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

Ellen Verhoef,[1,2] Chin Yang Shapland,[1,3,4] Simon E. Fisher,[1,5] Philip S. Dale,[6] and Beate St Pourcain[1,3,5]

[1]Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [2]International Max Planck Research School for Language Sciences, Nijmegen, The Netherlands; [3]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK; [4]Population Health Sciences, University of Bristol, Bristol, UK; [5]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands; [6]Speech and Hearing Sciences, University of New Mexico, Albuquerque, NM,USA

Background: The heritability of language and literacy skills increases from early-childhood to adolescence. The underlying mechanisms are little understood and may involve (a) the amplification of genetic influences contributing to early language abilities, and/or (b) 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 nonverbal 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 ($\leq 99\%$). 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 nonverbal 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.

Introduction

Individual differences in vocabulary during the preschool period are predictive of many later language- and literacy-related skills (Bleses, Makransky, Dale, Højen, & Ari, 2016; Bornstein, Hahn, Putnick, & Suwalsky, 2014; Duff, Reen, Plunkett, & Nation, 2015; Lee, 2011), an important component of academic achievement (Durham, Farkas, Hammer, Bruce Tomblin, & Catts, 2007). 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 (Duff et al., 2015). 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 (Lee, 2011).

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 (Gough & Tunmer, 1986). Early vocabulary is a central component of both these abilities (Jong & Leij, 2002). 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 (Jong & Leij, 2002). Listening comprehension (i.e. the understanding of spoken language), particularly bottom-up processing, necessarily begins with vocabulary comprehension (Dickinson & Neuman, 2007). Spelling performance is also closely related to phonological awareness and other

phonological abilities (Dich & Cohn, 2013). 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 (Harlaar, Hayiou-Thomas, Dale, & Plomin, 2008; Hayiou-Thomas, Dale, & Plomin, 2012; St Pourcain et al., 2014; Verhoef et al., 2019). Specifically, longitudinal twin studies, assessing heritability indirectly based on differences in phenotypic correlations between monozygotic and dizygotic twins (twin-h^2; Plomin, DeFries, Knopik, & Neiderhiser, 2016), have reported heritability estimates of 22% to 28% for a combined language measure including expressive vocabulary at 2, 3 and 4 years of age (Hayiou-Thomas et al., 2012). 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 (St Pourcain et al., 2014). 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% (Harlaar et al., 2008; Hayiou-Thomas et al., 2012) and SNP-h^2 estimates from 32% to 54% (Verhoef et al., 2019). 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 (Harlaar et al., 2008; Hayiou-Thomas et al., 2012).

The increase in heritability from early-childhood to adolescence has been reported for many cognitive skills (Briley & Tucker-Drob, 2013; Haworth et al., 2010), suggesting overarching aetiological mechanisms that may involve processes of genetic innovation and amplification (Plomin & DeFries, 1985). 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 (Briley & Tucker-Drob, 2013). 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 (Briley & Tucker-Drob, 2013). From eight years of age onwards (mid-childhood), there was evidence for enhanced genetic stability with dominance of amplification processes (Briley & Tucker-Drob, 2013). 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 (Hayiou-Thomas et al., 2012) and reading abilities (Harlaar et al., 2008), 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 (Harlaar et al., 2008; Hayiou-Thomas et al., 2012), 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, $N \leq 6,092$). 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, nonword repetition and verbal intelligence, as well as nonverbal intelligence (7–13 years).

Methods
Participants

All participants were drawn from ALSPAC, a UK population-based longitudinal pregnancy-ascertained birth cohort (estimated birth date: 1991–1992, Appendix S1; Boyd et al., 2013; Fraser et al., 2013). 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; Purcell et al., 2007; Appendix S2). After quality control, 465,740 SNPs and $\leq 6,092$ individuals with high-quality genetic and phenotypic data remained.

Measures

Early-childhood vocabulary.   Expressive and receptive vocabulary was assessed at 38 months using parent reports and age-specific defined word lists adapted from the

MacArthur Communicative Development Inventory Words & Sentences (CDI; Fenson et al., 1993). Parents were asked whether their child was able to (a) say, (b) understand, or (c) 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 (a) says and (c) 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 (b) understands and (c) says and understands. CDI expressive vocabulary scores have high validity, showing correlations with direct assessments of over 0.70 (Dale, 1991; Ring & Fenson, 2000). The correlation between parental and direct assessment of receptive vocabulary is 0.55 (Ring & Fenson, 2000). 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, nonword repetition and verbal intelligence were assessed from mid-childhood to early adolescence (7–13 years, $N \leq 5,749$) using both standardised and ALSPAC-specific instruments (Table 1, Appendix S3). Word reading accuracy and comprehension (age 7 years) were measured using the basic reading subtest of the Wechsler Objective Reading Dimensions (WORD) assessment. Word and nonword 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 nonword reading speed (age 13 years) was 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, and nonword 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.

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; Price et al., 2006). In addition, vocabulary scores were residualised for age squared, as vocabulary develops rapidly during early-childhood (Brooks & Meltzoff, 2008).

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 (Yang, Lee, Goddard, & Visscher, 2011). 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; Yang et al., 2011), which is a matrix with as many

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. 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.

columns and rows as individuals. Here, the GRM matrix was created with PLINK (v1.9; Purcell et al., 2007) using individuals with a genetic relationship < 0.05 ($N_{individuals} \leq 6,092$) 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 (Lee, Yang, Goddard, Visscher, & Wray, 2012) 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/covariance 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.stpourca in/gsem; St Pourcain et al., 2017). 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/covariance structures between unrelated individuals (Appendix S4; St Pourcain et al., 2017). 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 (Yang et al., 2011). In addition, we estimated SNP-h^2, genetic correlations, factorial coheritability (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, Figures S2a and S3a). Second, we carried out a meta-analysis of absolute GSEM path coefficients for these 13 models across predefined 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, Figures S4a and S5a) as part of sensitivity analyses. Finally, to compare LRA genetic covariance patterns with nonverbal 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/; Li and Ji, 2005). This corresponds to an experiment-wide significance threshold of .005 (0.05/9).

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, nonword 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 (Verhoef et al., 2019).

Consistent with phenotypic correlations between expressive and receptive vocabulary at 38 months (r_p = .63, Figure 1B), bivariate genetic correlations were strong (r_g = .86(SE = 0.15), p = .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 = .001) to 0.95 (SE = 0.23, p = 1 × 10^{-8}; Figure 1C). In contrast, expressive vocabulary was genetically correlated with VIQ at 8 years only (r_g = 0.38(SE = 0.14), p = .003, Figure 1C).

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 = 4 × 10^{-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 = 7 × 10^{-5}; Figure 2A,B), capturing approximately two thirds of its total genetic variance (factorial coheritability: 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-
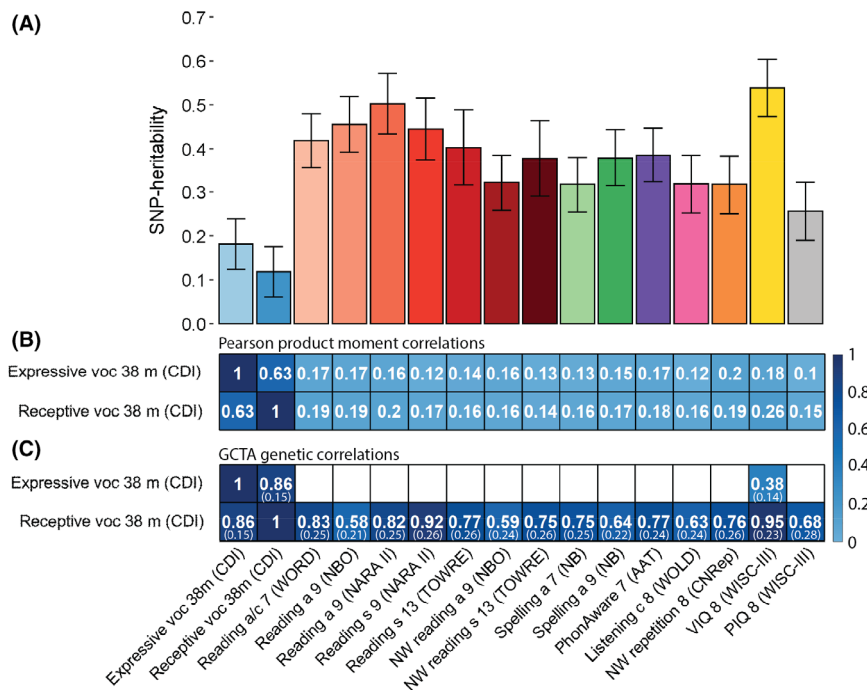
Figure 1 SNP heritability, phenotypic and genetic correlations. (A) SNP heritability was estimated with GCTA software, based on directly genotyped SNPs. Bars represent standard errors. (B and C) Bivariate correlations among rank-transformed measures passing the experiment-wide significance threshold ($p \leq .005$) 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

coefficient $a_{31} = -0.29(SE = 0.10)$, $p = .004$; 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 $a_{22} = -0.21(SE = 0.05)$, $p = 4 \times 10^{-6}$; Figure 2A,B). Thus, about a third of the genetic variance in receptive vocabulary is unique (factorial coheritability = 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, Figures S2 and 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 $a_{32} = -0.67$ ($SE = 0.06$), $p < 1 \times 10^{-10}$; 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 $a_{32} = -0.62(SE = 0.05)$, $p < 1 \times 10^{-10}$; Figure 2C,D), entailing nearly the entire SNP-h^2 of the measure (factorial coheritability: 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 nonword 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 (Figures S2 and 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 nonword reading speed at 13 years and spelling accuracy at 7 years (Figure S2d–i).

To evaluate overarching association patterns across mid-childhood/early-adolescent language and literacy skills (Figures S1 and S6), we meta-analysed absolute path coefficients across all 13 forward GSEM models (Table S1). This meta-analysis confirmed the amplification of genetic influences
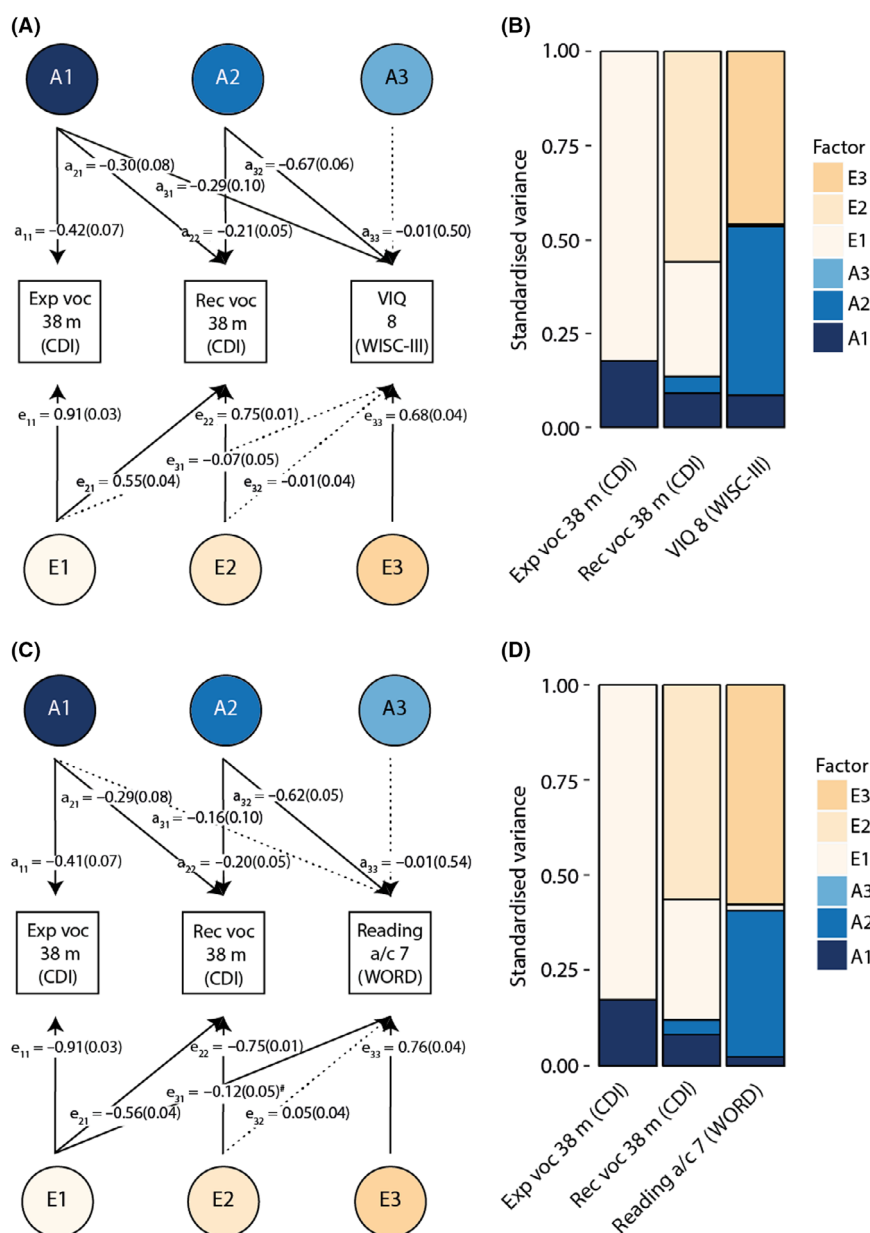
Figure 2 Path models and variance plots for early vocabulary and mid-childhood verbal intelligence or reading accuracy/comprehension. Trivariate Cholesky decomposition models (forward GSEM, GSEM software) were fitted based on all available observations for children across development ($N \leq 6,092$). (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 $p \leq .05$, and dashed lines indicate path coefficients with $p > .05$. (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 ($p \leq .05$), but not the experiment-wide significance threshold ($p \leq .005$). Trait abbreviations are described in Table 1

that are unique to receptive vocabulary at 38 months (meta-path-coefficient $a_{32} = 0.62(0.06)$, $p < 1 \times 10^{-10}$; 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 $a_{31} = 0.20(SE = 0.08)$, $p = .009$; 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 $a_{33} = 0.34(SE = 0.29)$, $p = .24$; 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).
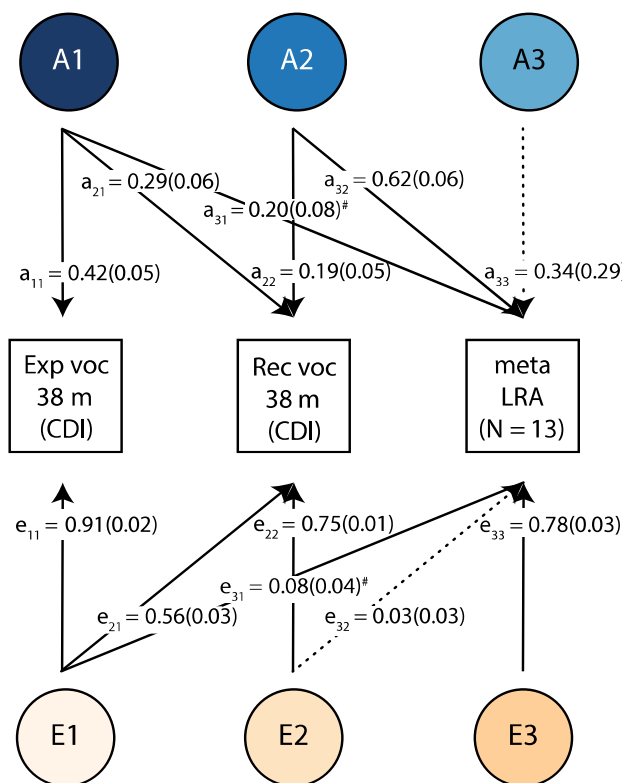
Cholesky decompositions are sensitive to the order of modelled traits, although SNP-h^2 estimations remain unchanged. We therefore created 13 additional GSEM models, as part of sensitivity analyses, reversing the order of expressive and receptive vocabulary at 38 months (reverse GSEM models, with path 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; Figures S4 and 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 (Figures S4 and 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 (Figures S4 and S5).

To identify the most predictive early genetic factors observed using either forward or reverse GSEM models, we studied factorial coheritabilities 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 nonverbal 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), $p < 1 \times 10^{-10}$; Figure S7a,b; factorial coheritability: 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).

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, nonword repetition, verbal and nonverbal 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 nonverbal 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 (Briley & Tucker-Drob, 2013; Hayiou-Thomas et al., 2012).
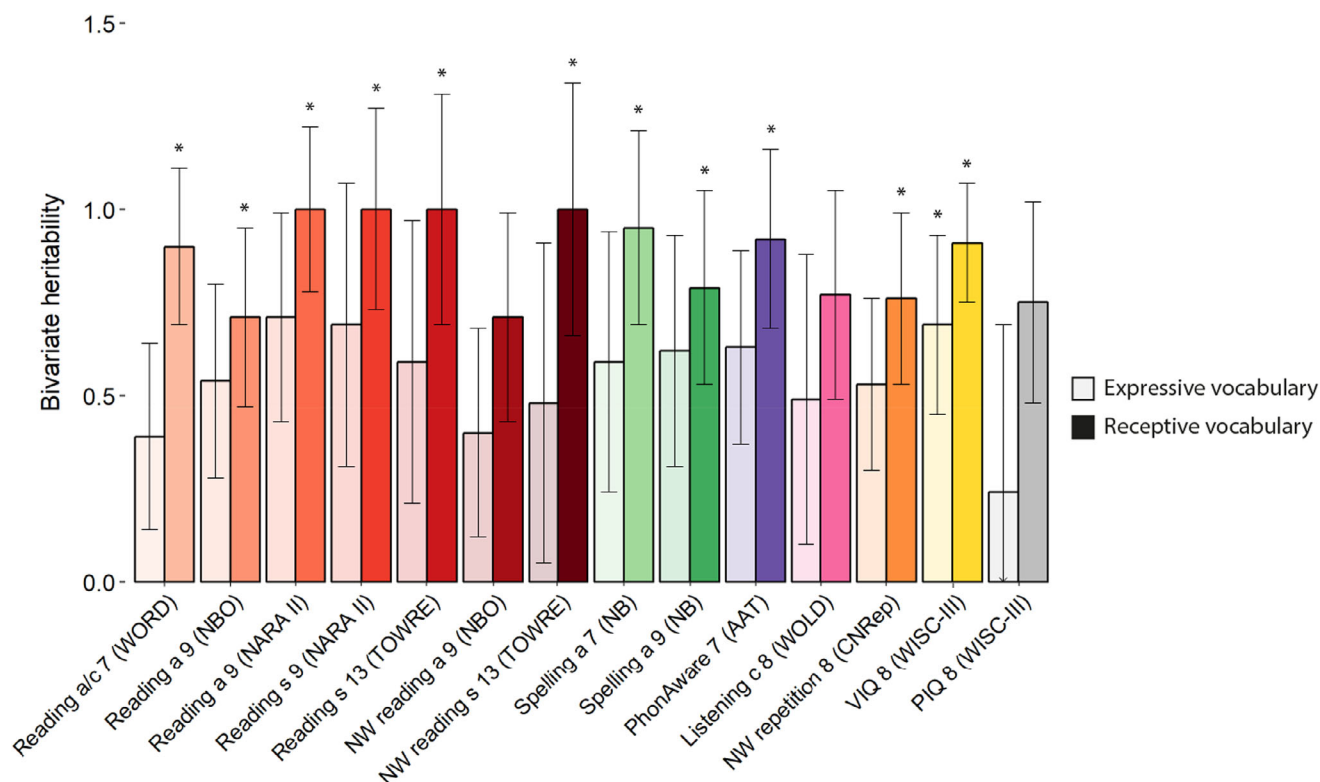
Figure 3 Meta-analyses of developmental structural models. Absolute path coefficients for 13 structural equation models (forward GSEM) corresponding to 13 LRAs in mid-childhood and early adolescence were meta-analysed, accounting for phenotypic interrelatedness. Detailed information, including estimates of effect heterogeneity, is shown in Table S4. # Path coefficient passing the nominal ($p \leq$ .05), but not the experiment-wide significance threshold ($p \leq$ .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 ($p \leq .005$). Trait abbreviations are described in Table 1

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 (Hayiou-Thomas et al., 2012) and reading (Harlaar et al., 2008). However, as genetic factors accounted only for about a third of the phenotypic correlations (Harlaar et al., 2008; Hayiou-Thomas et al., 2012), findings have been interpreted as evidence for genetic innovation (Hayiou-Thomas et al., 2012). In the present study, early vocabulary-related genetic factors, especially those related to receptive vocabulary, explained the majority of genetic variance ($\leq 99\%$ 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 (St Pourcain et al., 2017), 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 coheritability estimates, is consistent with overarching developmental patterns. According to the 'generalist genes' hypothesis, cognitive abilities are presumed to share genetic variance components (Plomin & Kovas, 2005). 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' (Plomin & Kovas, 2005), 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 (Frost, Madsbjerg, Niedersøe, Olofsson, & Sørensen, 2005). Notably, broadly defined early oral language, including receptive skills (Bryant, Maclean, & Bradley, 1990), has been shown to affect word recognition (NICHD Early Child Care Research Network, 2005), while vocabulary comprehension is also a precursor of listening comprehension (Dickinson & Neuman, 2007). 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' (Gough & Tunmer, 1986). 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 (Psyridou, Eklund, Poikkeus, & Torppa, 2018), and expressive and receptive vocabulary may be independently related to prereading skills (Wise, Sevcik, Morris, Lovett, & Wolf, 2007). Furthermore, variation in comprehension has been associated with nonlinguistic cognitive measures, such as tool use and symbolic play, compared to expressive vocabulary (Bates, Dale, & Thal, 2019). 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 (Plomin, DeFries, & Loehlin, 1977). Furthermore, the environmental variance may decrease with the start of schooling (Samuelsson et al., 2008). 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 (Walters, Churchhouse, & Hosking, 2019). 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 (Lee et al., 2018) 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 (Ring & Fenson, 2000; Sachse & Von Suchodoletz, 2008). A further limitation of the current study is that the CDI Words & Sentences were developed for vocabulary assessment in children up to 30 months (Fenson et al., 1993), 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 (Frank, Braginsky, Yurovsky, & Marchman, 2017), might be useful instruments to capture genetic variation in language, literacy and cognitive skills many years later in life.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Appendix S1. ALSPAC description.
Appendix S2. Genetic quality control.
Appendix S3. Mid-childhood/early-adolescent ALSPAC measures.
Appendix S4. Structural equation modelling.
Appendix S5. Factorial co-heritability.
Appendix S6. Bivariate heritability.
Appendix S7. Meta-analysis across mid-childhood/early-adolescent language- and literacy-related abilities.
Appendix S8. Power analyses.
Table S1. Meta-analysis domains of mid-childhood/early-adolescent language- and literacy-related abilities.
Table S2. SNP-heritability estimates.

Acknowledgements

Correspondence

Beate St Pourcain, Language and Genetics Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands; Email: Beate.StPourcain@mpi.nl

---

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 nonverbal 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.

References

Bates, E., Dale, P.S., & Thal, D. (2019). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 95–151). Hoboken, NJ: John Wiley & Sons Ltd.

Bleses, D., Makransky, G., Dale, P.S., Højen, A., & Ari, B.A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, 37, 1461–1476.

Bornstein, M.H., Hahn, C.-S., Putnick, D.L., & Suwalsky, J.T.D. (2014). Stability of core language skill from early childhood to adolescence: A latent variable approach. *Child Development*, 85(4), 1346–1356.

Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., Molloy, L. et al (2013). Cohort Profile: The 'children of the 90s'–the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, 42, 111–127.

Briley, D.A., & Tucker-Drob, E.M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science*, 24, 1704–1713.

Brooks, R., & Meltzoff, A.N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, 35, 207–220.

Bryant, P., Maclean, M., & Bradley, L. (1990). Rhyme, language, and children's reading. *Applied Psycholinguistics*, 11, 237–252.

Dale, P.S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research*, 34, 565–571.

de Jong, P.F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading*, 6, 51–77.

Dich, N., & Cohn, A.C. (2013). A review of spelling acquisition: Spelling development as a source of evidence for the psychological reality of the phoneme. *Lingua*, 133, 213–229.

Dickinson, D.K., & Neuman, S.B. (2007). *Handbook of early literacy research*, New York, NY: Guilford Press.

Duff, F.J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 56, 848–856.

Durham, R.E., Farkas, G., Hammer, C.S., Bruce Tomblin, J., & Catts, H.W. (2007). Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility*, 25, 294–305.

Fenson, L., Dale, P., Reznick, J.S., Thal, D., Bates, E., Hartung, J., Pethick, S. et al (1993). *User's guide and technical manual for the MacArthur communicative development inventories*. San Diego, CA: Singular Publishing.

Frank, M.C., Braginsky, M., Yurovsky, D., & Marchman, V.A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 44, 677–694.

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J. et al (2013). Cohort profile: The Avon longitudinal study of parents and children: ALSPAC mothers cohort. *International Journal of Epidemiology*, 42, 97–110.

Frost, J., Madsbjerg, S., Niedersøe, J., Olofsson, Å., & Sørensen, P.M. (2005). Semantic and phonological skills in predicting reading development: From 3–16 years of age. *Dyslexia*, 11, 79–92.

Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6–10.

Harlaar, N., Hayiou-Thomas, M.E., Dale, P.S., & Plomin, R. (2008). Why do preschool language abilities correlate with later reading? A twin study. *Journal of Speech, Language, and Hearing Research*, 51, 688–705.

Haworth, C., Wright, M., Luciano, M., Martin, N., de Geus, E., van Beijsterveldt, C., . . . & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular psychiatry*, 15, 1112–1120.

Hayiou-Thomas, M.E., Dale, P.S., & Plomin, R. (2012). The etiology of variation in language skills changes with development: A longitudinal twin study of language from 2 to 12 years. *Developmental Science*, 15, 233–249.

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, 32, 69–92.

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

Lee, S.H., Yang, J., Goddard, M.E., Visscher, P.M., & Wray, N.R. (2012). Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics (Oxford, England)*, 28, 2540–2542.

Li, J., & Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, 95, 221–227.

NICHD Early Child Care Research Network. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology*, 41, 428–442.

Plomin, R., & DeFries, J.C. (1985). The origins of individual differences in infancy; the Colorado adoption project. *Science*, 230, 1369.

Plomin, R., DeFries, J.C., Knopik, V.S., & Neiderhiser, J.M. (2016). *Behavioral genetics* (6th edn). New York: Worth Publishers.

Plomin, R., DeFries, J.C., & Loehlin, J.C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, 84, 309–322.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin*, 131, 592–617.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38, 904–909.

Psyridou, M., Eklund, K., Poikkeus, A.-M., & Torppa, M. (2018). Reading outcomes of children with delayed early vocabulary: A follow-up from age 2–16. *Research in Developmental Disabilities*, 78, 114–124.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., . . . & Sham, P.C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575.

Ring, E.D., & Fenson, L. (2000). The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language*, 20, 141–159.

Sachse, S., & Von Suchodoletz, W. (2008). Early identification of language delay by direct language assessment or parent report? *Journal of Developmental and Behavioral Pediatrics*, 29, 34–41.

Samuelsson, S., Byrne, B., Olson, R.K., Hulslander, J., Wadsworth, S., Corley, R., . . . & DeFries, J.C. (2008). Response to early literacy instruction in the United States, Australia, and Scandinavia: A behavioral-genetic analysis. *Learning and Individual Differences*, 18, 289–295.

St Pourcain, B., Cents, R.A., Whitehouse, A.J., Haworth, C.M., Davis, O.S., O'Reilly, P.F., . . . & Davey Smith, G. (2014). Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature Communications*, 5, 4831.

St Pourcain, B., Eaves, L.J., Ring, S.M., Fisher, S.E., Medland, S., Evans, D.M., & Smith, G.D. (2017). Developmental changes within the genetic architecture of social communication behaviour: A multivariate study of genetic variance in unrelated individuals. *Biological Psychiatry*, 83, 598–606.

Verhoef, E., Demontis, D., Burgess, S., Shapland, C.Y., Dale, P.S., Okbay, A., . . . & St Pourcain, B. (2019). Disentangling polygenic associations between attention-deficit/hyperactivity disorder, educational attainment, literacy and language. *Translational Psychiatry*, 9, 35.

Walters, R., Churchhouse, C., & Hosking, R. (2019, August 7). Heritability 101: What is 'heritability'? Available from http://www.nealelab.is/blog/2017/9/13/heritability-101-what-is-heritability. Accessed August 7, 2019.

Wise, J.C., Sevcik, R.A., Morris, R.D., Lovett, M.W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research*, 50, 1093–1109.

Yang, J., Lee, S.H., Goddard, M.E., & Visscher, P.M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, 88, 76–82.