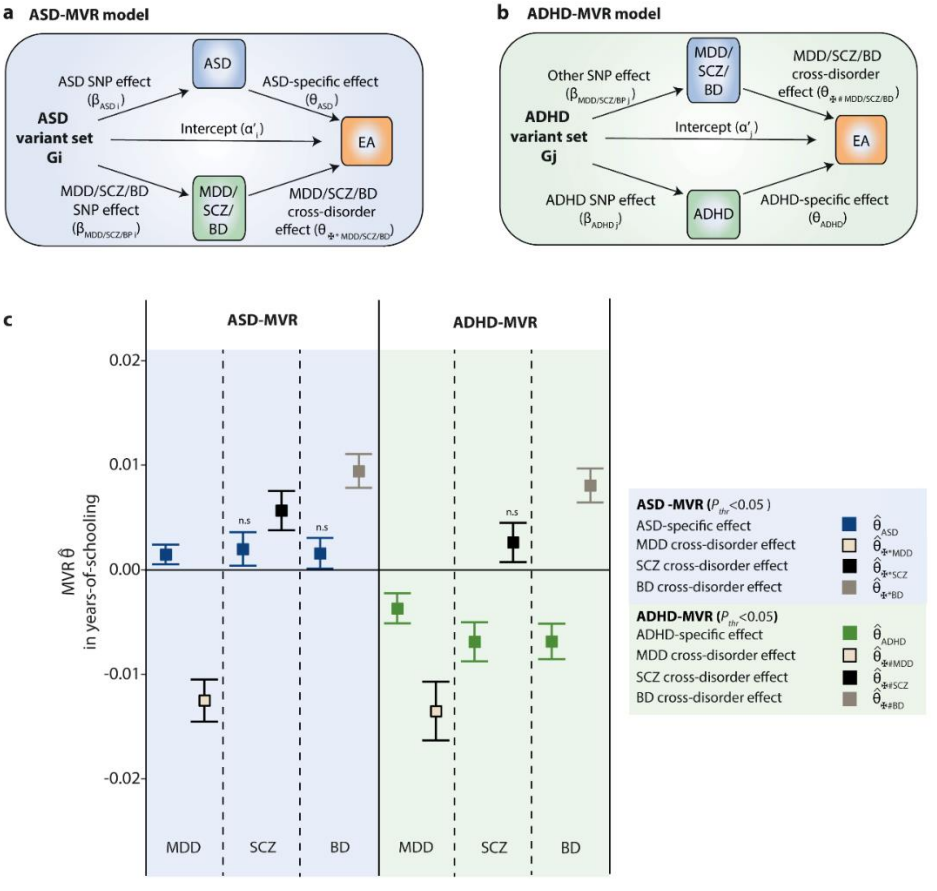
Supplementary Figure 2: Bivariate analyses of ASD, ADHD and EA genetic effects using ASD and ADHD concordant variant sets

Sets of independent ASD (Gi) and ADHD (Gj) genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively, based on a *P*-value threshold of 0.05. SNP estimates for ASD ($\hat{\beta}_{ASD}$), ADHD ($\hat{\beta}_{ADHD}$) and EA ($\hat{\beta}_{EA}$) were extracted from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) and EA(SSGAC) GWAS statistics respectively. Only genetic variants with the same risk-increasing allele for both ASD and ADHD (concordant variants) were included. **(a-c)** Using univariable weighted regression models and ASD variants (Gi), **(a)** ADHD SNP estimates ($\hat{\beta}_{ADHD}$) were regressed on ASD SNP estimates ($\hat{\beta}_{ASD}$), **(b)** EA SNP estimates ($\hat{\beta}_{EA}$) were regressed on ASD SNP estimates ($\hat{\beta}_{ASD}$), and **(c)** EA SNP estimates ($\hat{\beta}_{EA}$) were regressed on ADHD SNP estimates ($\hat{\beta}_{ADHD}$). **(d-f)** Using univariable weighted regression models and ADHD variants (Gj), **(d)** ADHD SNP estimates ($\hat{\beta}_{ADHD}$) were regressed on ASD SNP estimates ($\hat{\beta}_{ASD}$), **(e)** EA SNP estimates ($\hat{\beta}_{EA}$) were regressed on ASD SNP estimates ($\hat{\beta}_{ASD}$), and **(f)** EA SNP estimates ($\hat{\beta}_{EA}$) were regressed on ADHD SNP estimates ($\hat{\beta}_{ADHD}$). All regressions allowed for an intercept. A green regression line denotes a positive relationship between SNP estimates, a blue regression line reflects no significant relationship and a red line indicates a negative relationship. Regression estimates ($\hat{\theta}$), corresponding standard errors (SE) and *P*-values are shown for each regression model. Abbreviations: ASD, Autism Spectrum Disorder; ADHD, Attention-Deficit/Hyperactivity Disorder; EA, educational attainment; P_{thr}, *P*-value threshold



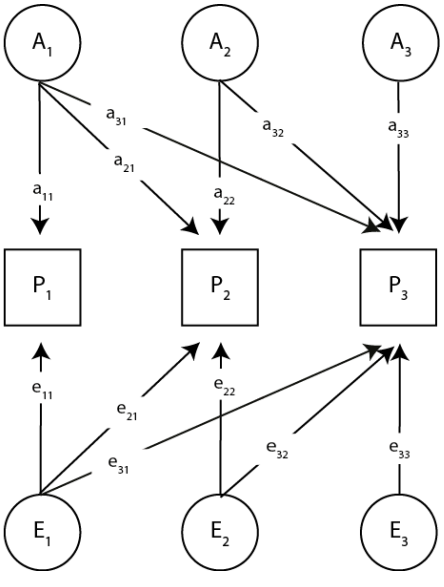
Supplementary Figure 3: Enrichment for microRNA targets

Variants selected at joint *P*-value threshold for both ASD and ADHD (P_{thr}<0.0015) included 83 SNPs that were mapped to 52 loci. Of those, 45 were aligned to unique Ensembl IDs (v92) and subjected to gene-set enrichment analysis (requesting at least 5 overlapping genes) within BrainSpan samples (29 different ages of brain samples and 11 developmental stages of brain samples) using FUMA software. The strongest evidence for enrichment was found for micro RNA targets, TTTGCAC_MIR19A_MIR19B and ACCAAAG_MIR9, as shown with 515 and 500 background genes respectively (Enrichment TTTGCAC_MIR19A_MIR19B *P*-unadjusted=3.3x10^{-6}, adjusted- *P*=0.00074; Enrichment ACCAAAG_MIR9 *P*-unadjusted = 0.00039, *P*-adjusted = 0.028).



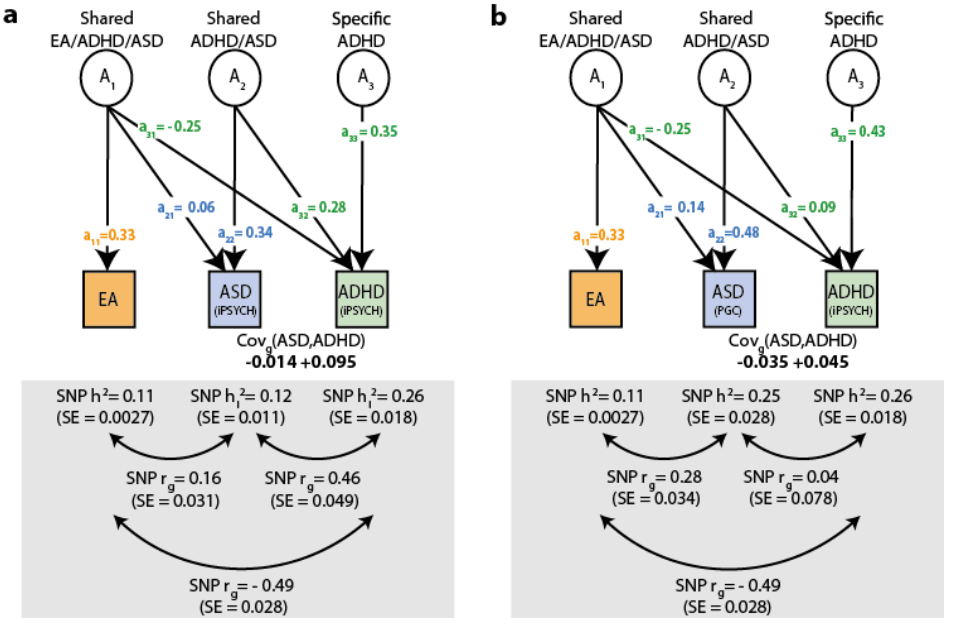
Supplementary Figure 4: Cross-disorder associations with educational attainment for other disorders. (a) Acyclic graph illustrating multivariable regression (MVR) for a set of ASD variants Gi (ASD-MVR), two independent variables (ASD risk and risk for either MDD, SCZ or BD) and the dependent variable EA. The genetic association effect of Gi on ASD, MDD, SCZ and BD risk are β_{ASDi}, β_{MDDi}, β_{SCZi} and β_{BDi} respectively. The genetic association effect of ASD risk on EA is the ASD-specific effect θ_{ASD}. The genetic association effects of MDD, SCZ or BD risk on EA are the cross-disorder effects $\theta_{✠*MDD}$, $\theta_{✠*SCZ}$ and $\theta_{✠*BD}$ respectively. The intercept α'_{ASD} represents the direct effect of ASD variants G_i on EA that are neither captured by θ_{ASD} nor any $\theta_{✠*}$. **(b)** Analogous acyclic graph illustrating MVR for a set of ADHD variants Gj (ADHD-MVR), two independent variables (ASD and ADHD risk) and the dependent variable EA. The genetic association effect of Gj on ADHD, MDD, SCZ and BD risk are β_{ADHDj}, β_{MDDj}, β_{SCZj} and β_{BDj} respectively. The genetic association effect of ADHD risk on EA is the ADHD-specific effect θ_{ADHD}. The genetic association effects of MDD, SCZ or BD risk on EA are the cross-disorder effects $\theta_{✠\#MDD}$, $\theta_{✠\#SCZ}$ and $\theta_{✠\#BD}$ respectively. The intercept α'_j represents the direct effect of ADHD variants G_j on EA that are neither captured by θ_{ADHD} nor any $\theta_{✠\#}$. **(c)** Estimated ASD-specific effect $\hat{\theta}_{ASD}$ and MDD, SCZ or BD cross-disorder effects $\hat{\theta}_{✠*MDD}$, $\hat{\theta}_{✠*SCZ}$ and $\hat{\theta}_{✠*BD}$ as fitted with ASD-MVR (a) and estimated ADHD-specific effect $\hat{\theta}_{ADHD}$ and MDD, SCZ or BD cross-disorder effects $\hat{\theta}_{✠\#MDD}$, $\hat{\theta}_{✠\#SCZ}$ and $\hat{\theta}_{✠\#BD}$ as fitted with ADHD-MVR (b). Sets of independent ASD (G_i) and ADHD (G_j) genetic variants were selected from ASD(iPSYCH, woADHD) and ADHD(iPSYCH) GWAS statistics respectively and are shown for two *P*-value thresholds (P_{thr}<0.0015, P_{thr}<0.05). SNP estimates $\hat{\beta}_{ASD}$, $\hat{\beta}_{ADHD}$, $\hat{\beta}_{MDD}$, $\hat{\beta}_{SCZ}$, $\hat{\beta}_{BD}$ and $\hat{\beta}_{EA}$ were extracted from ASD(iPSYCH, woADHD), ADHD(iPSYCH), MDD(PGC), SCZ(PGC), BD(PGC) and EA(SSGAC) GWAS statistics respectively. All MVR effects are presented as change in years-of-schooling per increase in log-odds of ASD, ADHD, MDD, SCZ or BD liability. Bars represent 95% confidence intervals. n.s. - MVR effects $\hat{\theta}$ that did not pass the multiple testing threshold of *P*< 0.0042. Abbreviations: ADHD, Attention-

Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; BD, Bipolar Disorder; EA, Educational attainment, iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; MDD; Major Depressive Disorder; MVR, multivariable regression; P_{thr}, *P*-value threshold; PGC, Psychiatric Genomics Consortium; SCZ, Schizophrenia; SSGAC, Social Science Genetic Consortium; woADHD; without ADHD



Supplementary Figure 5: Path diagram for a trivariate trait

The variance/covariance structure of multivariate trait consisting of three standardised measures P1, P2 and P3 can be described using a Cholesky decomposition consisting of three genetic factors (A1, A2 and A3) and three residual factors (E1, E2 and E3), shown here with genetic and residual factor loadings. The observed phenotypic measures are represented by squares, while all latent genetic and residual factors are represented by a circle. Single headed arrows ('paths') denote causal relationships between variables and are shown for genetic factor loadings (a) and residual factor loadings (e). Note that the variance of latent variables is constrained to unit variance, this is omitted from the diagrams to improve clarity.



Supplementary Figure 6: Multi-factor model of genetic interrelations between ASD, ADHD and educational attainment (allowing for ADHD-specific genetic influences)

The model predicts two sources of shared genetic influences between ASD and ADHD, as captured by common variants within an infinitely large population. The first genetic factor (A1, shared EA/ADHD/ASD) refers to shared genetic variation between EA, ADHD and ASD. It allows for a negative genetic covariance between ASD and ADHD. The second genetic factor (A2, shared ADHD/ASD) acts independently of A1, explaining positive genetic covariance between ASD and ADHD. Additional ADHD-specific effects are captured by a third factor (A3). Each factor loading ("a") for the Cholesky decomposition of a trivariate trait is described in the Methods. **(a)** Multi-factor model consistent with ASD(iPSYCH), ADHD(iPSYCH) and EA(SSGAC) summary statistics. **(b)** Multi-factor model consistent with ASD(PGC), ADHD(iPSYCH) and EA(SSGAC) summary statistics. Factor loadings ("a") were derived from LDSC SNP-heritability and genetic correlations, according to theory. Phenotypic measures are represented by squares, while latent genetic factors are represented by circles. Single headed arrows denote genetic factor loadings ("a"), double-headed arrows genetic correlations ("r_g"). Residual influences and unit variances for latent variables were omitted. Abbreviations: EA, educational attainment; ADHD, Attention-Deficit/Hyperactivity Disorder; ASD, Autism Spectrum Disorder; iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research; PGC, Psychiatric Genomics Consortium; SNP h^2, SNP heritability; SNP r_g, SNP genetic correlation, cov_g, genetic covariance

Supplementary References

1. Wray NR, Ripke S, Mattheisen M, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*. 2018;50(5):668. doi:10.1038/s41588-018-0090-3
2. Ruderfer DM, Ripke S, McQuillin A, et al. Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell*. 2018;173(7):1705-1715.e16. doi:10.1016/j.cell.2018.05.046
3. Savage JE, Jansen PR, Stringer S, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*. 2018;50(7):912-919. doi:10.1038/s41588-018-0152-6
4. Bulik-Sullivan BK, Loh P-R, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291-295. doi:10.1038/ng.3211
5. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47:1236–1241. doi:10.1038/ng.3406

Chapter 8

General discussion

Language is a complex human suite of skills that the majority of children acquire rapidly and effortlessly. However, there exist large individual differences in the pace and trajectory of language acquisition. This thesis aimed at broadening the knowledge of how genetic influences contribute to such differences during the first three years of life and investigates shared genetic links with subsequent language, literacy and cognitive development, as well as genetic associations of both language and related traits with childhood-onset neurodevelopmental disorders.

8.1. Summary

In chapter 1, I introduced the processes that underpin language development as well as genetic concepts that are relevant to this thesis. The literature review in chapter 2 provided an overview of prior knowledge about the genetic architecture underlying early language development and its links with later language- and literacy-related skills. Studies based on collections of both community-based twin samples and unrelated individuals reported low to modest heritability for individual differences in language skills during the first three years of life, suggesting an underlying genetic component including common variation. At the single-nucleotide polymorphism (SNP) level, a previous meta-genome-wide association study identified a genome-wide signal with expressive vocabulary at 15-18 months of age near *ROBO2*, despite modest statistical power for variant discovery (~38%). At the polygenic level, latent factor twin analyses reported some stability in genetic factors related to language and literacy skills assessed from toddlerhood till early adolescence, but also age-specific genetic characteristics. Beyond these latent factor twin analyses, knowledge about the composition of shared and specific genetic factors underlying language and literacy development has been scarce, highlighting the need for further research on this key issue. In particular, the genetic architecture of early receptive vocabulary and its relationship with subsequent language and literacy development has been little characterised.

Thus, in chapter 3, I assessed whether genetic factors underlying vocabulary skills during toddlerhood are related to a range of mid-childhood/early-adolescent language and literacy abilities, using the longitudinal dataset of the Avon Longitudinal Study of Parents and Children (ALSPAC). Applying a structural equation modelling approach based on unrelated individuals, Genetic-relationship-matrix Structural Equation Modelling (GSEM), I observed evidence for genetic stability during language and literacy development. Specifically, the majority of SNP-heritability (SNP-h^2) for many later complex abilities could be explained by genetic influences that are related to early-childhood vocabulary skills, especially in the receptive domain, suggesting a role for amplification processes. These findings extend the understanding of shared genetic links

between early-life vocabulary and subsequent language and literacy, and highlight the importance of studying receptive language skills early in life.

In chapter 4, I disentangled the genetic factors contributing to vocabulary skills during the first three years of life and assessed the emergence of genetic associations with mid-childhood reading, verbal and non-verbal intelligence, using a similar approach to that in chapter 3. Here, I studied measures of expressive vocabulary at 15, 24 and 38 months of age, as well as receptive vocabulary at 38 months of age. The genetic architecture underlying these vocabulary measures was highly dynamic, with evidence for age- and ability-specific genetic factors. Genetic influences identified for expressive vocabulary at 15 months also affected expressive vocabulary at 24 months, but did not affect vocabulary, reading or cognition measures beyond this age. An independent genetic factor contributing to expressive vocabulary at 24 months played a role in verbal processes throughout early and later childhood, including reading and verbal intelligence. Consistent with my findings in chapter 3, also genetic influences underlying receptive vocabulary at 38 months accounted for phenotypic covariance with mid-childhood reading, verbal intelligence and performance intelligence, independent of genetic factors identified for expressive vocabulary. Thus, the genetic foundations of mid-childhood reading cognitive skills are diverse. They involve at least two independent genetic factors that emerge at different stages during early language development, where part of the genetic influences related to verbal cognitive skills can already be captured by measures of expressive vocabulary in children as young as two years of age.

In chapter 5, I aimed to identify SNPs associated with early vocabulary and, as part of follow-up analyses, studied whether there are genetic links of early vocabulary with several later-life cognition-related traits, anthropometric traits and childhood-onset neurodevelopmental disorders. For this study, I performed the largest meta-GWAS of expressive and receptive vocabulary between 15 and 38 months of age to date, with increased study power compared to a previous effort[1]. The meta-GWAS was embedded within the Early Genetics and Life Course Epidemiology (EAGLE) Consortium and included seven independent European population- or community-based cohorts with genome-wide genotypes and early vocabulary scores, resulting in 37,913 observations across 17,298 individuals. To allow for age- and ability-specific genetic influences, I followed a stratified design investigating early-phase expressive vocabulary (15-18 months, N=8,799); late-phase expressive vocabulary (24-38 months, N=16,615); and late-phase receptive vocabulary (24-38 months, N=6,291). Single-trait meta-analyses confirmed a known GWAS signal[1], but did not result in the identification of novel SNP association signals. The strongest association with early phase expressive vocabulary was observed at rs9854781 (P=4x10^{-8}), a SNP located within an intergenic region at chr3p12.3, ~20 kb downstream of the 3' end of *ROBO2*, consistent with findings from the previous meta-GWAS on early expressive vocabulary in a largely overlapping sample[1]. Analysis across genetically highly correlated single-trait vocabulary GWAS

summary statistics using multi-trait analysis of genome-wide association (MTAG) increased the statistical power, equivalent to analysing a sample size of 26,206 individuals, but did not yield additional evidence for genetic associations with vocabulary. As part of follow-up analyses, I estimated low SNP-h^2 for both early-phase and late-phase expressive vocabulary, while SNP-h^2 for late-phase receptive vocabulary was consistent with zero. Genetic correlation analyses with later-life cognition-related traits, anthropometric traits and childhood-onset neurodevelopmental disorders revealed weak-to-moderate positive genetic links of late-phase expressive vocabulary with subsequent reading, educational attainment (EA), and intelligence. These results imply that some genetic variation related to later-life cognition can already be tagged by genetic influences underlying expressive vocabulary between the ages of two and three years, supporting my findings from chapter 3 and 4.

In chapter 6, I studied whether there was evidence for genetic links between Attention-Deficit/Hyperactivity Disorder (ADHD) and multiple mid-childhood/early-adolescent language- and literacy-related abilities (LRAs) in the general population, using a polygenic scoring approach. Investigating summary statistics from two independent ADHD collections, I showed that increased ADHD polygenic risk was associated with lower language and literacy task performance in ALSPAC, especially for reading-related skills. Next, I studied to what extent this genetic overlap could be attributed to shared genetic influences with EA, a trait with an underlying genetic architecture that is genetically related to both ADHD and LRAs. Using a multivariable regression approach that makes it possible to disentangle genetic effects between multiple genetically correlated traits while controlling for a potential collider bias[2], I dissected the genetic relationship between ADHD and LRAs into genetic effects that are also shared with EA (implying pleiotropy among ADHD, EA and LRAs) as well as genetic effects that capture overlap between ADHD and LRAs independent of EA. Polygenic associations of ADHD with language and literacy skills were largely shared with EA. However, conditional on these shared genetic effects, there was evidence for ADHD-specific associations with literacy-related skills. These findings suggest genetic overlap of ADHD with literacy performance, in particular reading, beyond genetic effects that are also shared with EA.

In chapter 7, I studied the complex genetic overlap between ADHD, ASD and EA that simultaneously result in ASD-related positive and ADHD-related negative genetic correlations with EA, a powerful genetic proxy for language and literacy skills. I first investigated whether the discordant association profiles between polygenic risk for each disorder and EA were attributable to the same or different underlying genetic loci. Using summary statistic data from large consortia and applying a multivariable regression approach, I showed that EA-related genetic variation was regionally shared across ASD and ADHD genetic architectures, implying the same genetic variants. Next, I studied whether ASD and ADHD effects were encoded with respect to the same or the opposite allele at a single GWAS marker. At the polygenic level, the same set of risk alleles

captured opposite EA-related association profiles, with negative genetic associations when modelling ADHD SNP estimates and positive genetic associations when modelling ASD SNP estimates. These results suggest that ASD-related positive and ADHD-related negative genetic correlations with EA may involve biological pleiotropy and/or co-localisation processes, implying that a single GWAS marker can capture different ASD and ADHD risk alleles due to linkage disequilibrium (LD). Similar findings were observed when studying genetic overlap with reading performance, though the statistical power of this analyses was low, due to a smaller sample size. The largest opposite associations with EA were observed for 83 SNPs that were associated with both ASD and ADHD risk at $P<0.0015$, showing evidence for enrichment of regulatory elements, such as miRNA targets. Furthermore, shared EA-related genetic variation between ASD and ADHD suggests local negative genetic covariance that may contribute to the total genome-wide correlation between ASD and ADHD, potentially cancelling out independent positive genetic covariance patterns. Thus, my work showed that EA-related polygenic variation is shared across ASD and ADHD genetic architectures at the same genetic variants, involving combinations of the same risk alleles, although the encoded effects fundamentally differ and may, through mechanisms consistent with pleiotropy or co-localisation, encode ASD-related positive and ADHD-related negative associations with EA.

Together, this thesis presents novel research on the genetic architecture underlying early language development and, by investigating proxy measures, examines the mechanisms shaping genetic overlap with ASD and ADHD. Studies include genome-wide association analyses of early vocabulary and in-depth characterisations of its developmental genetic architecture. In addition, I studied genetic overlap of early language development with later cognitive skills, and, in turn, using these more powerful traits as proxies, I investigated complex genetic mechanisms linking language and literacy skills with several childhood-onset neurodevelopmental disorders. Next, I discuss the broader implications of my work for future research. I will focus on the following topics (i) developmental stability and change in genetic factors related to early vocabulary, (ii) statistical power, (iii) biological interpretation of genome-wide association signals, (iv) pleiotropy and related mechanisms, and (v) potential sources of bias in genetic associations.

8.2. Developmental stability and change in genetic factors related to early vocabulary

The findings described in this thesis support both stability and change of language-related genetic factors during the course of child and adolescent development. In chapter 4, I showed that the genetic architecture underlying expressive and receptive

vocabulary during the first three years of life is highly dynamic, with evidence for genetic factors showing both age- and ability-specific effects. This pattern is consistent with rapid developmental changes that occur at the behavioural level during early life. Whereas children produce only a few single words around the age of 15 months[3], their expressive vocabulary has markedly increased to 100-600 words by the age of 24 months[4] with the concurrent production of first word combinations[5,6]. In addition, the development of receptive vocabulary skills is thought to precede the development of expressive skills[7] and the size of children's receptive vocabulary by far exceeds their expressive vocabulary[8]. Nonetheless, based on the outcomes of my studies, there is now evidence that from early childhood to early adolescence, genetic stability predominates over innovation processes, as presented in chapters 3 and 4. In particular, genetic influences underlying receptive vocabulary at 38 months, independent of expressive vocabulary, accounted for the majority of SNP-h^2 in many mid-childhood/early-adolescent language and literacy skills, suggesting genetic amplification. Previous research in twins on the genetic architecture underlying language and literacy development investigated links with respect to a latent factor consisting of early expressive language abilities only and found that the majority of genetic influences related to mid-childhood and adolescent language and reading skills could be attributed to innovation processes[9,10]. However, amplified receptive vocabulary-related genetic influences would have been missing from such an early latent factor, as they were not included in the study and largely independent of genetic sources underlying expressive vocabulary. These findings highlight the importance of studies researching the role of early receptive language skills and their shared genetic links with later cognitive skills.

Part of the genetic covariance between receptive vocabulary at 38 months and subsequent mid-childhood/early-adolescent language and literacy skills could already be captured by genetic factors contributing to vocabulary skills emerging even earlier during development. Longitudinal GSEM analyses of early vocabulary identified evidence for the presence of such links for a measure of expressive vocabulary at 24 months of age (chapter 4). These findings are consistent with my genome-wide genetic correlation analyses using GWAS summary statistic data derived from meta-analysing measures of expressive vocabulary between 24 and 38 months, where I identified weak-to-moderate positive genetic correlations with mid-childhood to early-adulthood reading, general intelligence across the lifespan and adult EA (chapter 5). Association patterns with later-life reading as well as verbal and non-verbal cognition, as reported in chapter 4, suggested that the earliest detectable genetic links with expressive vocabulary at 24 months might be driven by processes related to later-life verbal abilities, based on analyses of individual level genotype data from the ALSPAC sample. However, the power of these analyses was limited and 95%-confidence intervals of estimated factor loadings overlapped with those for later-life performance intelligence. The identified genetic links between infant/toddler vocabulary and subsequent language, literacy and cognition also

strengthen previously established evidence for longitudinal relationships reported in observational studies[11–14]. For example, broadly defined early oral language skills, including receptive abilities, have been shown to affect word recognition[11], while vocabulary comprehension is also a precursor of listening comprehension[12]. According to the "simple view of reading" theory, reading comprehension is the product of printed word recognition and oral language comprehension[15]. Consistent with this theoretical framework, early language skills at the age of three, including vocabulary, comprehension and sentence construction, have been linked to adolescent reading comprehension[13]. In addition, a delay in both expressive and receptive vocabulary at the age of two years is much more likely to lead to problems with later literacy, compared to delays in expressive vocabulary alone[14]. Variation in language comprehension has also been associated with more non-linguistic cognitive measures, such as tool use and symbolic play, compared to expressive vocabulary[16]. Thus, genetic variation influencing receptive vocabulary at 38 months may share genetic foundations with several key skills that are important for future cognitive development and only partially overlap with genetically predictable cognitive mechanisms underlying early expressive vocabulary. Nonetheless, the stability in genetic influences related to both early expressive and receptive vocabulary across developmental periods suggests that their joint analysis will increase power to detect shared genetic influences, as applied in chapter 5 and discussed in more detail in section 8.3.

To gain further insight into how genetic factors captured by early language skills relate to other aspects of development, future studies could consider investigating behavioural and motor skills that are acquired during the first years of life[17]. In particular, the onset of walking is thought to initiate development in social interactions, personal-social skills and language learning[18–22]. For example, an earlier age of mastering motor skills, such as sitting and walking, is associated with earlier attainment of language developmental milestones[20–22]. This relationship might reflect an increased level of interaction between children and their mothers once children are able to walk independently, both in terms of quantity and quality, where the latter may include pointing behaviour to direct mothers' attention to particular objects[23]. To investigate whether interlinked motor, social and language developmental processes are also related at the genetic level and to assess the role of pre-linguistic abilities on language development, additional research based on large longitudinal samples is required.

8.3. Statistical power

To identify SNPs associated with vocabulary during the first three years of life, I performed a meta-GWAS in chapter 5. The study presented in this chapter extended a previous meta-GWAS effort on early expressive vocabulary[1] by increasing the sample

size by almost 50%. In addition, I carried out the first GWAS on receptive vocabulary during the first few years of life. Finally, I maximised the statistical power by adopting a multivariate analysis approach implemented in MTAG to analyse genetically related expressive and receptive vocabulary traits. The power of this multi-trait vocabulary analyses corresponded to an estimated sample size of 26,206 children, but did not result in the identification of genome-wide significant associations. Although a sample of this size has 99% power to detect association with a genetic variant explaining as little as 0.3% of the trait variance (assuming an additive model and an increaser allele frequency of 0.1, with complete LD with marker and genetic risk variant)[24], the power to detect variants with smaller contributions to trait variance is only modest (e.g. 37% power to detect a genetic variant explaining 0.1% of the trait variance)[24]. In addition, MTAG analyses may provide biased estimates and have high maximum false discovery rates, especially when combining traits with low inherent power based on sample size and polygenic signal captured (mean $\chi^2 < 1.02$)[25], such as some of the vocabulary GWAS statistics derived in chapter 5, although I did not observe evidence for such bias. However, despite an increase in statistical power compared to the previous meta-GWAS effort[1], current sample sizes are still underpowered for identifying genetic variants with small effects on early vocabulary development.

Compared to data on expressive vocabulary, the available information on receptive language skills is limited, both with respect to available psychological instruments and data collections (chapter 2). For children below the age of two years, low validity of receptive vocabulary assessments has been reported, based on a comparison of parent-report with child performance on a preferential looking task[26]. In addition, I observed little evidence for SNP-h^2 for a measure of receptive vocabulary size at 15 months (chapter 4), potentially reflecting a relatively high random error rate in phenotypic assessment[27] that would be consistent with low validity and/or reliability of parental report. However, there is evidence for validity of parental-reported receptive vocabulary scores in older children. A study comparing parental assessment and child task performance for receptive vocabulary in 25-month old children reported a correlation of 0.55[28]. In addition, I observed evidence for low SNP-h^2 of receptive vocabulary at 38 months, as well as strong genetic relationships with expressive vocabulary at 24 and 38 months, suggesting utility of receptive vocabulary scores assessed in early childhood (chapter 4). Nonetheless, receptive vocabulary scales are absent from many frequently used psychological instruments assessing early language development, including forms of the MacArthur vocabulary scales (e.g. the MacArthur Communicative Development Inventory: Words & Sentences[29] (CDI-WS)), but also the Language Development Survey[30] (LDS), prohibiting investigation of the underlying genetic factors. Opportunities for reliable large-scale assessments of early receptive vocabulary development could involve the collection of app-based vocabulary data in combination with DNA samples. For expressive vocabulary size, this data ascertainment

route has already been successfully established, with app-based assessments showing high reliability and validity, even when based on 25 words only (drawn from CDI-WS items)[31]. Such approaches are cost- and time-effective and thus hold great promise for future sample collections.

Another way to increase power could be by adopting a broader definition of early language skills. An observational study of 1,137 children showed that oral language conceptualised as vocabulary, grammar and semantics had increased predictive value for reading skills compared to measures of vocabulary alone[11]. At the genetic level, vocabulary and grammar in two- and three-year olds are moderately-to-strongly genetically correlated[32,33], suggesting shared underlying factors. This supports joint analyses across language and grammatical scores of the MacArthur CDI[5]. Other questionnaires assessing broad aspects of early language competence include, for example, the Bayley Scales of Infant and Toddler Development[34] and the Parent Report of Children's Abilities-Revised[35]. Cohorts with data on such questionnaires and DNA samples may be approached to participate in future genetic studies. Importantly, genetic analyses can be performed using language data collected in different countries. Children follow similar patterns of language acquisition across various languages[36] and CDI vocabulary assessments are comparable across different cultures[37]. Furthermore, there was little evidence for heterogeneity in SNP associations with early expressive vocabulary across English and Dutch languages[1]. In addition to studies across different language skills in early development, it might also be worth considering a joint analysis of early vocabulary skills with several language and literacy abilities assessed later in life. Based on my findings in chapter 3, 4 and 5, there is evidence for substantial genetic overlap between early-childhood vocabulary and for example reading performance. Thus, broadening the definition of early language skills as well as joint analysis with genetically correlated traits assessed later in life may increase the power to detect genetic factors underlying shared developmental processes of early language development, although it may limit the identification of vocabulary-specific genetic influences.

8.4. Biological interpretation of genome-wide association signals

Fixed-effect meta-analysis of early phase expressive vocabulary (15-18 months), a developmental window during which children typically speak words in isolation[3], provided evidence for association with rs9854781. This genetic variant is located within an intergenic region at chr3p12.3, ~20kb near *ROBO2*. It is located only 976 base pairs apart from, and in high linkage disequilibrium (LD-r^2=0.78) with, rs764282, a known signal for early expressive vocabulary reported by the prior meta-GWAS effort using largely similar samples[1]. *ROBO2* is highly expressed across different human brain regions,

according to the Genotype-Tissue Expression project[38] (GTEx, v8). During the course of development, *ROBO2* expression peaks in the first trimester[1] and has been associated with new-born neurons[39]. Analyses of gene expression and regulatory chromatin states indicated that variation at rs7642482 might be related to regulatory mechanisms in embryonic cell types[1]. However, an experimental validation of the functional relevance of rs764282 on *ROBO2* is still lacking and searches based on the latest blood and (fetal) brain gene-expression data sets did not provide evidence for an effect of rs764282 (chapter 2). Future experiments of developing neuronal cells are required to characterise the expression pattern of *ROBO2* in great detail and to investigate the consequences of rs764282.

8.5. Pleiotropy and related mechanisms

Pleiotropy arises when a gene or genetic variant influences more than one trait[40]. Genetic links between early-life vocabulary measures and subsequent language and literacy skills, as well as other cognition-related later-life outcomes (chapters 3, 4 and 5) are thus consistent with pleiotropy. Longitudinal genetic links between early vocabulary and later language and literacy abilities (chapters 3 and 4) may reflect a specific form of pleiotropy, mediated pleiotropy, that arises due to a causal association between two traits[40], in line with evidence for longitudinal relationships from observational studies[11–14] (see section 8.2). Beside mediated pleiotropy, genetic overlap may occur due to other processes. For example, biological pleiotropy captures a direct biological influence of a genetic factor on multiple phenotypes[40]. Furthermore, both mediated and biological pleiotropy might be mimicked by genetic confounding, a process where genetic associations between two traits are induced due to their genetic correlations with a third phenotype[41], or by co-localisation, where different risk variants are tagged by the same GWAS marker allele, due to high LD[40]. In addition, there could be a false association between a genetic factor and multiple traits due to various sources of bias, so-called spurious pleiotropy[40].

In chapter 6, I observed strong and consistent evidence for genetic links between ADHD and mid-childhood/early-adolescent language and literacy (as proxy for early-childhood language), consistent with previous reports[42–45]. I showed that this genetic association was, to a large extent, attributable to genetic effects that were also shared with EA, consistent with genetic confounding. The results imply that genetic associations of ADHD with language and literacy skills might be inflated due to a relationship of both ADHD and LRAs, such as reading abilities, with genetically predicted EA[45–48]. A similar association profile was found between schizophrenia, bipolar disorder and EA, where the genetic correlation between schiozphrenia and EA could be attributed to shared genetic effects between schizophrenia and bipolar disorder[49]. Other possible

explanations for shared genetic variance between ADHD, EA, language and literacy abilities may involve an intergenerational multiple-deficit model proposed for reading disability[50,51]. For example, children growing up in disadvantaged environments, genetically predictable through polygenic EA scores[52], might be more vulnerable to psychiatric illness including ADHD[53] that affects, in turn, their language and literacy performance. In addition, adolescents with ADHD might be more likely to leave school at an earlier age, with lower educational achievement and, subsequently, pass on an increased genetic load to their own children[54]. Conditionally on shared genetic effects with EA, I observed evidence for genetic links between ADHD and especially reading abilities. Thus, their genetic overlap cannot be fully attributed to shared genetic effects with EA, possibly reflecting other forms of pleiotropy. These findings suggest that there is a potential need to target reading skills for improvement in children diagnosed with ADHD, beyond general training programmes aiming broadly at enhancing schooling outcomes[55].

In chapter 7, I investigated complex genetic mechanisms underlying the discordant association profile of two positively genetically correlated neurodevelopmental disorders, ASD and ADHD, with EA; with positive genetic correlations between ASD and EA, and negative genetic correlations between ADHD and EA. In this study, EA was included as a genetic proxy for language and literacy skills, supported by high genetic correlations of childhood language and literacy skills with EA as reported in chapter 6. The pattern of results in chapter 7 suggested that discordant genetic association patterns with EA are encoded at the same GWAS marker alleles for both ASD and ADHD risk, suggesting biological pleiotropy and/or co-localisation. The 83 genetic variants capturing the strongest association effects were enriched for miRNA targets, and also comprised several miRNA and lncRNA loci, but replication of these findings is warranted. These findings suggest multiple regulatory sites in close genomic proximity or different regulations of the same genetic markers via epistatic or gene-environment interaction effects[56]. Although the power to carry out similar investigations for early vocabulary and mid-childhood language and reading skills was too low, it can be speculated whether similar gene sets for language and literacy performance are also implicated in ASD and ADHD genetic architectures, but involve different regulation. For example, I observed evidence for a positive genetic link of ASD risk-increasing-variants and reading performance when accounting for shared genetic effects with ADHD, although this finding needs to be replicated (chapter 7). In addition, extensive research of genetic overlap is required to uncover links of early vocabulary with neurodevelopmental disorders that are characterised by a primary deficit in speech and/or language abilities, once large-scale genetic studies for these disorders have become available. Such efforts could, for example, include studies of developmental dyslexia (also known as reading disability) and specific language impairment (also known as developmental language disorder).

The identification of discordant genetic association profiles with EA for polygenic ASD and ADHD risk, encoded at the same marker alleles, suggests local negative genetic covariance between ASD and ADHD (chapter 7). Negative covariance has been previously observed among other complex traits[57], even in the absence of genome-wide correlation[58]. Thus, genetic correlation patterns between ASD and ADHD may vary across the genome. Local negative genetic covariance may also contribute to genetic overlap of ASD and ADHD risk with other neuropsychiatric disorders, such as schizophrenia, major depressive disorder and bipolar disorder (chapter 7). More generally, patterns of local genetic covariance may also play a role in genetic relationships between many other human traits, including psychiatric disorders and brain phenotypes[59]. Characterising parts of the genome that are positively and negatively correlated with each other may provide additional insight into the biological mechanisms underlying genetic links between traits. This can for example be achieved by applying computational approaches that decompose shared and specific genetic effects, such as GSEM[60]. An alternative approach, genomic structural equation modelling[61], is based on GWAS summary statistic data and might therefore be more widely applicable, although it is less precise, as it is based on estimated genetic variance/co-variance structures only. However, shared genetic variance between traits does not automatically imply similar biological processes, but may instead capture genetic confounding, biological pleiotropy and/or co-localisation. Genes are known to have multiple biological functions, and dynamic gene expression patterns over time and space have been shown for multiple brain-related gene expression modules[62]. Thus, specifically designed gene-based studies are warranted to reveal the actual biological processes captured by genetic variance/co-variance structures.

8.6. Potential sources of bias in genetic associations

Recently, the influence of population-based phenomena including dynastic effects, population stratification and assortative mating (chapter 2) on heritability estimates and genetic correlations has gained attention in the literature[63–66]. Studies of dynastic effects primarily investigated the role of genetically predicted parental EA on offspring phenotype via modulation of the environmental niche. EA is genetically related to many other traits, including language and literacy skills (chapter 5, 6), cognitive ability and family socioeconomic status[67]. The latter is known to have an influence on the quality and quantity of language input a child receives in their home environment[68]. This input, in turn, is related to language acquisition[69] and indeed, language processing was attenuated in 18 month old infants that grew up in families with low socioeconomic status compared to high socioeconomic status families[70]. Thus, parents who carry a higher load of EA increasing genetic variants may create more supportive environments

for language learning than parents with a lower load, which will affect their children's performance indirectly. Many methods estimating genetic variance components based on genome-wide information from unrelated individuals, including polygenic scoring[71], GCTA restricted maximum likelihood analyses[72–74] and GSEM[60], as applied within this thesis, will capture dynastic effects as part of additive genetic variation[63–66]. This may lead to inflated estimates of heritability and genetic overlap, especially for cognition-related traits when studying samples of unrelated individuals[64]. Other population phenomena that may bias associations include population stratification and assortative mating. In this thesis, I accounted for population stratification by adjusting for the first two principal components, capturing genetic differences due to ancestry. However, adjustment for up to twenty[63] or even more principal components might be required to fully control for population structure[75]. Finally, there is evidence for assortative mating on EA within ALSPAC, which may lead to increased heritability estimates[63]. Compared to population stratification or assortative mating, however, dynastic effects related to family socioeconomic status have been found to be the major source of inflated genetic associations with cognition-related traits[64]. So far, the extent to which dynastic effects, population stratification and assortative mating affect genetic studies on language development is unknown. Future investigations of large family-based samples with spouse, sibling and parent/child information as well as language data, such as the Twins Early Develpoment Study[76] and the Netherlands Twin Register[77], could specifically address this bias.

8.7. Conclusion

The studies of common genetic markers in this thesis show that several processes underlying early language development are genetically shared with multiple later language and literacy skills, as well as later-life cognition-related abilities such as adulthood EA, forming an early biological foundation. The mechanisms linking early vocabulary skills to subsequent language and literacy abilities are likely to involve amplification processes and mediated pleiotropy. However, the genetic overlap of language and literacy-related skills with several childhood-onset neurodevelopmental disorders is more complex, involving mechanisms of biological pleiotropy, co-localisation and/or shared genetic variation with a third variable such as EA. To further our knowledge of aetiological mechanisms underlying early language development, future studies may consider adopting a broader phenotype definition. It is, however, also important to carry out analyses that allow for developmental genetic heterogeneity, which is characteristic for early vocabulary development. Thus, the genetic landscape underlying early-life vocabulary is dynamic and includes genetic components that serve as foundation for subsequent cognitive development.

References

1. St Pourcain, B. *et al.* Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature communications* **5**, 4831 (2014).
2. Burgess, S. *et al.* Dissecting Causal Pathways Using Mendelian Randomization with Summarized Genetic Data: Application to Age at Menarche and Risk of Breast Cancer. *Genetics* genetics.300191.2017 (2017) doi:10.1534/genetics.117.300191.
3. Clark, E. V. *First language acquisition*. (Cambridge University Press, 2016).
4. Goldfield, B. A. & Reznick, J. S. Early lexical acquisition: rate, content, and the vocabulary spurt. *J Child Lang* **17**, 171–183 (1990).
5. Fenson, L. *et al.* Variability in early communicative development. *Monogr Soc Res Child Dev* **59**, 1–173; discussion 174-85 (1994).
6. Hoff, E. *Language Development*. (Cengage Learning, 2013).
7. Bergelson, E. & Swingley, D. At 6-9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3253–3258 (2012).
8. Burger, A. & Chong, I. Receptive Vocabulary. in *Encyclopedia of Child Behavior and Development* (eds. Goldstein, S. & Naglieri, J. A.) 1231–1231 (Springer US, 2011). doi:10.1007/978-0-387-79061-9_2359.
9. Harlaar, N., Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. Why Do Preschool Language Abilities Correlate With Later Reading? A Twin Study. *J Speech Lang Hear Res* **51**, 688–705 (2008).
10. Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev Sci* **15**, 233–49 (2012).
11. NICHD Early Child Care Research Network. Pathways to reading: the role of oral language in the transition to reading. *Dev Psychol* **41**, 428–442 (2005).
12. Dickinson, D. K. & Neuman, S. B. *Handbook of Early Literacy Research*. (Guilford Press, 2007).
13. Frost, J., Madsbjerg, S., Niedersøe, J., Olofsson, Å. & Sørensen, P. M. Semantic and phonological skills in predicting reading development: from 3–16 years of age. *Dyslexia* **11**, 79–92 (2005).
14. Psyridou, M., Eklund, K., Poikkeus, A.-M. & Torppa, M. Reading outcomes of children with delayed early vocabulary: A follow-up from age 2–16. *Research in Developmental Disabilities* **78**, 114–124 (2018).
15. Gough, P. B. & Tunmer, W. E. Decoding, Reading, and Reading Disability. *Remedial and Special Education* **7**, 6–10 (1986).
16. Bates, E., Dale, P. S. & Thal, D. Individual Differences and their Implications for Theories of Language Development. in *The Handbook of Child Language* 95–151 (John Wiley & Sons, Ltd, 2019). doi:10.1111/b.9780631203124.1996.00005.x.
17. Iverson, J. M. Developing language in a developing body: the relationship between motor development and language development. *J Child Lang* **37**, 229–261 (2010).
18. Libertus, K. & Violi, D. A. Sit to Talk: Relation between Motor Skills and Language Development in Infancy. *Front. Psychol.* **7**, (2016).
19. Walle, E. A. & Campos, J. J. Infant language development is related to the acquisition of walking. *Dev Psychol* **50**, 336–348 (2014).
20. Flensborg-Madsen, T. & Mortensen, E. L. Developmental milestones during the first three years as precursors of adult intelligence. *Dev Psychol* **54**, 1434–1444 (2018).
21. Flensborg-Madsen, T., Grønkjær, M. & Mortensen, E. L. Predictors of early life milestones: Results from the Copenhagen Perinatal Cohort. *BMC Pediatrics* **19**, 420 (2019).
22. Oudgenoeg-Paz, O., Volman, M. C. J. M. & Leseman, P. P. M. Attainment of sitting and walking predicts development of productive vocabulary between ages 16 and 28 months. *Infant Behav Dev* **35**, 733–736 (2012).
23. Clearfield, M. W. Learning to walk changes infants' social interactions. *Infant Behavior and Development* **34**, 15–25 (2011).
24. Purcell, S., Cherny, S. S. & Sham, P. C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003).
25. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics* **50**, 229–237 (2018).

26. Houston-Price, C., Mather, E. & Sakkalou, E. Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *J Child Lang* **34**, 701–724 (2007).
27. Walters, R., Churchhouse, C. & Hosking, R. Heritability 101: What is 'heritability'? http://www.nealelab.is/blog/2017/9/13/heritability-101-what-is-heritability (2019).
28. Ring, E. D. & Fenson, L. The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language* **20**, 141–159 (2000).
29. Fenson, L. *et al. User's Guide and Technical Manual for the MacArthur Communicative Development Inventories*. (Singular Publishing, 1993).
30. Rescorla Leslie. The Language Development Survey. *Journal of Speech and Hearing Disorders* **54**, 587–599 (1989).
31. Mayor, J. & Mani, N. A short version of the MacArthur–Bates Communicative Development Inventories with high validity. *Behav Res* **51**, 2248–2255 (2019).
32. Dale, P. S., Dionne, G., Eley, T. C. & Plomin, R. Lexical and grammatical development: a behavioural genetic perspective. *Journal of Child Language* **27**, 619–642 (2000).
33. Dionne, G., Dale, P. S., Boivin, M. & Plomin, R. Genetic Evidence for Bidirectional Effects of Early Lexical and Grammatical Development. *Child Development* **74**, 394–412 (2003).
34. Albers, C. A. & Grieve, A. J. Test Review: Bayley, N. (2006). Bayley Scales of Infant and Toddler Development– Third Edition. San Antonio, TX: Harcourt Assessment: *Journal of Psychoeducational Assessment* (2016) doi:10.1177/0734282906297199.
35. Johnson, S. *et al.* Standardisation of the Parent Report of Children's Abilities–Revised (PARCA-R): a norm-referenced assessment of cognitive and language development at age 2 years. *The Lancet Child & Adolescent Health* **3**, 705–712 (2019).
36. Dale, P. S. & Goodman, J. C. Commonality and Individual Differences in Vocabulary Growth. in *Beyond nature-nurture: Essays in honor of Elizabeth Bates* 41–78 (Lawrence Erlbaum Associates Publishers, 2005).
37. Bleses, D. *et al.* Early vocabulary development in Danish and other languages: a CDI-based comparison. *J Child Lang* **35**, 619–650 (2008).
38. Aguet, F. *et al.* The GTEx Consortium atlas of genetic regulatory effects across human tissues. *bioRxiv* 787903 (2019) doi:10.1101/787903.
39. Pollen, A. A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nature Biotechnology* **32**, 1053–1058 (2014).
40. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* **14**, 483–495 (2013).
41. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nature Genetics* 1–10 (2019) doi:10.1038/s41588-019-0481-0.
42. Martin, N. C., Levy, F., Pieka, J. & Hay, D. A. A Genetic Study of Attention Deficit Hyperactivity Disorder, Conduct Disorder, Oppositional Defiant Disorder and Reading Disability: Aetiological overlaps and implications. *International Journal of Disability, Development and Education* **53**, 21–34 (2006).
43. Willcutt, E. G., Pennington, B. F. & DeFries, J. C. Twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **96**, 293–301 (2000).
44. Willcutt, E. G., Pennington, B. F., Olson, R. K. & DeFries, J. C. Understanding comorbidity: A twin study of reading disability and attention-deficit/hyperactivity disorder. *Am. J. Med. Genet.* **144B**, 709–714 (2007).
45. Gialluisi, A. *et al.* Genome-wide association scan identifies new variants associated with a cognitive predictor of dyslexia. *Transl Psychiatry* **9**, (2019).
46. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics* 1 (2018) doi:10.1038/s41588-018-0269-7.
47. Selzam, S. *et al.* Genome-Wide Polygenic Scores Predict Reading Performance Throughout the School Years. *Sci Stud Read* 1–16 (2017) doi:10.1080/10888438.2017.1299152.
48. Luciano, M. *et al.* Single Nucleotide Polymorphisms Associated with Reading Ability Show Connection to Socio-Economic Outcomes. *Behav Genet* **47**, 469–479 (2017).
49. Nieuwboer, H. A., Pool, R., Dolan, C. V., Boomsma, D. I. & Nivard, M. G. GWIS: Genome-Wide Inferred Statistics for Functions of Multiple Phenotypes. *Am. J. Hum. Genet.* **99**, 917–927 (2016).

50. Pennington, B. F. From single to multiple deficit models of developmental disorders. *Cognition* **101**, 385–413 (2006).
51. van Bergen, E., van der Leij, A. & de Jong, P. F. The intergenerational multiple deficit model and the case of dyslexia. *Front Hum Neurosci* **8**, 346 (2014).
52. Ayorech, Z., Krapohl, E., Plomin, R. & Stumm, S. von. Genetic Influence on Intergenerational Educational Attainment. *Psychol Sci* **28**, 1302–1310 (2017).
53. Reiss, F. Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. *Soc Sci Med* **90**, 24–31 (2013).
54. Russell, G., Ford, T., Rosenberg, R. & Kelly, S. The association of attention deficit hyperactivity disorder with socioeconomic disadvantage: alternative explanations and evidence. *J Child Psychol Psychiatr* **55**, 436–445 (2014).
55. Daley, D. & Birchwood, J. ADHD and academic performance: why does ADHD impact on academic performance and what can be done to support ADHD children in the classroom? *Child: Care, Health and Development* **36**, 455–464 (2010).
56. Lee, P. H. *et al.* Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482.e11 (2019).
57. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *The American Journal of Human Genetics* **101**, 737–751 (2017).
58. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* **20**, 567–581 (2019).
59. Brainstorm Consortium *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**, (2018).
60. St Pourcain, B. *et al.* Developmental changes within the genetic architecture of social communication behaviour: A multivariate study of genetic variance in unrelated individuals. *Biological Psychiatry* **83**, 598–606 (2017).
61. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour* **3**, 513–525 (2019).
62. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, (2018).
63. Morris, T. T., Davies, N. M., Hemani, G. & Smith, G. D. Why are education, socioeconomic position and intelligence genetically correlated? *bioRxiv* 630426 (2019) doi:10.1101/630426.
64. Selzam, S. *et al.* Comparing Within- and Between-Family Polygenic Score Prediction. *The American Journal of Human Genetics* **105**, 351–363 (2019).
65. Bates, T. C. *et al.* The Nature of Nurture: Using a Virtual-Parent Design to Test Parenting Effects on Children's Educational Attainment in Genotyped Families. *Twin Research and Human Genetics* **21**, 73–83 (2018).
66. Kong, A. *et al.* The nature of nurture: Effects of parental genotypes. *Science* **359**, 424–428 (2018).
67. Selzam, S. *et al.* Predicting educational achievement from DNA. *Mol. Psychiatry* **22**, 267–272 (2017).
68. Hoff, E. The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Dev* **74**, 1368–1378 (2003).
69. Hoff, E. How social contexts support and shape language development. *Developmental Review* **26**, 55–88 (2006).
70. Fernald, A., Marchman, V. A. & Weisleder, A. SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science* **16**, 234–248 (2013).
71. Choi, S. W., Mak, T. S.-H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nature Protocols* 1–14 (2020) doi:10.1038/s41596-020-0353-1.
72. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
73. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
74. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540–2542 (2012).

75. Haworth, S. *et al.* Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nature Communications* **10**, 333 (2019).
76. Haworth, C. M. A., Davis, O. S. P. & Plomin, R. Twins Early Development Study (TEDS): A Genetically Sensitive Investigation of Cognitive and Behavioral Development From Childhood to Young Adulthood. *Twin Research and Human Genetics* **16**, 117–125 (2013).
77. Boomsma, D. I. *et al.* Netherlands Twin Register: From Twins to Twin Families. *Twin Research and Human Genetics* **9**, 849–857 (2006).

Summary

Language is a complex human capacity that allows us to acquire knowledge, share thoughts, convey feelings, and report experiences. Mastering the components of language is a complex task and there are large individual differences in children's language abilities during the first few years of life, which are predictive of future language development, literacy performance and cognitive skills. In addition, children diagnosed with a neurodevelopmental disorder, such as Attention-Deficit/Hyperactivity Disorder (ADHD) or Autism Spectrum Disorder (ASD), often experience problems with language and cognitive development.

Individual differences in early language development are a consequence of variation in both genetic and environmental factors, as described in chapters 1 and 2. Previous studies showed that a small proportion of variation in language skills during the first three years of life is attributable to genetic factors, including common genetic variation. Using latent factor constructs, twin studies demonstrated that some of the underlying genetic influences remain stable from toddlerhood to early adolescence and affect later language and literacy skills, in addition to age-specific genetic contributions. However, beyond these findings, the genetic landscape of early language development is little characterised. In particular, developmental changes in genetic factor structures across early life and the role of receptive language skills are poorly understood. Molecular research of early language has so far been limited to a single meta- genome-wide association study (meta-GWAS) that identified a genome-wide signal for expressive vocabulary size at 15-18 months near *ROBO2*, a gene involved in axon guidance. In this thesis, I analysed genome-wide genotype information to broaden knowledge of genetic factors underlying language abilities during the first three years of life, and their role during subsequent language, literacy and cognitive development, as well as links of language and related traits with childhood-onset neurodevelopmental disorders, such as ADHD and ASD.

In chapter 3, I studied whether genetic factors underlying vocabulary skills during toddlerhood are related to mid-childhood/early-adolescent language and literacy abilities. Using a longitudinal structural equation modelling approach investigating genetic relationship information from unrelated children, I showed that the majority of genetic influences contributing to mid-childhood/early-adolescent language and literacy skills can already be captured by genetic factors underlying expressive and especially receptive vocabulary during early childhood. The phenotypic variation explained by these early genetic influences increased during development, suggesting not only genetic stability but amplification. Meta-regression analyses of genetic factor contributions across a spectrum of language and literacy abilities showed that these amplification patterns were strongest for mid-childhood reading-related measures.

In chapter 4, I investigated the developmental genetic architecture of early-life vocabulary skills in depth, using a similar structural equation modelling approach as in chapter 3. I showed that the genetic landscape underlying expressive and receptive vocabulary development from 15 to 38 months of age is dynamic, involving multiple independent age- and/or ability-specific genetic factors. Two of these genetic factors were also associated with mid-childhood reading and cognitive skills. Genetic influences identified for receptive vocabulary at 38 months accounted for the majority of genetic variation in mid-childhood reading and cognitive skills, consistent with my findings in chapter 3. However, genetic links with mid-childhood verbal processes, such as reading and verbal intelligence, could already be detected for expressive vocabulary at the age of two years. Thus, the genetic foundations of mid-childhood reading and cognitive skills involve at least two independent genetic factors that emerge at different stages during early language development.

In chapter 5, I performed the most powerful meta-GWAS of early-life expressive and receptive vocabulary (15 and 38 months of age) to date. My analyses confirmed the previously identified GWAS signal at *ROBO2*, but did not result in the identification of novel SNPs. Genetic correlation analyses provided evidence for a genetic link between expressive vocabulary from 24 to 38 months of age and later-life verbal and cognition-related traits, such as reading, educational attainment and intelligence, confirming my findings from chapters 3 and 4 using a GWAS approach.

In chapter 6, I studied the complex pattern of genetic overlap between ADHD and mid-childhood language- and literacy-related abilities, which are powerful genetic proxies of early-life language abilities. Applying a polygenic scoring approach, I showed that increased polygenic ADHD risk was associated with lower language and literacy task performance, especially for reading-related skills. Using a multivariable regression approach, analogous to Mendelian Randomization methodologies, I showed that polygenic associations of ADHD with language and literacy skills could be largely attributed to shared genetic variation with educational attainment, a trait that is genetically related to both ADHD and language- and literacy-related abilities. Conditional on these shared genetic effects, risk for ADHD was predominantly linked to reading skills. Thus, my findings suggest that ADHD and reading performance share genetic factors beyond those that are related to educational attainment.

In chapter 7, I investigated the genetic overlap between ASD, ADHD and educational attainment by disentangling genetic mechanisms that may result in known opposite genetic correlation patterns between educational attainment and both ASD (positive links) and ADHD (negative links), given a positive genetic correlation between ASD and ADHD. I studied these discordant association profiles in detail using summary statistic data and a multivariable regression approach. This analysis demonstrated that educational attainment-related polygenic variation is shared between ASD and ADHD. I showed that different combinations of the same ASD and ADHD risk-increasing alleles

can simultaneously re-capture known ASD-related positive and ADHD-related negative associations with educational attainment. These findings suggest pleiotropic mechanisms and/or co-localisation of different risk variants, where the same polygenic sites can encode multiple independent, even discordant, association patterns.

The work presented in this thesis shows that processes underlying early language abilities are genetically complex and partially share genetic factors with multiple later-life language and literacy skills as well as cognition-related abilities. Furthermore, I showed that the genetic overlap of mid-childhood/early-adolescent language and literacy-related skills with neurodevelopmental disorders is highly intricate, involving mechanisms of biological pleiotropy, co-localisation and/or shared genetic variation with other traits. Based on the findings in this thesis, future studies of early language acquisition may consider adopting a broader phenotype definition of early language, including receptive vocabulary skills, while allowing for genetic heterogeneity, which is characteristic of early language development.

Nederlandse samenvatting

Taal is een complexe menselijke eigenschap die ons in staat stelt nieuwe kennis te vergaren, en ook gedachten, gevoelens en ervaringen te delen. Het aanleren van de verschillende componenten van taal is ingewikkeld en er bestaan grote individuele verschillen in de taalvaardigheden van kinderen in de eerste levensjaren. Deze verschillen zijn voorspellend voor taalvaardigheden, geletterdheid en cognitieve capaciteiten later in het leven. Daarnaast ervaren kinderen die gediagnosticeerd zijn met een ontwikkelingsstoornis als Attention-Deficit/Hyperactivity Disorder (ADHD) of Autism Spectrum Disorder (ASD, ookwel autisme genoemd) vaak problemen met taal en cognitieve vaardigheden.

Individuele verschillen in de taalontwikkeling van jonge kinderen zijn een gevolg van variatie in zowel genetische als omgevingsfactoren, zoals beschreven in de hoofdstukken 1 en 2. Eerdere studies hebben aangetoond dat een klein deel van de variatie in taalvaardigheden in jonge kinderen kan worden verklaard door genetische factoren, onder andere door genetische varianten die veelvoorkomend zijn in de populatie. Tweelingstudies hebben zowel stabiliteit als specificiteit gerapporteerd voor genetische factoren die van invloed zijn op taal- en leesvaardigheden van de peutertijd tot aan adolescentie. Buiten deze tweelingstudies om, is de kennis over het genetische landschap onderliggend aan vroege taalontwikkeling beperkt. Er is met name weinig kennis over veranderingen in genetische factoren die van invloed zijn op taalontwikkeling in de eerste paar levensjaren en de rol van receptieve taalvaardigheden. Moleculair onderzoek gericht op taalontwikkeling in jonge kinderen bestond tot nu toe uit een enkele meta-analyse van genoom-wijde associatie studies (meta-GWAS). Deze studie heeft een associatie gevonden tussen een genetische variant gelegen nabij *ROBO2*, een gen betrokken bij axonale groei, en het aantal woorden dat een kind in de leeftijd van 15 tot 18 maanden spreekt (woordproductie). In dit proefschrift, heb ik genoom-wijde genotype data geanalyseerd met als doel het verkrijgen van meer inzicht in de genetische factoren die een rol spelen bij taalontwikkeling in de eerste drie levensjaren en hun relatie met latere taal-, lees- en cognitieve vaardigheden, als wel ontwikkelingsstoornissen zoals ADHD en ASD.

In hoofdstuk 3 heb ik onderzocht of de genetische factoren onderliggend aan woordenschat in de peutertijd (38 maanden) ook een bijdrage leveren aan verschillende taal- en leesvaardigheden in kinderen tussen de 7 en 13 jaar. Hiervoor heb ik gebruik gemaakt van longitudinale data van niet verwante kinderen en structurele vergelijkingsmodellen. Mijn resultaten laten zien dat de meerderheid van genetische invloeden op taalvaardigheden en geletterdheid van midden kindertijd tot vroege adolescentie kan worden toegeschreven aan genetische factoren die al rol spelen bij woordenschat in de vroege kindertijd, met name bij woordbegrip. De phenotypische variatie die door deze vroege genetische factoren verklaard kan worden nam toe gedurende de ontwikkeling, hetgeen zowel genetische stabiliteit als amplificatie suggereert. Meta-regressie analyses voor een spectrum van taalvaardigheden en

geletterdheid toonden aan dat deze amplificatiepatronen het sterkst waren voor variabelen gerelateerd aan leesvaardigheid.

In hoofdstuk 4 heb ik de samenstelling van genetische factoren onderliggend aan woordenschat in de vroege ontwikkeling in detail bestudeerd met behulp van een vergelijkbare onderzoeksaanpak als in hoofdstuk 3. Het genetische landschap onderliggend aan woordproductie en woordbegrip in de leeftijd van 15 tot 38 maanden was dynamisch, met bewijs voor het bestaan van zowel leeftijd- als dimensie-specifieke genetische factoren. Twee van deze genetische factoren waren ook gerelateerd aan lees- en cognitieve vaardigheden in midden kindertijd (7-8 jaar). Vooral genetische invloeden die geïdentificeerd zijn voor woordbegrip bij een leeftijd van 38 maanden droegen bij aan leesvaardigheid en cognitieve capaciteiten halverwege de kindertijd, consistent met mijn bevindingen in hoofdstuk 3. Echter, genetische factoren die met name van invloed zijn op latere verbale processen, zoals leesvaardigheid en verbale intelligentie, waren al detecteerbaar op een leeftijd van 24 maanden, voor woordproductie. Dus, genetische invloeden die een rol spelen in leesvaardigheid en cognitieve capaciteiten halverwege de kindertijd kunnen al vroeg in de ontwikkeling gedetecteerd worden en omvatten minstens twee onafhankelijke genetische factoren.

In hoofdstuk 5 heb ik een meta-GWAS voor woordproductie en woordbegrip (15 tot 38 maanden) uitgevoerd, statistisch gezien de meest krachtige analyse hiervoor tot nu toe. Mijn analyses bevestigden de associatie van een genetische variant nabij *ROBO2* met woordproductie gerapporteerd door een eerdere studie, maar leidden niet tot de identificatie van nieuwe genetische varianten. Genetische correlatie analyses leverden bewijs voor gedeelde genetische factoren tussen woordproductie in de leeftijd van 24-38 maanden en cognitie-gerelateerde eigenschappen later in het leven, zoals leesvaardigheid, opleidingsniveau en intelligentie. Deze bevindingen zijn in lijn met mijn bevindingen in de hoofdstukken 3 en 4.

In hoofdstuk 6 heb ik het complexe patroon van genetische overlap tussen ADHD en taal- en leesvaardigheden halverwege de kindertijd bestudeerd. Laatstgenoemde zijn krachtige genetische proxies voor taalvaardigheden van jonge kinderen. Gebruikmakend van een onderzoeksaanpak genaamd polygenic scoring, heb ik laten zien dat een verhoogd polygeen risico op ADHD was geassocieerd met mindere taalvaardigheden en geletterdheid, met name leesvaardigheden. Vervolgens heb ik multivariabele regressie toegepast en aangetoond dat de polygene associaties tussen ADHD, taalvaardigheden en geletterdheid grotendeels ook gedeeld waren met opleidingsniveau. Echter, na inachtneming van deze gedeelde genetische effecten was er ook bewijs voor ADHD-specifieke associaties met leesvaardigheid. Deze bevindingen suggereren dat er een genetische relatie bestaat tussen ADHD en leesvaardigheden, die verder gaat dan gedeelde genetische effecten met opleidingsniveau.

In hoofdstuk 7 heb ik de genetische overlap tussen ADHD, ASD en opleidingsniveau ontrafeld, die gelijktijdig resulteert in positieve ASD-gerelateerde en negatieve ADHD-gerelateerde genetische correlaties met opleidingsniveau, ondanks een positieve genetische correlatie tussen ASD en ADHD. Gebruikmakend van summary

statistics en multivariabele regressie, heb ik laten zien dat opleidingsniveau-gerelateerde genetische variatie gedeeld is tussen ADHD en ASD. Verschillende combinaties van dezelfde ASD en ADHD risico verhogende allelen leidden tot zowel positieve ASD-gerelateerde als negatieve ADHD-gerelateerde associatie patronen met opleidingsniveau. Deze bevindingen suggereren een rol voor biologische pleiotropie en/of co-lokalisatie processen (dezelfde GWAS variant codeert voor een verschillende ASD en ADHD risico variant door linkage disequilibrium).

De bevindingen gepresenteerd in dit proefschrift laten zien dat de processen die ten grondslag liggen aan vroege taalontwikkeling genetisch gezien complex zijn. Voor een deel zijn ze ook gerelateerd aan taalvaardigheden en geletterdheid later in het leven, als wel cognitie-gerelateerde capaciteiten. Daarnaast heb ik aangetoond dat de genetische overlap van ontwikkelingsstoornissen met taalvaardigheden en geletterdheid van midden kindertijd tot vroege adolescentie erg complex is en mechanismen als biologische pleiotropie, co-lokalisatie en/of genetische overlap met een derde eigenschap omvat. Het uitbreiden van de phenotypische definitie van vroege taalvaardigheden, bijvoorbeeld met receptief taalgebruik, kan een waardevolle toevoeging zijn voor toekomstig onderzoek naar genetische factoren onderliggend aan vroege taalontwikkeling. Daarnaast is het van belang om rekening te houden met genetische heterogeniteit, een eigenschap die kenmerkend is voor het genetisch landschap onderliggend aan taalontwikkeling in de eerste levensjaren.

Biography

Ellen Verhoef was born in Enschede, the Netherlands, on March 10th, 1993. She obtained a BSc degree in Biology at the Rijksuniversiteit Groningen in 2014, followed by a cum laude MSc degree in Medical Biology at Radboud University Nijmegen in 2016. As part of her master education she performed two research projects, both in the area of genomics research. In her first master project, she aimed to identify and biologically characterise genetic variants related to ADHD using whole exome sequencing data from a multiplex family at the department of Human Genetics of the Radboud UMC. In Ellen's second master project, she studied the use of circulating tumor DNA to monitor disease recurrence in colorectal cancer patients at Medlon, a company in medical diagnostics. In June 2016, Ellen started her PhD project in the population genetics of human communication group led by Dr. Beate St Pourcain, one of the themes embedded in the Language and Genetics department of the Max Planck Institute for Psycholinguistics. Under supervision of Dr. Beate St Pourcain and Prof. Dr. Simon E. Fisher, she studied genetic factors underlying early language development and their relationships with subsequent language and literacy skills, as well as related traits within the context of neurodevelopmental disorders.

List of publications

Peer reviewed journals

Verhoef, E., Shapland, C.Y., Fisher, S.E., Dale, P.S., St Pourcain, B. (2020, accepted). The developmental genetic architecture of vocabulary skills during the first three years of life: Capturing emerging associations with later-life reading and cognition. *PLOS Genetics*. doi: 10.1371/journal.pgen.1009144.

Verhoef, E., Shapland, C. Y., Fisher, S. E., Dale, P., & St Pourcain, B. (2020, in press). The developmental origins of genetic factors influencing language and literacy: Associations with early-childhood vocabulary. *Journal of Child Psychology and Psychiatry*. doi: 10.1111/jcpp.13327

Grasby, K. L., Jahanshad, N., Painter, J. N., Colodro-Conde, L., Bralten, J., Hibar, D. P., Lind, P. A., Pizzagalli, F., Ching, C. R. K., McMahon, M. A. B., Shatokhina, N., Zsembik, L. C. P., Thomopoulos, S. I., Zhu, A. H., Strike, L. T., Agartz, I., Alhusaini, S., Almeida, M. A. A., Alnæs, D., Amlien, I. K., Andersson, M., Ard, T., Armstrong, N. J., Ashley-Koch, A., Atkins, J. R., Bernard, M., Brouwer, R. M., Buimer, E. E. L., Bülow, R., Bürger, C., Cannon, D. M., Chakravarty, M., Chen, Q., Cheung, J. W., Couvy-Duchesne, B., Dale, A. M., Dalvie, S., De Araujo, T. K., De Zubicaray, G. I., De Zwarte, S. M. C., Den Braber, A., Doan, N. T., Dohm, K., Ehrlich, S., Engelbrecht, H.-R., Erk, S., Fan, C. C., Fedko, I. O., Foley, S. F., Ford, J. M., Fukunaga, M., Garrett, M. E., Ge, T., Giddaluru, S., Goldman, A. L., Green, M. J., Groenewold, N. A., Grotegerd, D., Gurholt, T. P., Gutman, B. A., Hansell, N. K., Harris, M. A., Harrison, M. B., Haswell, C. C., Hauser, M., Herms, S., Heslenfeld, D. J., Ho, N. F., Hoehn, D., Hoffmann, P., Holleran, L., Hoogman, M., Hottenga, J.-J., Ikeda, M., Janowitz, D., Jansen, I. E., Jia, T., Jockwitz, C., Kanai, R., Karama, S., Kasperaviciute, D., Kaufmann, T., Kelly, S., Kikuchi, M., Klein, M., Knapp, M., Knodt, A. R., Krämer, B., Lam, M., Lancaster, T. M., Lee, P. H., Lett, T. A., Lewis, L. B., Lopes-Cendes, I., Luciano, M., Macciardi, F., Marquand, A. F., Mathias, S. R., Melzer, T. R., Milaneschi, Y., Mirza-Schreiber, N., Moreira, J. C. V., Mühleisen, T. W., Müller-Myhsok, B., Najt, P., Nakahara, S., Nho, K., Olde Loohuis, L. M., Orfanos, D. P., Pearson, J. F., Pitcher, T. L., Pütz, B., Quidé, Y., Ragothaman, A., Rashid, F. M., Reay, W. R., Redlich, R., Reinbold, C. S., Repple, J., Richard, G., Riedel, B. C., Risacher, S. L., Rocha, C. S., Mota, N. R., Salminen, L., Saremi, A., Saykin, A. J., Schlag, F., Schmaal, L., Schofield, P. R., Secolin, R., Shapland, C. Y., Shen, L., Shin, J., Shumskaya, E., Sønderby, I. E., Sprooten, E., Tansey, K. E., Teumer, A., Thalamuthu, A., Tordesillas-Gutiérrez, D., Turner, J. A., Uhlmann, A., Vallerga, C. L., Van der Meer, D., Van Donkelaar, M. M. J., Van Eijk, L., Van Erp, T. G. M., Van Haren, N. E. M., Van Rooij, D., Van Tol, M.-J., Veldink, J. H., **Verhoef, E.**, Walton, E., Wang, M., Wang, Y., Wardlaw, J. M., Wen, W.,

Westlye, L. T., Whelan, C. D., Witt, S. H., Wittfeld, K., Wolf, C., Wolfers, T., Wu, J. Q., Yasuda, C. L., Zaremba, D., Zhang, Z., Zwiers, M. P., Artiges, E., Assareh, A. A., Ayesa-Arriola, R., Belger, A., Brandt, C. L., Brown, G. G., Cichon, S., Curran, J. E., Davies, G. E., Degenhardt, F., Dennis, M. F., Dietsche, B., Djurovic, S., Doherty, C. P., Espiritu, R., Garijo, D., Gil, Y., Gowland, P. A., Green, R. C., Häusler, A. N., Heindel, W., Ho, B.-C., Hoffmann, W. U., Holsboer, F., Homuth, G., Hosten, N., Jack Jr., C. R., Jang, M., Jansen, A., Kimbrel, N. A., Kolskår, K., Koops, S., Krug, A., Lim, K. O., Luykx, J. J., Mathalon, D. H., Mather, K. A., Mattay, V. S., Matthews, S., Mayoral Van Son, J., McEwen, S. C., Melle, I., Morris, D. W., Mueller, B. A., Nauck, M., Nordvik, J. E., Nöthen, M. M., O'Leary, D. S., Opel, N., Paillère Martinot, M.-L., Pike, G. B., Preda, A., Quinlan, E. B., Rasser, P. E., Ratnakar, V., Reppermund, S., Steen, V. M., Tooney, P. A., Torres, F. R., Veltman, D. J., Voyvodic, J. T., Whelan, R., White, T., Yamamori, H., Adams, H. H. H., Bis, J. C., Debette, S., Decarli, C., Fornage, M., Gudnason, V., Hofer, E., Ikram, M. A., Launer, L., Longstreth, W. T., Lopez, O. L., Mazoyer, B., Mosley, T. H., Roshchupkin, G. V., Satizabal, C. L., Schmidt, R., Seshadri, S., Yang, Q., Alzheimer's Disease Neuroimaging Initiative, CHARGE Consortium, EPIGEN Consortium, IMAGEN Consortium, SYS Consortium, Parkinson's Progression Markers Initiative, Alvim, M. K. M., Ames, D., Anderson, T. J., Andreassen, O. A., Arias-Vasquez, A., Bastin, M. E., Baune, B. T., Beckham, J. C., Blangero, J., Boomsma, D. I., Brodaty, H., Brunner, H. G., Buckner, R. L., Buitelaar, J. K., Bustillo, J. R., Cahn, W., Cairns, M. J., Calhoun, V., Carr, V. J., Caseras, X., Caspers, S., Cavalleri, G. L., Cendes, F., Corvin, A., Crespo-Facorro, B., Dalrymple-Alford, J. C., Dannlowski, U., De Geus, E. J. C., Deary, I. J., Delanty, N., Depondt, C., Desrivières, S., Donohoe, G., Espeseth, T., Fernández, G., Fisher, S. E., Flor, H., Forstner, A. J., Francks, C., Franke, B., Glahn, D. C., Gollub, R. L., Grabe, H. J., Gruber, O., Håberg, A. K., Hariri, A. R., Hartman, C. A., Hashimoto, R., Heinz, A., Henskens, F. A., Hillegers, M. H. J., Hoekstra, P. J., Holmes, A. J., Hong, L. E., Hopkins, W. D., Hulshoff Pol, H. E., Jernigan, T. L., Jönsson, E. G., Kahn, R. S., Kennedy, M. A., Kircher, T. T. J., Kochunov, P., Kwok, J. B. J., Le Hellard, S., Loughland, C. M., Martin, N. G., Martinot, J.-L., McDonald, C., McMahon, K. L., Meyer-Lindenberg, A., Michie, P. T., Morey, R. A., Mowry, B., Nyberg, L., Oosterlaan, J., Ophoff, R. A., Pantelis, C., Paus, T., Pausova, Z., Penninx, B. W. J. H., Polderman, T. J. C., Posthuma, D., Rietschel, M., Roffman, J. L., Rowland, L. M., Sachdev, P. S., Sämann, P. G., Schall, U., Schumann, G., Scott, R. J., Sim, K., Sisodiya, S. M., Smoller, J. W., Sommer, I. E., St Pourcain, B., Stein, D. J., Toga, A. W., Trollor, J. N., Van der Wee, N. J. A., van 't Ent, D., Völzke, H., Walter, H., Weber, B., Weinberger, D. R., Wright, M. J., Zhou, J., Stein, J. L., Thompson, P. M., & Medland, S. E. (2020). The genetic architecture of the human cerebral cortex. *Science*, 367(6484). doi: 10.1126/science.aay6690.

Verhoef, E.[‡], Demontis, D.[‡], Burgess, S., Shapland, C. Y., Dale, P. S., Okbay, A., Neale, B. M., Faraone, S. V., iPSYCH-Broad-PGC ADHD Consortium, Stergiakouli, E., Davey Smith, G., Fisher, S. E., Børglum, A., & St Pourcain, B. (2019). Disentangling polygenic associations between Attention-Deficit/Hyperactivity Disorder, educational attainment, literacy and language. *Translational Psychiatry*, 9: 35. doi: 10.1038/s41398-018-0324-2.
[‡] E. Verhoef and D. Demontis share first authorship.

Klein, M., Van Donkelaar, M., **Verhoef, E.**, & Franke, B. (2017). Imaging genetics in neurodevelopmental psychopathology. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 174(5), 485-537. doi:10.1002/ajmg.b.32542.

Preprints under review

Shapland, C.Y., **Verhoef, E.**, Davey Smith, G., Fisher, S.E., Verhulst, B., Dale, P.S., St Pourcain, B. (2020). The multivariate genome-wide architecture of interrelated literacy, language and working memory skills reveals distinct etiologies. *bioRxiv*. doi: doi.org/10.1101/2020.08.14.251199.

Verhoef, E., Grove, J., Shapland, C.Y., Demontis, D., Burgess, S., Rai, D., Børglum, A.D., St Pourcain, B. (2019). Shared polygenetic variation between ASD and ADHD exerts opposite association patterns with educational attainment. *bioRxiv*. doi: doi.org/10.1101/580365.

Research data management

This thesis research has been carried out under the institute research data management policy of the Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands. A summary of this policy can be found at https://www.mpi.nl/page/research-data-management.

As part of an ongoing collaboration within the Early Genetics and Lifecourse Epidemiology (EAGLE) Consortium genome-wide association summary statistics were created for early-life vocabulary and later-life reading skills as part of chapter five. These will be made publically available via the MPI Language Archive (https://archive.mpi.nl/tla/) together with the corresponding manuscript, which is currently in preparation. All analyses presented in this thesis were performed using freely accessible software.

Acknowledgements

I would not have been able to write this PhD thesis without the tremendous support of many and I would like to thank all that contributed in any way to help me achieve this.

This journey all started with Beate and Simon, who had enough trust in me to give me the opportunity to pursue a PhD. Thank you for all your support, inspiration and guidance during the past years. From the first day, I felt very welcome and I was excited about all that had to come. Now, almost five years later, I still feel the same and look forward to spend another couple of years doing genetics research at the MPI.

I would also like to thank the members of my manuscript committee, Prof. dr. Jan Buitelaar, Prof. dr. Bertram Müller-Myhsok and Dr. Emma Hayiou-Thomas, for their time and willingness to review my PhD thesis.

A special thanks goes out to all the collaborators who have contributed to the research presented in this thesis, especially Philip, Ditte, Jakob, Anders, Steven, Wolfgang and Chin Yang. Thank you for your very useful insights regarding language development, genetics of neurodevelopmental disorders and statistical models. In addition, I would like to thank everybody that is involved in ALSPAC, 1958BC, PNC, PGC, iPSYCH, SSGAC and EAGLE. The work presented in this thesis would not have been possible without these efforts to collect large-scale genetic and behavioral data. A huge thanks also goes to all other consortia and researchers that made their data available for analyses.

Thank you also to all colleagues from Language & Genetics! I had the privilage to enjoy a very pleasant working environment, which is created by you all. Apart from the critical discussions and useful insights regarding research topics, I would also like to thank you for putting up decorations in my office for my birthday and all the non-research conversations during lunch. These things really mean much to me. Margot, ik wil je bedanken voor alle gesprekken, je openheid en de leuke tijd die we in Boulder hebben gehad. Fenja, het is super gezellig om met jou een kantoor te delen en enorm bedankt voor alle taken die je van me hebt overgenomen terwijl ik mijn thesis aan het afronden was. Else, we zijn op dezelfde dag begonnen bij het MPI en ik heb veel aan je gehad de afgelopen jaren. Ik wil je bedanken voor je adviezen over hoe een PhD aan te pakken, de fijne, nuchtere gesprekken over onze carrières, en bovenal alle gezelligheid. Chin Yang, Marjolein, Lucía and Mariska, thanks for being such great colleagues, it is a pleasure to work in the same research group as you. Martina, thank you for everything. You have been a tremendous support throughout and your kindness and selflessness are inspirational. To all other MPI'ers that helped me in the past few years, thank you! The IMPRS, Operations, Library and TG teams are truely fantastic and I always felt supported.

Naast alle steun die ik heb ontvangen in de scientific community, hebben ook mijn familie en vrienden een belangrijke rol gespeeld in de afgelopen jaren. Bedankt aan alle ooms, tantes, oma's en opa die elke keer weer de moeite hebben genomen om te begrijpen waar ik nu precies mee bezig was. Anne, Iris, Kelly, Maarten, Tim, Pascal, Patrick, Unique en Heleen, bedankt voor alle gezelligheid de afgelopen jaren! Genoeg ontspanning is van groot belang en daarin mag jullie bijdrage niet onderschat worden. Anne, Iris en Kelly ook enorm bedankt voor jullie hulp in de afrondende fase, zij het nu lay-out tips of het last-minute doorlezen van mijn Nederlandse Samenvatting. Lieve schoonfamilie, bedankt dat jullie er altijd zijn en overal mee willen helpen. Ik ben trots dat ik me ook een beetje Noordink mag noemen. Papa en mama, bedankt dat jullie me altijd aanmoedigen om mijn dromen na te jagen, niet op te geven en de lat hoog te leggen. Jullie zullen hemel en aarde bewegen om me hierbij te helpen en zijn altijd een rots in de branding geweest. Lianne en Marleen, bedankt dat jullie zulke lieve zussen zijn! Tom, je bent een super leuke zwager. Ik weet dat ik altijd op jullie kan rekenen en dat waardeer ik enorm. Tot slot wil ik mijn gezin bedanken. Luca, dankzij jou heb ik in de afrondende fase ook heel wat nachtelijke uurtjes over mijn proefschrift kunnen nadenken. Maar vooral bedankt dat je me er met je fantastische lach en ondeugende blik elke dag weer aan herinnert wat er echt belangrijk is in het leven. Ramon, jou ben ik een proefschrift opzich aan bedankjes schuldig. Bedankt voor alle keren dat je thuis een stapje harder deed, zodat ik tijd aan mijn onderzoek kon besteden. Maar bovenal, bedankt dat je er altijd voor me bent, in me gelooft en me door de laatste loodjes hebt heen gesleept.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*

26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*

53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. *Split intransitivity in Rotokas, a Papuan language of Bougainville. Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space**.** *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*

77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*
85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*
90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. Incrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*

101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*
105. Modelling Multimodal Language Processing. *Alastair Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*
107. Situational variation in non-native communication. *Huib Kouwenhoven*
108. Sustained attention in language production. *Suzanne Jongman*
109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*
110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*
111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*
112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*
113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*
114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*
115. Music and language comprehension in the brain. *Richard Kunert*
116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H.L. Lam*
117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*
118. Language processing in a conversation context. *Lotte Schoot*
119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*
120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*
121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*
122. Sensorimotor experience in speech perception. *Will Schuerman*
123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*
124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*
125. Lapse organization in interaction. *Elliott Hoey*
126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*

127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*
128. The role of voice and word order in incremental sentence processing. *Sebastian Saupp*
129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*
130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*
131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stoehr*
132. A holistic approach to understanding pre-history. *Vishnupriya Kolipakam*
133. Characterization of transcription factors in monogenic disorders of speech and language. *Sara Busquets Estruch*
134. Indirect request comprehension in different contexts. *Johanne Tromp*
135. Envisioning Language - An Exploration of Perceptual Processes in Language Comprehension. *Markus Ostarek*
136. Listening for the WHAT and the HOW: Older adults' processing of semantic and affective information in speech. *Juliane Kirsch*
137. Let the agents do the talking: on the influence of vocal tract anatomy on speech during ontogeny and glossogeny. *Rick Janssen*
138. Age and hearing loss effects on speech processing. *Xaver Koch*
139. Vocabulary knowledge and learning: Individual differences in adult native speakers. *Nina Mainz*
140. The face in face-to-face communication: Signals of understanding and non-understanding. *Paul Hömke*
141. Person reference and interaction in Umpila/Kuuku Ya'u narrative. *Clair Hill*
142. Beyond the language given: The neurobiological infrastructure for pragmatic inferencing. *Jana Bašnáková*
143. From Kawapanan to Shawi: Topics in language variation and change. *Luis Miguel Rojas-Berscia*
144. On the oscillatory dynamics underlying speech-gesture integration in clear and adverse listening conditions. *Linda Drijvers*
145. Understanding temporal overlap between production and comprehension. *Amie Fairs*
146. The role of exemplars in speech comprehension. *Annika Nijveld*
147. A network of interacting proteins disrupted in language-related disorders. *Elliot Sollis*
148. Fast speech can sound slow: Effects of contextual speech rate on word recognition. *Merel Maslowski*
149. Reason-giving in everyday activities. *Julija Baranova*

150. Speech planning in dialogue - Psycholinguistic studies of the timing of turn taking. *Mathias Barthel*
151. The role of neural feedback in language unification: How awareness affects combinatorial processing. *Valeria Mongelli*
152. Exploring social biases in language processing. *Sara Iacozza*
153. Vocal learning in the pale spear-nosed bat, *Phyllostomus discolor*. *Ella Lattenkamp*
154. The effect of language contact on speech and gesture: The case of Turkish-Dutch bilinguals in the Netherlands. *Elif Zeynep Azar*
155. Language and society: How social pressures shape grammatical structure. *Limor Raviv*
156. The moment in between: Planning speech while listening. *Svetlana-Lito Gerakaki*
157. How speaking fast is like running: Modelling control of speaking rate. *Joe Rodd*
158. The power of context: How linguistic contextual information shapes brain dynamics during sentence processing. *René Terporten*
159. Neurobiological models of sentence processing. *Marvin Uhlmann*
160. Individual differences in syntactic knowledge and processing: The role of literacy experience. *Saoradh Favier*
161. Memory for speaking and listening. *Eirini Zormpa*
162. Masculine generic pronouns: Investigating the processing of an unintended gender cue. *Theresa Redl*
163. Properties, structures and operations: Studies on language processing in the brain using computational linguistics and naturalistic stimuli. *Alessandro Lopopolo*
164. Investigating spoken language comprehension as perceptual inference. *Greta Kaufeld*
165. What was that Spanish word again? Investigations into the cognitive mechanisms underlying foreign language attrition. *Anne Mickan*
166. A tale of two modalities: How modality shapes language production and visual attention. *Francie Manhardt*
167. Why do we change how we speak? Multivariate genetic analyses of language and related traits across development and disorder. *Ellen Verhoef*

MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS

VISITING ADDRESS
Wundtlaan 1
6525 XD Nijmegen
The Netherlands

POSTAL ADDRESS
P.O. Box 310
6500 AH Nijmegen
The Netherlands

CONTACT
T +31(0)24 3521 911
F +31(0)24 3521 213
E info@mpi.nl
Twitter @MPI_NL
www.mpi.nl