

RESEARCH

Open Access



Genetic and phenotypic heterogeneity in early neurodevelopmental traits in the Norwegian Mother, Father and Child Cohort Study

Laura Hegemann^{1,2,3*}, Elizabeth C. Corfield^{1,2}, Adrian Dahl Askelund^{1,2,3}, Andrea G. Allegrini^{4,5}, Ragna Bugge Askeland¹, Angelica Ronald⁷, Helga Ask^{1,8}, Beate St Pourcain^{9,10,11}, Ole A. Andreassen^{6,12}, Laurie J. Hannigan^{1,2,10} and Alexandra Havdahl^{1,2,8}

Abstract

Background Autism and different neurodevelopmental conditions frequently co-occur, as do their symptoms at sub-diagnostic threshold levels. Overlapping traits and shared genetic liability are potential explanations.

Methods In the population-based Norwegian Mother, Father, and Child Cohort study (MoBa), we leverage item-level data to explore the phenotypic factor structure and genetic architecture underlying neurodevelopmental traits at age 3 years (N=41,708–58,630) using maternal reports on 76 items assessing children's motor and language development, social functioning, communication, attention, activity regulation, and flexibility of behaviors and interests.

Results We identified 11 latent factors at the phenotypic level. These factors showed associations with diagnoses of autism and other neurodevelopmental conditions. Most shared genetic liabilities with autism, ADHD, and/or schizophrenia. Item-level GWAS revealed trait-specific genetic correlations with autism (items r_g range = -0.27–0.78), ADHD (items r_g range = -0.40–1), and schizophrenia (items r_g range = -0.24–0.34). We find little evidence of common genetic liability across all neurodevelopmental traits but more so for several genetic factors across more specific areas of neurodevelopment, particularly social and communication traits. Some of these factors, such as one capturing prosocial behavior, overlap with factors found in the phenotypic analyses. Other areas, such as motor development, seemed to have more heterogenous etiology, with specific traits showing a less consistent pattern of genetic correlations with each other.

Conclusions These exploratory findings emphasize the etiological complexity of neurodevelopmental traits at this early age. In particular, diverse associations with neurodevelopmental conditions and genetic heterogeneity could inform follow-up work to identify shared and differentiating factors in the early manifestations of neurodevelopmental traits and their relation to autism and other neurodevelopmental conditions. This in turn could have implications for clinical screening tools and programs.

*Correspondence:

Laura Hegemann
alha@fhi.no

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Recent versions of international diagnostic classification systems have introduced an umbrella category of neurodevelopmental conditions. Conditions classified in this category typically manifest from childhood and are characterized by divergent trajectories of development. Generally, they are diagnosed based on significant difficulties in developmental skills in areas such as language, social abilities, learning, or motor activity. Neurodevelopmental conditions include autism spectrum conditions (autism) as well as conditions such as attention-deficit hyperactivity disorder (ADHD), intellectual disabilities, specific learning disabilities, developmental coordination disorder, and tic conditions. Some of these conditions had previously been conceptualized as independent and mutually exclusive conditions. For example, under DSM-IV, autism was an exclusion criterion for ADHD preventing their co-diagnosis. However, the recent shift towards their co-classification aligns diagnostic systems, such as the DSM, with a longstanding clinical awareness that observations of the specific traits and clinical features of neurodevelopmental conditions co-occur across diagnostic boundaries.

Neurodevelopmental conditions frequently co-occur [1, 2] and share symptoms at sub-diagnostic threshold levels [3, 4]. While the etiology of this co-occurrence is not well understood, some observations have implicated shared genetic liability between neurodevelopmental conditions. Unidentified latent genetic factors [5] as well as identified common [6–9] and rare genetic variants [10–12] are shared amongst many clinically-distinct neurodevelopmental conditions. Revisions of the formal diagnostic classification of neurodevelopmental conditions, such as those mentioned above, are in part a reflection of developments in our understanding of their common features [13–15]. However, it is also important to recognize that initial classifications were intended to describe characteristic symptom profiles rather than intended to imply inherent distinctions reflecting biological realities. That is, investigating the co-occurrence and shared etiology of different neurodevelopmental conditions can result in highly clinically-relevant insights without necessarily calling into question the distinctiveness and clinical utility of the conditions as separate entities.

Investigating the nosological and genetic bases for co-occurring neurodevelopmental conditions requires detailed data on their traits. Population-based registries, which collect diagnostic information from health care use for a given population, are typically limited to diagnostic (yes/no) outcomes. Clinical cohorts, which may have more detailed data, are generally smaller and commonly ascertain individuals based on a single condition. Thus, meaningful analyses of common genetic variants and

shared etiology across areas of development and specific traits are difficult. Data collected in population-based cohorts, which are sampled from the general population and focused on longitudinal collection of data, typically have more breadth and depth of information that can help explore shared etiology of neurodevelopmental traits, but relatively fewer individuals with neurodevelopmental conditions. Still, relevant traits—capturing individual differences in language and motor development, attention, hyperactivity, social behavior, and repetitive, restricted behaviors and interests—can be observed in all children. These traits are likely influenced by some of the same underlying genetic liabilities as neurodevelopmental conditions [16–18]. The prospective nature of population-based birth cohorts means these traits can be studied early—prior to or around the age at which neurodevelopmental diagnoses are most commonly made [19, 20]. Exploring the relationships between neurodevelopmental traits early in life, investigating their genetic liabilities, and exploring links to neurodevelopmental conditions can give new insights into etiological mechanisms underlying the development and differentiation of such conditions.

Previous studies have examined the phenotypic factor structure of behaviors related to multiple neurodevelopmental conditions, primarily using items from questionnaires for both autism and ADHD in school-aged or older children. Out of five studies, four found differentiated dimensions of social communication, restricted and repetitive interests and behaviors, attention, and hyperactivity-impulsivity [21–24], while one found a common dimension for restricted and repetitive interests and behaviors with hyperactivity-impulsivity [25].

Regarding genetic factor structure, studies have found both shared and differentiating genetic factors between different domains of autism as well as with other neurodevelopmental conditions. Evidence from both twin and molecular genetic studies suggests that communication and repetitive interest and behavior traits of autism have genetically dissociable domains [26, 27]. Findings across a range of methodologies support correlated but separate genetic contributions to ADHD and autism [24, 28, 29]. Although the status and history of schizophrenia's conceptualization as having neurodevelopmental origins is complex and warrants a more fulsome discussion [30, 31], it is worth noting that schizophrenia also shares genetic liability with both autism [7, 32, 33] and ADHD [29]. Initial evidence shows that this overlap contributes to different aspects of the phenotypic heterogeneity seen in autism [18, 27, 33]. Finally, across neurodevelopment more broadly, Pettersson et al. [5] found both a shared latent genetic factor across a range of different neurodevelopmental traits as well as specific genetic latent factors

for impulsivity, learning problems, and autism and tics in a general population twin sample. In the present study, we leverage information on multiple traits related to different neurodevelopmental conditions. We investigate the phenotypic factor structure and genetic architecture underlying these early (age 3 years) neurodevelopmental traits in a large population-based birth cohort. We additionally investigate associations of these early signs with neurodevelopmental conditions at both the phenotypic and genotypic levels.

Methods

Measures and sample

Sample

The Norwegian Mother, Father and Child Cohort Study (MoBa) is a population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health [34, 35]. Participants were recruited from all over Norway from 1999 to 2008. The women consented to participation in 41% of the pregnancies. Blood samples were obtained from both parents during pregnancy and from mothers and children (umbilical cord) at birth. The cohort includes approximately 114,500 children, 95,200 mothers and 75,200 fathers. The current study is based on version 12 of the quality-assured data files released for research in January 2019. The establishment of MoBa and initial data collection was based on a license from the Norwegian Data Protection Agency and approval from The Regional Committees for Medical and Health Research Ethics. The MoBa cohort is currently regulated by the Norwegian Health Registry Act. The current study was approved by The Regional Committees for Medical and Health Research Ethics (2016/1702).

The present study was conducted on a subset of the cohort ($n=58,630$) who had information available from the 36-month questionnaire. The children were an average of 3.1 years ($SD=0.18$) old when mothers completed the questionnaire. The sample had a 1.04:1 male-to-female ratio. Genetic analyses were conducted using a further quality controlled genotyped subset of the cohort ($n=42,934$). For more information on the genotyping of the MoBa sample and for the family-based quality control pipeline used to prepare these data for analysis, see Corfield et al. [36].

Measures for neurodevelopmental traits

We included items from all maternal report scales related to neurodevelopment in the 3-year questionnaire that asked about children's observable behavior (as opposed to maternal concerns). Items were selected to cover areas of motor, language, social, communication, attention, activity regulation, sensory perception, and flexibility of behaviors and interests across multiple scales

when possible (Fig. 1). This included items from the Social Communication Questionnaire (SCQ) [37], Ages and Stages Questionnaire (ASQ) [38], Non-Verbal Communication Checklist (NVCC) [39], Modified Checklist for Autism in Toddlers (M-CHAT) [40], Early Screening for Autistic Traits Questionnaire (ESAT) [41], the attention and hyperactivity subscale from the Child Behavior Checklist (CBCL) [42], the prosocial behaviors subscale of the Strength and Difficulties Questionnaire (SDQ) [43] as well as several MoBa-specific questions. All items included had either dichotomous (e.g., yes/no) or trichotomous (e.g., not true/sometimes true/often true) response categories. Items were reverse coded where needed so that higher values reflected greater endorsement of the trait.

Measures for diagnostic and clinically relevant outcomes

Diagnostic data was ascertained from the Norwegian Patient Register (NPR) between 2008 and June 2021 based on ICD-10 criteria using the R package *phenotools* [44]. Therefore, for the youngest in the cohort, diagnostic data was available from birth until approximately age 12, and for the oldest from approximately age 8 to 21. For those without diagnostic data from birth, the first diagnosis may be missing from registry data, but most will have the diagnostic code registered in subsequent health-care use. Diagnostic groups were defined for receiving a diagnostic code at least one time for ADHD (F90), autism (F84.0, F84.1, F84.5, F84.8, and F84.9), intellectual disability and general developmental delay (F7 and F83), specific conditions of speech and language (F80, F98.5, F98.6), specific conditions of scholastic skills (F81), specific conditions of motor function (F82), and tic conditions (F95). The validity of autism diagnoses in MoBa has been studied previously [45]. Of 61 children identified with at least one instance of autism code in NPR, 58 (95%) received an autism diagnosis based on an independent multi-disciplinary standardized diagnostic assessment, and record review for another 567 children showed that for 86% the diagnostic criteria were well-documented in the health records [46]. No exclusions were made on the basis of other co-occurring disabilities or potentially contributing causes of disability (e.g., cerebral palsy, identified genetic syndromes, premature birth, birth complications) because this might lead to a biased or incomplete representation of children with neurodevelopmental conditions.

Sensitivity analyses for the main diagnostic outcomes (ADHD, autism, and intellectual disability/general developmental delay) were run restricting to individuals who had received a diagnostic code more than once to address the possibility of misdiagnosis or coding errors. Most diagnoses will have occurred after age 3, but some

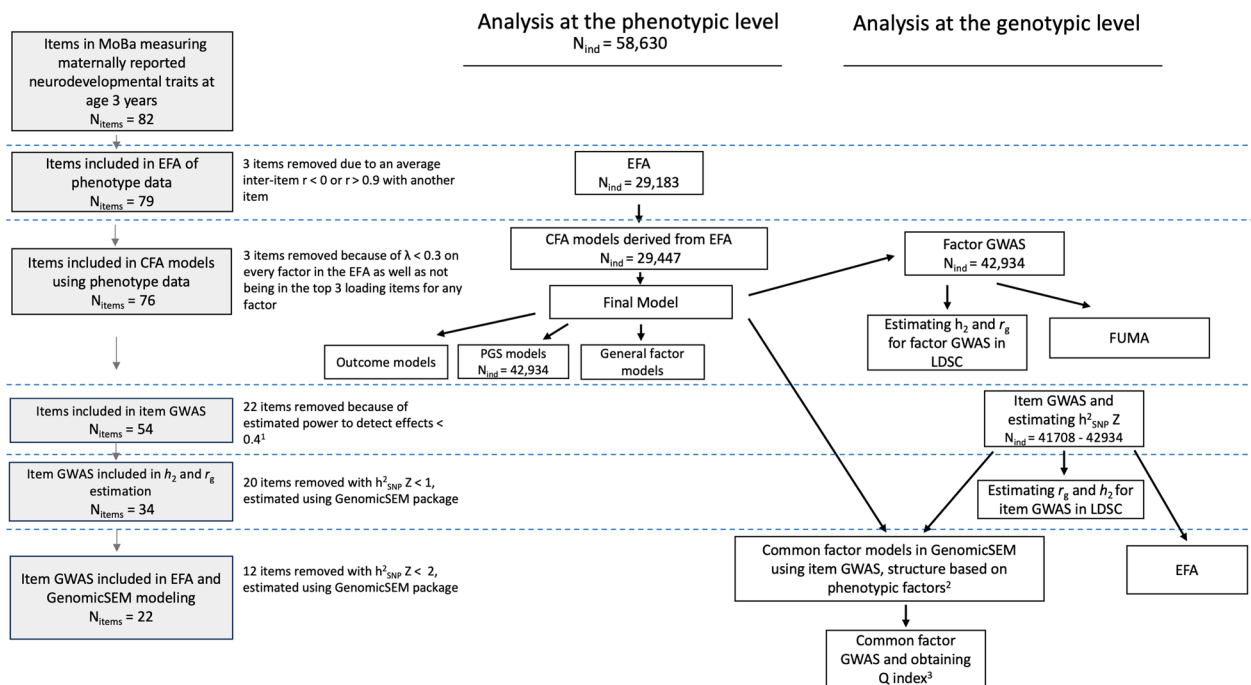


Fig. 1 Outline of study design and main analyses at the phenotypic and genotypic levels. Grey boxes outline the steps where questionnaire items were removed with the exclusion thresholds listed to the right. Boxes indicate an analysis with the arrows denoting analyses which are based on (i.e., factor structure) or used results (i.e., summary statistics) from a previous analysis. Analyses conducted at the phenotypic level with no sample size listed were conducted in the full sample ($N = 58,630$). Half-samples for the EFA/CFA conducted in the phenotypic level were randomly selected halves of the full sample. Estimating r_g refers to estimation of genetic correlations of the items/factors with neurodevelopmental conditions. 1 With the assumptions of an OR of 1.2, MAF of 0.01, and alpha of 0.01 in a logistic model with additive genetic effects. 2 Only common factor models with 3+ items run. 3 Common factor GWAS only run on models with good fits and significant factor loadings

children will have already had a diagnosis at the time the questionnaire data was collected. This impacts different diagnostic outcomes differently. For example, using only individuals with NPR data available at age 3 so that all diagnoses by age 3 are captured, the percent of the diagnostic group who had received a diagnosis before age 4 was as low as 0.2% for ADHD, 6% for autism, 11% for intellectual disability/global developmental delay and up to 26% for specific conditions of motor function. Percentages for all outcomes are available in Additional file 1: Table S2.

Additionally, clinically relevant outcomes for having multiple neurodevelopmental conditions diagnoses registered as well as any psychiatric hospitalization were derived from NPR. Several clinically relevant outcomes were also coded using the MoBa questionnaire data. These included measures of maternal report of early (by age 3) referral to service use (habilitation service, educational psychology service, or child psychiatric clinic/department) as well as later maternally perceived impact and impairment from difficulties in development and behavior in their child's life in the age 5 & 8 questionnaires. Further information on the scales,

the items used in the factor models, and diagnostic and clinically relevant outcomes are available in Additional file 2: supplementary methods and Additional file 1: Tables S1–3.

Polygenic scores

Polygenic scores (PGS) were estimated with the software PRSice2 [47] based on summary statistics from the most recent Psychiatric Genomic Consortium GWAS for ADHD [6], autism [7], and schizophrenia [48]. ADHD and autism were included as they are neurodevelopmental conditions with well powered and publicly available GWAS summary statistics. Schizophrenia was included given neurodevelopmental aspects to its development [31, 49, 50]. Scores were regressed on the first 10 genomic principal components (PCs) and genotype batch. The first principal component of 11 scores, constructed based on p-value thresholds between 5×10^{-8} and 1, was used for the subsequent analyses. This approach controls for type one error rate arising from optimization of pruning and thresholding while still maintaining prediction performance [51].

Analyses

An overview of the analyses performed as well as thresholds for item inclusion in each analytic step are presented in Fig. 1. Lenient thresholds for item selection were chosen to maximize the number of traits across different areas of development. Analytical code can be found at https://github.com/psychgen/neurodevelopment_traits_structure.

Exploratory and confirmatory factor analyses

Exploratory factor analysis (EFA) was performed in one randomly selected half of the full sample ($n=29,183$). Confirmatory factor analyses (CFA) were run in the other half of the full sample ($n=29,447$) for possible viable models derived from the EFA. Using standard fit indices (CFI, TLI, RMSEA) the best fitting model out of these possible models was used as the final model for all downstream analyses. In the full sample, both bifactor and higher-order models were run alongside the final selected correlated factor model to assess a unidimensional factor. To address potential sex differences in the measurement of these factors, we conducted measurement invariance testing in the full sample. A multi-group CFA (MG-CFA) of the correlated factor model by sex ($N_{\text{males}}=29,955$, $N_{\text{females}}=28,589$) was used to test for configural invariance and invariance of thresholds and loadings [52]. See Additional file 2: supplementary methods for further details on the factor analyses, criteria for model selection, and measurement invariance testing.

Measurement models with neurodevelopmental diagnoses, clinically relevant outcomes, and polygenic scores

The factor associations with diagnostic outcomes served two purposes: (1) validation and further characterization of the factors and (2) insight into how specific areas of development at age 3 are related to receiving a particular neurodevelopmental condition diagnosis. A correlated factor and a higher-order general factor model were run specifying the factors to predict neurodevelopmental diagnoses and other clinically relevant outcomes. In the correlated factor models, both univariate models with the factors predicting the outcomes individually and multiple regression models with factors predicting the outcome simultaneously were run. Due to collinearity concerns in the multiple regression models arising from groups of highly correlated factors, the magnitude of the factors' effects within those groups were constrained to be equal in the correlated factor model. A higher-order model was run to assess if factors moderated the effect of a general factor on the outcomes as well as gain some insight into the factor effects on outcomes that are unique to the specific factor between highly correlated factors in

the correlated factor model. In the higher-order model, general and specific factors were specified to predict outcomes separately in two models. Measurement models including PGS as explanatory variables for the factors were run in the correlated factors and higher-order model. Models were run in a multi-group SEM framework, grouped by sex with both regression effects and model parameters estimated for each sex separately.

Factor analyses software

EFA analyses were all run using the weighted least square mean and variance adjusted (WLSMV) estimation method and with a geomin oblique rotation applied in the Mplus statistical software (Muthén & Muthén, 2011). All CFA and measurement invariance models were run using the *lavaan* (v0.6–14) and *semTools* (v0.5–6) packages in R with the WLSMV estimation method [53, 54]. Missing data was handled using pairwise deletion for both the EFA and CFA, as it is the default in Mplus for categorical data.

Genome-wide association studies

Genome-wide association studies (GWAS) were run on each individual item (item GWAS) for which power calculations indicated sufficient statistical power, and on factor scores estimated for each factor (factor GWAS). This was done both to investigate the genetic effects underlying the factors we identified as well as to investigate the specificity of genetic effects between the factor and item levels of analysis. Factor scores were estimated using parameters for each sex from the correlated factor model multi-group CFA using the Empirical Bayes Model approach, the *lavaan* default method for categorical indicators. All GWAS included sex, genotype batch, and the first 10 PCs as covariates. Additional sex specific GWAS were run as sensitivity analyses for the factors. GWAS were run using version 3.1 of the REGENIE software, a computationally efficient linear mixed model method of conducting multi-trait GWAS. REGENIE can handle relatedness in the sample and correct for unbalanced case-control phenotypes in binary phenotypes [55]. For all factor and feasible item GWAS, SNP-based heritability (h^2_{SNP}) and genetic correlations (r_g) with ADHD [6], autism [7], and schizophrenia [48] were estimated using linkage disequilibrium score regression (LDSC) [56]. Estimated h^2_{SNP} for the item GWAS was on the liability scale. Functional mapping and annotation of the factor GWAS results were performed with FUMA (v1.5.3) [57]. Further information on sample sizes, prevalence estimates for LDSC, and power estimates used for the above analyses are listed in the Additional file 2: supplementary methods and Additional file 1: Table S4.

Genomic factor modeling and specificity of SNP effects

Genomic factor modeling used selected item GWAS. A lenient power inclusion threshold of $Z > 2$ as opposed to a more standard heuristic of $Z > 4$ for the item GWAS meant that power was borderline for genomic factor modeling. Because of this, an EFA was conducted on the estimated smoothed genetic correlation matrix of all chromosomes as opposed to only on even or odd chromosomes, which has been done to guard against overfitting if performing downstream analyses based on the EFA [9]. Therefore, no further downstream analyses (e.g., CFA) were conducted based on the results. Version 4.1.2 of the R package *stats* [58] was used to run the EFA and a promax rotation was applied. Common factor models based on factors from the phenotypic models that had at least three items meeting the item GWAS power threshold were run. For those with good fits and significant factor loadings, a common factor GWAS was run estimating SNP and Q_{SNP} effects. Q_{SNP} being a measure of how well the association of the SNP and the individual trait is accounted for by the factor [9, 59]. All confirmatory genomic factor modeling and GWAS were conducted using diagonally weighted least squares (DWLS) estimation in version 0.0.5 of the *GenomicSEM* R package [59].

Results

Phenotypic factor structure underlying early neurodevelopmental traits

Results of the EFA (Additional file 1: Tables S5–6) and CFA models indicated high dimensionality underlying early neurodevelopmental traits. Procedures to determine the optimal number of factors to retain indicated between 1 and 15 factors (Additional file 2: Figure S1) and fit indices from the EFA showed models with more than 9 factors met good fit criteria (Additional file 1: Table S5). Balancing these results with the interpretability of the factors, 3 models (9, 10, and 11-factor models) were selected to be run as confirmatory factor models in the other half of the sample. The 11-factor showed the best fit for complexity-penalized fit indices out of the three in both the EFA (Additional file 1: Table S5) and CFA (Additional file 1: Table S7). The 11-factor model was selected to be used in the downstream analyses.

The 11-factor model included factors roughly corresponding to areas of prosocial behavior (prosocial), motor development (motor), nonverbal communication and joint attention (NVcom), social attention and interest (Social-Att), language and verbal communication (language), play, repetitive and restricted behaviors and interests (RepBehavior), repetitive and idiosyncratic speech (RepSpeech), waiting, inattention and overactivity (inattention), and impulsivity. Most items (73/76) loaded well ($\lambda > 0.4$) onto their respective factors (Additional file 2: Figure S2).

Additionally, all factors except the idiosyncratic speech and impulsivity factors had moderate to high positive correlations with most other factors (Fig. 2). Factors covering the broad domains of social/communication, ADHD traits, and repetitive behaviors and speech were highly correlated amongst themselves but showed differing patterns of correlation with factors outside their broad domains. Parameter estimates of the final model are presented in Additional file 1: Tables S8–11. Finally, measurement invariance testing showed that invariance of thresholds and loadings held, so factors were assumed to largely represent the same constructs between males and females (Additional file 1: Table S12).

An additional general factor explaining all covariance between the different factors of early neurodevelopment had poor model fit indices (Hierarchical CFI: 0.621, TLI: 0.609, RMSEA: 0.032; Bifactor: CFI: 0.644, TLI: 0.624, RMSEA: 0.031) compared with the correlated factor model (CFI: 0.888, TLI: 0.883, RMSEA: 0.018) in the full sample. Besides fit indices, anomalous results in parameter estimates, non-uniform ($\lambda = 0.07–0.89$) loadings, and several specific factors with variances estimated close to zero indicated misspecification of the bifactor model to the data. For indicators other than model fit, this was less apparent in the hierarchical model (Additional file 1: Tables S13–14); therefore, it was used for further analyses. However, the general factor still exhibited varied loadings ($\lambda = 0.313–0.787$) and was characterized by factors encompassing social, communication, and motor development, which all had strong loadings from items with low endorsement in the general population.

Factor validation and correlations with later outcomes

We found that nearly all early neurodevelopmental factors were associated with receiving a diagnosis of any of the neurodevelopmental conditions, higher perceived impact in daily life at ages 5 and 8, later psychiatric inpatient services, and reported early referral to habilitation, special education, and psychiatric services (Additional file 2: Figures S3–S5). In multiple regression models, all outcomes were still associated with at least one factor or group of highly correlated factors, and many were associated with multiple (Fig. 3; Additional file 2: Figures S6, S7). For example, both the highly correlated groups of the ADHD-trait factors, and social and communication factors were still associated with later receiving a diagnosis of ADHD. Estimates of these associations did not differ when restricting the sample to those who had received a diagnostic code at least twice, although precision decreased slightly (Additional file 2: Figure S8). Some of these associations also differed by sex, such as the motor factor being associated with an autism diagnosis only in girls in this model.

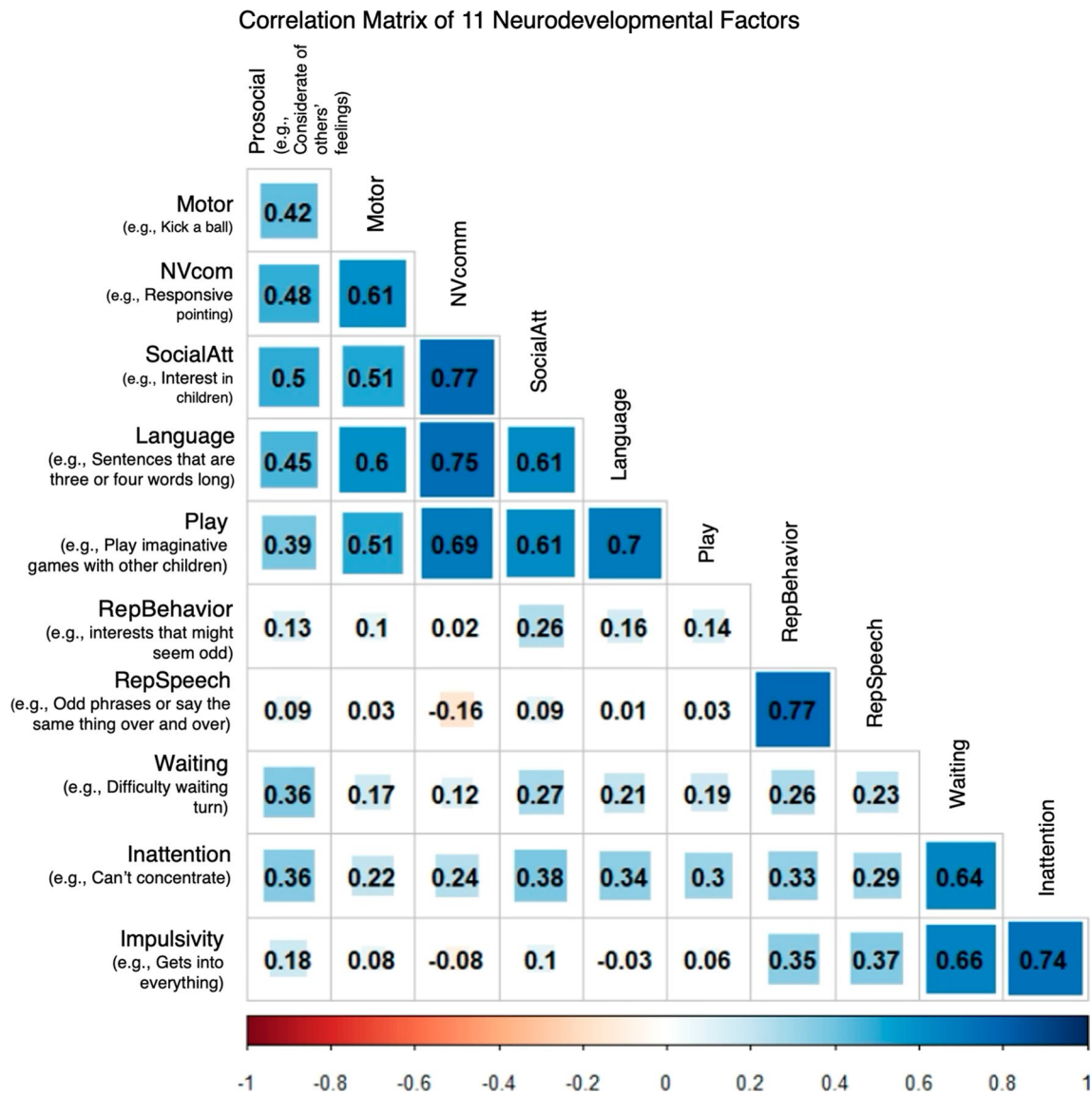


Fig. 2 A correlation matrix of the 11 factors from the correlated factor model in the full population. Factors include prosocial behavior (prosocial), motor development (motor), nonverbal communication and joint attention (NVcomm), social attention and interest (Social Att), language and verbal communication (language), play, repetitive and restricted behaviors and interests (RepBehavior), repetitive and idiosyncratic speech (RepSpeech), waiting, inattention and overactivity (inattention), and impulsivity. An example item from the factor is listed for each factor

In the hierarchical model, where specific factors simultaneously predicted the outcomes, all factors were still associated with at least one outcome and some factors within the highly correlated factor groups had differing magnitude and direction of effects from each other (Additional file 2: Figures S9–S11). For example, out of the highly correlated social and communication factors, only the play and language factors were associated with ADHD. These two factors also had the most associations in the higher-order models, both being

significantly associated with most of the diagnostic outcomes. The general factor was associated with all outcomes (Additional file 2: Figures S9–S11). However, the effect of a general factor on the outcome, when moderated by the specific factors, primarily explained additional variance in the outcomes related to early referral, general developmental delay/intellectual disability, and, in girls, specific language conditions when compared to the correlated factors model (Additional file 2: Figure S12).

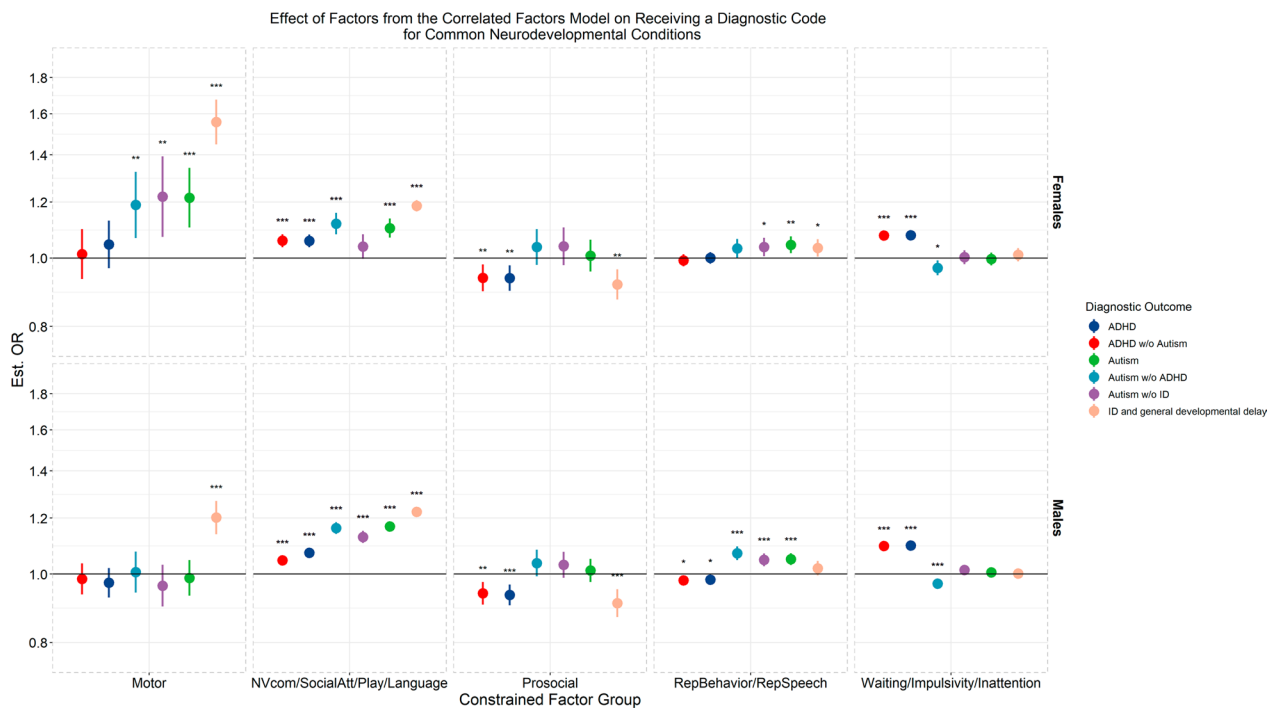


Fig. 3 Estimated effects of factors from the correlated factor model in a multivariate regression controlling for the effects of all factors on the outcome for 5 selected diagnostic outcomes. Effects are presented as odds ratios calculated from the exponential of the standardized beta value from the logistic regression in the measurement models. 95% percent confidence intervals are shown. Due to high correlations amongst domains in the broad areas of social communication (the language & verbal communication, nonverbal communication and joint attention, play, and social attention and interest factors), ADHD-associated traits (the inattention and overactivity, waiting, impulsivity factors), and repetitive and restricted behaviors (the repetitive and idiosyncratic speech and repetitive and restricted behaviors and interests factors) effects of these factors were constrained to be equal to avoid collinearity issues. “*” “**” “***” “****” denote adjusted $p < 0.05$, < 0.01 , and < 0.001 respectively, after multiple testing correction. For full results of the outcome models, see the supplementary results (Additional file 2)

Common genetic variance underlying early neurodevelopmental traits

GWAS of the factor scores from the 11-factor model had low h^2_{SNP} estimates. Four factors had estimated confidence intervals that crossed 0 (Additional file 1: Table S15). The highest estimate was the non-verbal communication factor ($h^2_{SNP}=0.037$ [0.013–0.061], $p=0.003$). Four unique genome-wide significant loci identified across the factors, three of which were associated with multiple factors (Additional file 1: Table S16). Results from gene-based association analyses implemented in FUMA (significant $p < 2.682 \times 10^{-6}$; Table S17), identified *CNGB3* ($p=1.53 \times 10^{-6}$) as associated with the motor factor as well as *RSRC1* ($p=3.95 \times 10^{-7}$) and *ADAMTS17* ($p=8.19 \times 10^{-7}$) with the prosocial behavior factor. Sex-stratified factor GWAS showed high genetic correlations with the factors in the full sample. These GWAS showed some differences in h^2_{SNP} estimates by sex, but these differences did not reach statistical significance (Additional file 1: Table S18). 34 item GWAS reached our greater than 1 h^2_{SNP} Z threshold (Additional file 1: Table S19), of which 21 items had h^2_{SNP} that

reached statistical significance. These items had a large range of estimated h^2_{SNP} (range: 0.02–0.27; Additional file 1: Table S20) with differing levels of precision.

Early neurodevelopmental traits relationships with genetic liability for neurodevelopmental conditions

Genetic correlations between early neurodevelopmental traits and neurodevelopmental conditions were observed across multiple domains, as shown in Fig. 4, and were evident at both the factor (Additional file 1: Table S21) and item-level (Additional file 1: Table S22). ADHD had the highest genetic correlation with the inattentive and overactivity factor ($r_g=0.95$ [0.13–1]). The prosocial behavior factor had the highest significant association for both autism ($r_g=0.56$ [0.29–0.83]) and schizophrenia ($r_g=0.20$ [0.05–0.34]). We find some instances of differing effects across conditions, such as the positive genetic correlation between the motor factor and autism ($r_g=0.42$ [0.11–0.72]), and to a lesser extent, schizophrenia ($r_g=0.17$ [0–0.34]) but a negative correlation with ADHD ($r_g=-0.32$ [–0.58 to –0.01]).

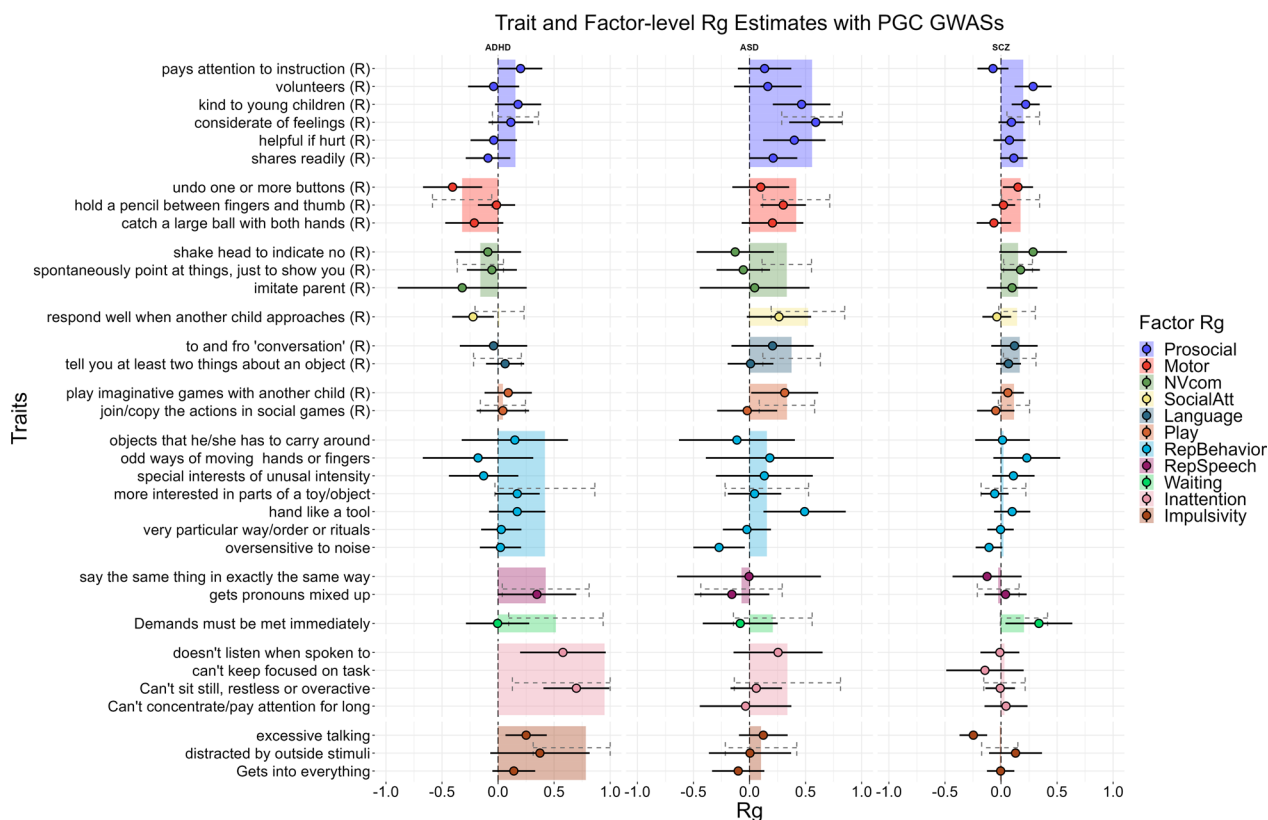


Fig. 4 Estimated item and factor loading GWAS genetic correlation with PGS GWAS. 95% percent confidence intervals are presented. Results of multiple testing corrections are presented in Additional file 1: Tables S21 and S22 as a reference for the strength of statistical significance. Items are represented by points and factors are represented by bars. Bar width only reflects the number of items from the factor that were included. (R) denotes reversed coded items. The inattention factor had an estimated genetic correlation above one but is shown just below 1.0. This factor as well as the impulsivity factor had upper bounds of the confidence interval estimated over 1. Item-level estimates were removed if confidence intervals were estimated as having a range larger than 1.5

The factors from the sex-stratified GWAS displayed similar genetic correlations with the neurodevelopmental conditions as in the entire sample but with slightly higher correlation estimates in males than females with autism and slightly higher in females than males with schizophrenia (Additional file 1: Table S23). These correlations were accompanied by large, overlapping confidence intervals but were in accordance with the findings of the PGS analyses. In these analyses, effects surviving multiple testing corrections were found exclusively in males for the autism PGS and in females for the schizophrenia PGS (Additional file 2: Figures S13–S14 and Additional file 1: Tables S24–S25).

At the item-level, the item “*considerate of feelings*” had the highest genetic correlation with autism, the item “*can’t sit still, restless or overactive*” with ADHD, and “*volunteers to help others*” with schizophrenia (Additional file 1: Table S22). A few item GWAS had differing effects compared to their specified factor’s GWAS. For example, the item measuring “*excessive talking*”, which was a part

of the CBCL and loaded onto the impulsivity factor, was significantly negatively correlated ($r_g = -0.25$ [−0.37 to −0.124]) with schizophrenia after multiple testing corrections while the impulsivity factor was uncorrelated with schizophrenia ($r_g = -0.01$ [−0.17–0.15]).

Genomic structure modeling and specificity of SNP effects

Given power constraints, the EFA was run on the smoothed estimated genetic correlation matrix of all chromosomes and no further downstream analyses were performed. Genetic correlations between all items were estimated and are presented in Fig. 5. Two to three clusters of items seem to emerge from this, the most obvious being the prosocial behavior items and the item “*uses hand like a tool.*” These items were notably the items with the highest genetic correlations with autism. Extracting one “general” factor in the EFA left many items unrepresented. Further extractions of factors beyond a single factor were hard to interpret and frequently had factors

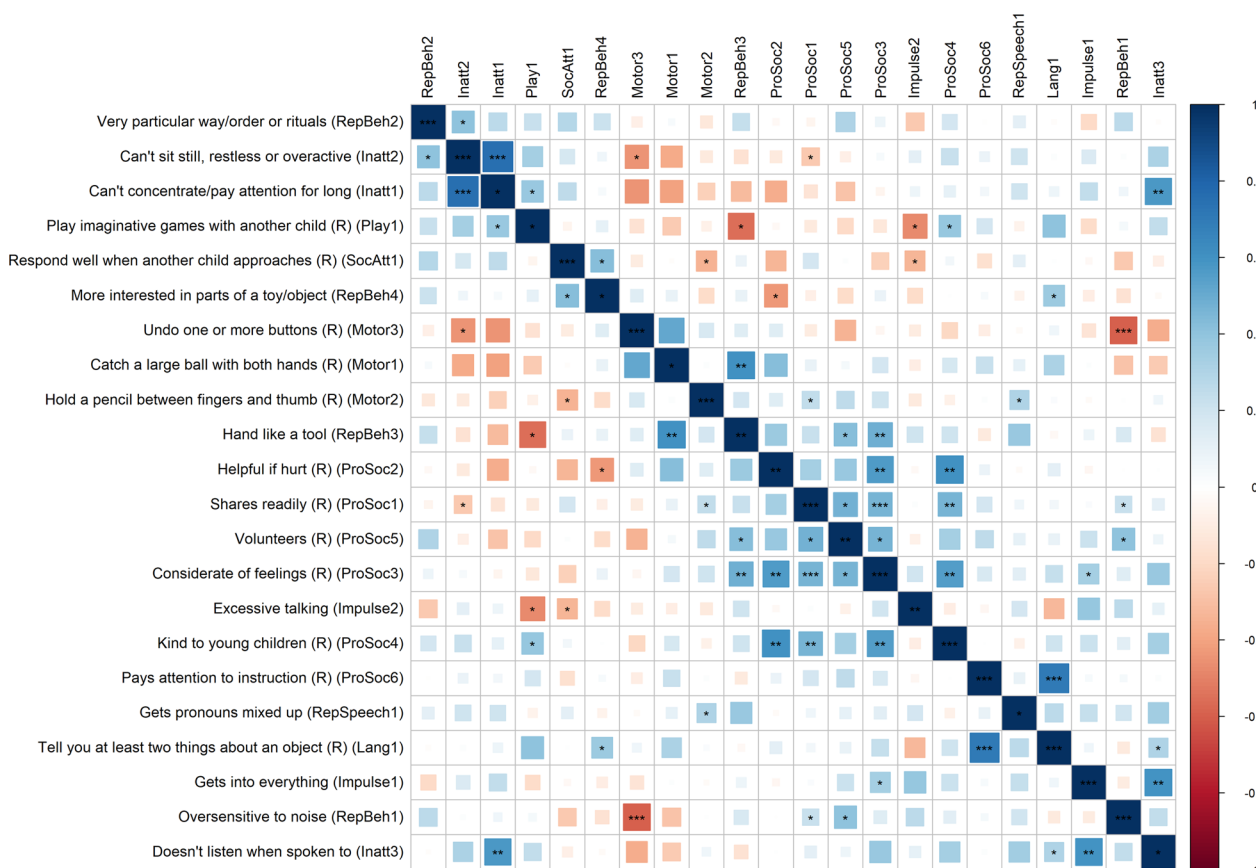


Fig. 5 The estimated smoothed genetic correlations matrix for the 22 neurodevelopmental items used in the EFA and genetic factor modeling. Items order using angular order of the eigenvectors (AOE). “*” “**” “***” “****” “*****” denote uncorrected $p < 0.05$, < 0.01 , and < 0.001 respectively

defined by a few items, frequent cross-loadings, and strong negative loadings.

The motor, prosocial behavior, RepBehavior, and inattention factors were recreated via a CFA at the genomic level. Among these, only the prosocial behavior factor demonstrated an exceptional fit (CFI=1, SRMR=0.095; Additional file 2: Figure S15) and significant loadings for most items. The other common factor models either had loadings that did not reach statistical significance, fit indices that could not be estimated, or, in the case of the motor factor, did not converge. Based on this, we only performed a subsequent common factor GWAS for the prosocial behavior factor, which did not yield any genome-wide significant loci but identified more Q_{snp} than SNP hits at a suggestive association threshold ($p < 5 \times 10^{-5}$; see Additional file 2: supplementary results and Additional file 1: Tables S26–S27).

Discussion

We leverage the item-level questionnaire data in up to 58,630 MoBa children to investigate patterns of relationships between specific traits from different areas of

development in early childhood, the underlying genetic contributions, and potential shared etiology to clinically diagnosed neurodevelopmental conditions. We find that difficulties across all areas of early neurodevelopment are associated with receiving diagnoses across a range of neurodevelopmental conditions. Particularly, early difficulties in social and communication domains are associated with receiving a diagnosis of almost all neurodevelopmental conditions, suggesting that these are trans-diagnostically relevant in neurodevelopmental conditions. Additionally, the genetic contributions underlying variation in the general population for several early neurodevelopmental domains are genetically correlated with ADHD, autism, and schizophrenia. Finally, we find limited evidence of shared common genetic effects across all areas of neurodevelopment. From these results, we draw two over-arching conclusions. First, at both the phenotypic and genotypic levels, there is high heterogeneity in the underlying effects on variation in these traits—higher than would be expected if these traits were neatly aligned with distinct neurodevelopmental conditions. Secondly, despite their etiological and structural

heterogeneity, early neurodevelopmental traits in a general population sample *are* phenotypically and genetically associated with neurodevelopmental diagnoses.

Heterogeneity underlying early neurodevelopmental traits in a population-based sample

We find that most domains of neurodevelopment traits are at least moderately correlated with each other at the phenotypic level. The simplest etiological explanation for this would be shared liability across all areas of neurodevelopment, such as a general genetic neurodevelopment factor, which has been suggested based on a previous twin study [5]. However, we find substantial heterogeneity underlying early neurodevelopmental traits both at the phenotypic and genetic levels, and little evidence supporting a general factor of liability to all early neurodevelopmental traits at either level of analysis. Notwithstanding the question of the existence of a general factor for neurodevelopmental traits, we observe increased heterogeneity compared to what would be expected based on etiological factors that neatly lined up with diagnostic criteria. The observed factors are highly correlated amongst themselves in domains related to commonly separated neurodevelopmental domains (i.e., social and communication, repetitive behaviors, ADHD-related traits). However, these factors are differentially correlated with the factors outside of their domain and, when correlations between factors were accounted for by a general factor, have differing associations with later diagnoses. To note, while the neurodevelopmental traits are associated with and share some genetic variance with neurodevelopmental conditions, traits are non-specific to conditions. This pattern is consistent with co-occurrence between neurodevelopmental conditions being commonplace—in many cases "being the rule, rather than the exception" [3, 60, 61].

The results of the genomic factor modeling point towards a similar level of heterogeneity in the genetic architecture of early neurodevelopmental traits. We find some evidence for a genetic factor that resembles the prosocial factor identified in the phenotypic models and some shared genetic loci across areas of neurodevelopment, particularly across social and communication and prosocial behavior. However, in other areas, such as motor development, increased heterogeneity is observed at the item-level, potentially suggesting different genetic mechanisms underlying different aspects of motor skills. Lastly, even among the items measuring prosocial behavior, the higher number of Q_{SNP} hits compared to SNP hits contributing to the common genetic factor at a suggestive association threshold emphasizes the possibility of item-level specificity of genetic effects, even within the most coherent genetic factor.

Early neurodevelopmental traits are associated with neurodevelopmental conditions

The factors identified as underpinning early neurodevelopmental traits in our sample were all associated with receiving a clinical diagnosis for different neurodevelopmental conditions. We find stronger associations between conditions and factors that contain items that overlap with diagnostic criteria of that condition, such as the inattention, impulsivity, and waiting factors with ADHD. One notable exception to this concerns the repetitive speech and behavior factors. Although these factors were independently associated with receiving an autism diagnosis, the effect sizes were lower compared to factors covering social communication difficulties. Further, when controlling for variation in other areas of development, these factors were no longer associated with receiving a diagnosis in some of the subgroups of the diagnostic outcome, such as girls who received an autism diagnosis but not a diagnosis for ADHD. What the repetitive speech and behavior factors are capturing in the general population should be considered in the interpretation of this finding. Most of the items that make up repetitive speech and behavior factors are from the SCQ, which, as a diagnostic screener, has mixed findings on the validity of its use for children under 4 [62]. Additionally, items in these factors are endorsed relatively frequently compared to the other SCQ items in our sample, meaning maternal reports of these items may be primarily capturing behaviors in the more typical range of these traits.

Stronger associations between factors and the diagnostic outcomes are seen for conditions that have higher rates of earlier referral or diagnosis in our sample, such as intellectual disability and specific motor conditions. The strength of these associations is likely impacted by the overlap in items with diagnostic criteria, however, we still find associations of factors with conditions with later average age of diagnosis, such as specific learning conditions, and with conditions that do not have diagnostic criteria overlapping with the factor, such as the social and communication factors with an ADHD diagnosis without co-occurring autism.

We find early childhood neurodevelopmental traits share common genetic liability with ADHD, autism, and schizophrenia. We identified associations between autism, and to a lesser extent, schizophrenia genetic liability with early motor, language, and social traits in contrast to some work in general populations, including in a smaller subset of MoBa, which have found few associations between these conditions and early childhood behaviors [17, 63]. Although the previous finding of autism genetic liability being associated with motor difficulties at age 3 in MoBa [17] remains in our larger sample. There are mixed findings on the association between

ADHD genetic liability and social, communication, and repetitive behaviors and interests [64, 65]. We primarily find weak evidence for ADHD genetic liability contributing to early repetitive and idiosyncratic speech. Whilst it is tempting to read into discrepancies, it is important to note here the potential impact of well-characterized issues with the portability of polygenic scores across different study samples and populations [66], as well as the potential impact of chance variation—given the small effect sizes with which polygenic liability appears to manifest early in life [68]. Based on these inconsistent results across samples and given the current stage of genomic discovery for neurodevelopmental conditions, this result should not be over-interpreted. However, our findings are in line with previous findings of inattention and hyperactivity traits as well as social and communication behavior in the general population sharing genetic variation with ADHD and autism, respectively [16, 65]. Finally, we observe other consistent findings with the literature, such as autism and schizophrenia genetic liability being associated with lower prosocial behavior [64, 67] and ADHD with fewer early motor difficulties [63].

Many items suggest similar associations to neurodevelopmental conditions as their factor. However, we do observe some trait-level heterogeneity similar to the item-level or sub-domain level genomic analysis of neuroticism [69] and impulsivity [70]. For example, while the prosocial factor was genetically correlated with autism, only the factor's items "*kind to young children*", "*helpful when hurt*", and "*considerate of feelings*" had associations with autism after multiple testing corrections. Items in the motor and repetitive behavior factors also show some trait-level heterogeneity. These observations offer some potential areas for follow-up work in clinical samples identifying differentiating mechanisms of early development across conditions. Our findings also identify some potential for shared mechanisms across domains at the sub-diagnostic level: for instance, repetitive behaviors and speech with ADHD traits. Genetic liability to ADHD has also been associated with repetitive behaviors and interests in clinical samples of autistic individuals [18, 33]. However, as previously mentioned, the potential impact of chance variation and the validity of items in a general population should also be considered with this observation. For example, items such as "*says the same thing over and over*" could be misinterpreted by parents, resulting in it capturing activity level or more common behaviors, rather than the idiosyncratic speech typically associated with autism.

Limitations

There are some limitations of our study that should be considered. Despite splitting our sample into discovery and test halves, the exploratory factor analysis of such

a diverse set of items in a large sample is likely to have led to some level of overfitting. Because of this, we do not suggest interpreting all identified factors as necessarily definitive distinct factors but instead put forward that there is increased dimensionality across areas of development with differing relationships to each other and neurodevelopmental conditions that may be lost at the diagnostic or scale level. Although we included measures of most behavioral domains of neurodevelopmental conditions, we did not have measures of all domains, such as cognitive ability and tics. Another consideration is that we used registry data to create our diagnostic outcomes. A limitation of this is that we cannot distinguish between subsequent co-occurring diagnoses and substitution diagnoses. Further, data being only available after 2008 means we do not know for certain how many had a diagnosis at the time mothers filled out the questionnaire. As these individuals were included this could impact the strength of the associations seen for some of the diagnostic outcomes.

Low power to detect signal for many of the item GWAS limits the claims. For the effects we identified the increased variation due to underpowered GWAS may contribute to the large range of estimated genetic correlations. However, power concerns are unlikely to fully explain the lack of a single general genetic factor. Our genetic analyses were also limited to common genetic variants, which may contribute to a lack of a general genetic factor as there is considerable overlap of rare variants associated with different neurodevelopmental conditions [10–12]. Finally, the genetic analyses were limited to participants in MoBa of European genetic ancestry, limiting the generalizability of our results across ancestries.

Conclusions

Our exploratory results reveal the multidimensionality underlying early neurodevelopmental traits in a population-based birth cohort. These dimensions are broadly associated with receiving a diagnosis of neurodevelopmental conditions, and many are genetically correlated with ADHD, autism, and/or schizophrenia. We find little support for a shared common genetic liability across all traits in the general population. Instead, we observe multiple specific factors with certain shared genetic loci identified across, particularly, the social and communication domains of neurodevelopment, but none that are evidently relevant across all domains. Our trait-level analyses highlight the role of heterogeneous genetic effects underlying early neurodevelopmental traits and their relationships to neurodevelopmental conditions. These findings provide areas for further investigation to identify shared and distinct mechanisms across neurodevelopmental conditions.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13229-024-00599-0>.

Additional file 1: Supplementary Tables.

Additional file 2: Supplementary Methods, Results, and Figures.

Acknowledgements

The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the participating families in Norway who take part in this on-going cohort study. We thank the Norwegian Institute of Public Health (NIPH) for generating high-quality genomic data. This research is part of the HARVEST collaboration, supported by the Research Council of Norway (#229624). We also thank the NORMENT Centre for providing genotype data, funded by the Research Council of Norway (#223273), South East Norway Health Authorities and Stiftelsen Kristian Gerhard Jebsen. We further thank the Center for Diabetes Research, the University of Bergen for providing genotype data and performing quality control and imputation of the data funded by the ERC AdG project SELECTionPREDISPOSED, Stiftelsen Kristian Gerhard Jebsen, Trond Mohn Foundation, the Research Council of Norway, the Novo Nordisk Foundation, the University of Bergen, and the Western Norway Health Authorities. This work was performed on the TSD (Tjeneste for Sensitive Data) facilities, owned by the University of Oslo, operated and developed by the TSD service group at the University of Oslo, IT Department (USIT). (tsd-drift@usit.uio.no). Analyses were performed on resources provided by Sigma2—the National Infrastructure for High-Performance Computing and Data Storage in Norway. Disclaimer: Data from the Norwegian Patient Registry has been used in this publication. The interpretation and reporting of these data are the sole responsibility of the authors, and no endorsement by the Norwegian Patient Registry is intended nor should be inferred.

Author contributions

L.H., L.J.H., and A.H. contributed to the design of the research questions. L.H., L.J.H., A.H., E.C.C., A.D.A., and A.G.A. contributed to the methodological design of the work. E.C.C. and L.J.H. contributed to data curation and quality control. L.H. conducted analyses for the work and drafted the main manuscript text. L.H., E.C.C., A.D.A., A.G.A., R.B.A., A.R., H.A., B.S., O.A.A., L.J.H., and A.H. contributed to the interpretation of the data and substantive revisions of the manuscript. L.J.H. and A.H. provided supervision for the work.

Funding

Open access funding provided by Norwegian Institute of Public Health (FHI). The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The South-Eastern Norway Regional Health Authority supported LEH (#2020022), ADA (2020023), AH (#2020022), LJH (#2019097, #2922083), and ECC (#2021045). The Research Council of Norway supported AH, RBA, and ECC (#274611), HA (#324620), and OAA (#324499, #324252). AR was supported by the Simons Foundation Autism Research Initiative (724306).

Availability of data and materials

The consent given by the participants does not allow for storage of data on an individual level in repositories. Researchers can apply for access to data for replication purposes via MoBa, in line with MoBa data access policies. Analytical code for this study can be found at https://github.com/psychgen/neurodevelopment_traits_structure. Documentation for MoBa questionnaires can be found at <https://www.fhi.no/en/ch/studies/moba/for-forskere-artikler/questionnaires-from-moba/>. Publicly available summary statistics provided by the Psychiatric Genomics Consortium (<https://pgc.unc.edu/>) were used for the creation of PGS and estimation of genetic correlations of autism, ADHD, and schizophrenia with the early neurodevelopmental traits.

Declarations

Competing interests

Ole A. Andreassen discloses that he is a consultant of cortechs.ai, and has received speaker's honorarium from Lundbeck, Janssen and Sunovion with no conflict of interest relevant to this work. The other authors declare no conflicts of interest.

Author details

¹PsychGen Centre for Genetic Epidemiology and Mental Health, Norwegian Institute of Public Health, Oslo, Norway. ²Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway. ³Department of Psychology, University of Oslo, Oslo, Norway. ⁴Division of Psychology & Language Sciences, Department of Clinical, Educational & Health Psychology, Faculty of Brain Sciences, University College London, London, UK. ⁵Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK. ⁶Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. ⁷School of Psychology, Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK. ⁸PROMENTA Research Centre, Department of Psychology, University of Oslo, Oslo, Norway. ⁹Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands. ¹⁰MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, UK. ¹¹Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands. ¹²Institute of Clinical Medicine, University of Oslo, Oslo, Norway.

Received: 22 November 2023 Accepted: 18 April 2024

Published online: 07 June 2024

References

- Thapar A, Cooper M, Rutter M. Neurodevelopmental disorders. *Lancet Psychiatry*. 2017;4(4):339–46. [https://doi.org/10.1016/S2215-0366\(16\)30376-5](https://doi.org/10.1016/S2215-0366(16)30376-5).
- Lundström S, Reichenberg A, Melke J, Råstam M, Kerekes N, Lichtenstein P, Gillberg C, Anckarsäter H. Autism spectrum disorders and coexisting disorders in a nationwide Swedish twin study. *J Child Psychol Psychiatry*. 2015;56(6):702–10. <https://doi.org/10.1111/jcpp.12329>.
- Brimo K, Dinkler L, Gillberg C, Lichtenstein P, Lundström S, Åsberg Johnels J. The co-occurrence of neurodevelopmental problems in dyslexia. *Dyslexia*. 2021;27(3):277–93. <https://doi.org/10.1002/dys.1681>.
- Reiersen AM, Constantino JN, Volk HE, Todd RD. Autistic traits in a population-based ADHD twin sample. *J Child Psychol Psychiatry*. 2007;48(5):464–72. <https://doi.org/10.1111/j.1469-7610.2006.01720.x>.
- Petterson E, Anckarsäter H, Gillberg C, Lichtenstein P. Different neurodevelopmental symptoms have a common genetic etiology. *J Child Psychol Psychiatry*. 2013;54(12):1356–65. <https://doi.org/10.1111/jcpp.12113>.
- Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, Baldursdottir G, Belliveau R, Bybjerg-Grauholm J, Bækvad-Hansen M, Cerrato F, Chambert K, Churchhouse C, Dumont A, Eriksson N, Gandal M, Goldstein JI, Grasby KL, Grove J, Neale BM. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet*. 2019. <https://doi.org/10.1038/s41588-018-0269-7>.
- Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, Pallesen J, Agerbo E, Andreassen OA, Anney R, Awasthi S, Belliveau R, Bettella F, Buxbaum JD, Bybjerg-Grauholm J, Bækvad-Hansen M, Cerrato F, Chambert K, Christensen JH, et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet*. 2019;51(3):431–44. <https://doi.org/10.1038/s41588-019-0344-8>.
- Doust C, Fontanillas P, Eising E, Gordon SD, Wang Z, Alagöz G, Molz B, Pourcain BS, Francks C, Marioni RE, Zhao J, Paracchini S, Talcott JB, Monaco AP, Stein JF, Gruen JR, Olson RK, Willcutt EG, DeFries JC, et al. Discovery of 42 genome-wide significant loci associated with dyslexia. *Nat Genet*. 2022. <https://doi.org/10.1038/s41588-022-01192-y>.
- Grotzinger AD, Mallard TT, Akingbuwa WA, Ip HF, Adams MJ, Lewis CM, McIntosh AM, Grove J, Dalsgaard S, Lesch K-P, Strom N, Meier SM, Mattheisen M, Børglum AD, Mors O, Breen G, Lee PH, Kendler KS, Smoller JW, et al. Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nat Genet*. 2022;54(5):548–59. <https://doi.org/10.1038/s41588-022-01057-4>.

10. Satterstrom FK, Walters RK, Singh T, Wigdor EM, Lescai F, Demontis D, Kosmicki JA, Grove J, Stevens C, Bybjerg-Grauholm J, Bækvad-Hansen M, Palmer DS, Maller JB, Nordentoft M, Mors O, Robinson EB, Hougaard DM, Werge TM, Bo Mortensen P, et al. Autism spectrum disorder and attention deficit hyperactivity disorder have a similar burden of rare protein-truncating variants. *Nat Neurosci*. 2019;22(12):1961–5. <https://doi.org/10.1038/s41593-019-0527-8>.
11. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, Rubeis SD, An J-Y, Peng M, Collins R, Grove J, Klei L, Stevens C, Reichert J, Mulhern MS, Artomov M, Gerges S, Sheppard B, Xu X, Bhaduri A, Norman U, et al. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell*. 2020;180(3):568–584.e23. <https://doi.org/10.1016/j.cell.2019.12.036>.
12. Martin J, Cooper M, Hamshere ML, Pocklington A, Scherer SW, Kent L, Gill M, Owen MJ, Williams N, O'Donovan MC, Thapar A, Holmans P. Biological overlap of attention-deficit/hyperactivity disorder and autism spectrum disorder: evidence from copy number variants. *J Am Acad Child Adolesc Psychiatry*. 2014;53(7):761–770.e26. <https://doi.org/10.1016/j.jaac.2014.03.004>.
13. Ismail FY, Shapiro BK. What are neurodevelopmental disorders? *Curr Opin Neurol*. 2019;32(4):611–6. <https://doi.org/10.1097/WCO.00000000000000710>.
14. Reiersen A. How should we classify complex neurodevelopmental disorders? *Scand J Child Adolesc Psychiatry Psychol*. 2017. <https://doi.org/10.21307/sjcap-2017-005>.
15. Mullin AP, Gokhale A, Moreno-De-Luca A, Sanyal S, Waddington JL, Faundez V. Neurodevelopmental disorders: mechanisms and boundary definitions from genomes, interactomes and proteomes. *Transl Psychiatry*. 2013;3(12):e329–e329. <https://doi.org/10.1038/tp.2013.108>.
16. Robinson EB, St Pourcain B, Anttila V, Kosmicki JA, Bulik-Sullivan B, Grove J, Maller J, Samocha KE, Sanders SJ, Ripke S, Martin J, Hollegaard MV, Werge T, Hougaard DM, Neale BM, Evans DM, Skuse D, Mortensen PB, Børglum AD, et al. Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat Genet*. 2016;48(5):552–5. <https://doi.org/10.1038/ng.3529>.
17. Askeland RB, Hannigan LJ, Ask H, Ayorech Z, Tesli M, Corfield E, Magnus P, Njølstad PR, Andreassen OA, Davey Smith G, Reichborn-Kjennerud T, Havdahl A. Early manifestations of genetic risk for neurodevelopmental disorders. *J Child Psychol Psychiatry*. 2022;63(7):810–9. <https://doi.org/10.1111/jcpp.13528>.
18. Thomas TR, Koormar T, Casten LG, Tener AJ, Bahl E, Michaelson JJ. Clinical autism subscales have common genetic liabilities that are heritable, pleiotropic, and generalizable to the general population. *Transl Psychiatry*. 2022. <https://doi.org/10.1038/s41398-022-01982-2>.
19. van't Hof M, Tisseur C, van Berckelaer-Onnes I, van Nieuwenhuizen A, Daniels AM, Deen M, Hoek HW, Ester WA. Age at autism spectrum disorder diagnosis: a systematic review and meta-analysis from 2012 to 2019. *Autism*. 2021;25(4):862–73. <https://doi.org/10.1177/1362361320971107>.
20. Rocco I, Corso B, Bonati M, Minicuci N. Time of onset and/or diagnosis of ADHD in European children: a systematic review. *BMC Psychiatry*. 2021;21(1):575. <https://doi.org/10.1186/s12888-021-03547-x>.
21. Krakowski AD, Cost KT, Anagnostou E, Lai M-C, Crosbie J, Schachar R, Georgiades S, Duku E, Szatmari P. Inattention and hyperactive/impulsive component scores do not differentiate between autism spectrum disorder and attention-deficit/hyperactivity disorder in a clinical sample. *Mol Autism*. 2020;11(1):28. <https://doi.org/10.1186/s13229-020-00338-1>.
22. Ghanizadeh A. Factor analysis on ADHD and autism spectrum disorder DSM-IV-derived items shows lack of overlap. *Eur Child Adolesc Psychiatry*. 2010;19(10):797–8. <https://doi.org/10.1007/s00787-010-0127-3>.
23. Ghanizadeh A. Co-morbidity and factor analysis on attention deficit hyperactivity disorder and autism spectrum disorder DSM-IV-derived items. *J Res Med Sci*. 2012;17(4):368–72.
24. Ronald A, Larsson H, Anckarsäter H, Lichtenstein P. Symptoms of autism and ADHD: a Swedish twin study examining their overlap. *J Abnorm Psychol*. 2014;123(2):440–51. <https://doi.org/10.1037/a0036088>.
25. Martin J, Hamshere ML, O'Donovan MC, Rutter M, Thapar A. Factor structure of autistic traits in children with ADHD. *J Autism Dev Disord*. 2014;44(1):204–15. <https://doi.org/10.1007/s10803-013-1865-0>.
26. Ronald A, Happé F, Plomin R. The genetic relationship between individual differences in social and nonsocial behaviours characteristic of autism. *Dev Sci*. 2005;8(5):444–58.
27. Warrier V, Toro R, Won H, Leblond CS, Cliquet F, Delorme R, De Witte W, Bralten J, Chakrabarti B, Børglum AD, Grove J, Poelmans G, Hinds DA, Bourgeron T, Baron-Cohen S. Social and non-social autism symptoms and trait domains are genetically dissociable. *Commun Biol*. 2019. <https://doi.org/10.1038/s42003-019-0558-4>.
28. Mattheisen M, Grove J, Als TD, Martin J, Voloudakis G, Meier S, Demontis D, Bendl J, Walters R, Carey CE, Rosengren A, Strom NI, Hauberg ME, Zeng B, Hoffman G, Zhang W, Bybjerg-Grauholm J, Bækvad-Hansen M, Agerbo E, et al. Identification of shared and differentiating genetic architecture for autism spectrum disorder, attention-deficit hyperactivity disorder and case subgroups. *Nat Genet*. 2022;54(10):1470–8. <https://doi.org/10.1038/s41588-022-01171-3>.
29. Demontis D, Walters GB, Athanasiadis G, Walters R, Therrier K, Nielsen TT, Farajzadeh L, Voloudakis G, Bendl J, Zeng B, Zhang W, Grove J, Als TD, Duan J, Satterstrom FK, Bybjerg-Grauholm J, Bækved-Hansen M, Gudmundsson OO, Magnusson SH, et al. Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat Genet*. 2023;55(2):198–208. <https://doi.org/10.1038/s41588-022-01285-8>.
30. Lewis DA, Levitt P. Schizophrenia as a disorder of neurodevelopment. *Annu Rev Neurosci*. 2002;25(1):409–32. <https://doi.org/10.1146/annurev.neuro.25.112701.142754>.
31. Murray RM, Bora E, Modinos G, Vernon A. Schizophrenia: A developmental disorder with a risk of non-specific but avoidable decline. *Schizophr Res*. 2022;243:181–6. <https://doi.org/10.1016/j.schres.2022.03.005>.
32. Weiner DJ, Wigdor EM, Ripke S, Walters RK, Kosmicki JA, Grove J, Samocha KE, Goldstein JL, Okbay A, Bybjerg-Grauholm J, Werge T, Hougaard DM, Taylor J, Skuse D, Devlin B, Anney R, Sanders SJ, Bishop S, Mortensen PB, et al. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat Genet*. 2017;49(7):978–85. <https://doi.org/10.1038/ng.3863>.
33. Warrier V, Zhang X, Reed P, Havdahl A, Moore TM, Cliquet F, Leblond CS, Rolland T, Rosengren A, Rowitch DH, Hurles ME, Geschwind DH, Børglum AD, Robinson EB, Grove J, Martin HC, Bourgeron T, Baron-Cohen S. Genetic correlates of phenotypic heterogeneity in autism. *Nat Genet*. 2022. <https://doi.org/10.1038/s41588-022-01072-5>.
34. Magnus P, Irgens LM, Haug K, Nystad W, Skjærven R, Stoltenberg C, The Moba Study Group. Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol*. 2006;35(5):1146–1150. <https://doi.org/10.1093/ije/dyl170>.
35. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, Handal M, Haugen M, Høiseth G, Knudsen GP, Paltiel L, Schreuder P, Tambs K, Vold L, Stoltenberg C. Cohort profile update: the Norwegian mother and child cohort study (MoBa). *Int J Epidemiol*. 2016;45(2):382–8. <https://doi.org/10.1093/ije/dyw029>.
36. Corfield EC, Frei O, Shadrin AA, Rahman Z, Lin A, Athanasios L, Akdeniz BC, Hannigan L, Wootton RE, Austerberry C, Hughes A, Tesli M, Westlye LT, Stefánsson H, Stefánsson K, Njølstad PR, Magnus P, Davies NM, Appadurai V, et al. The Norwegian Mother, Father, and Child cohort study (MoBa) genotyping data resource: MoBaPsychGen pipeline v.1. *bioRxiv*. 2022. <https://doi.org/10.1101/2022.06.23.496289>.
37. Rutter M, Lord C, Bailey A. SCQ the social communication questionnaire: manual. Los Angeles: Western Psychological Services; 2003.
38. Squires J, Potter L, Briker D. The ASQ User's Guide. 2nd ed. Baltimore: Paul H. Brookes Publishing Co.; 1999.
39. Schjølberg S. Early identification of autism spectrum disorders. Gøteborg: The Social Brain; 2003.
40. Robins DL, Fein D, Barton ML, Green JA. The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *J Autism Dev Disord*. 2001;31(2):131–44. <https://doi.org/10.1023/a:1010738829569>.
41. Swinkels SHN, Dietz C, van Daalen E, Kerkhof IHGM, van Engeland H, Buitelaar JK. Screening for autistic spectrum in children aged 14 to 15 months. I: the development of the Early Screening of Autistic Traits Questionnaire (ESAT). *J Autism Dev Disord*. 2006;36(6):723–32. <https://doi.org/10.1007/s10803-006-0115-0>.
42. Achenbach TM. Manual for child behavior checklist. Wuhan: Scientific Research Publishing; 1992.
43. Goodman R. The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry*. 1997;38(5):581–6. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>.

44. Hannigan LJ, Corfield EC, Askelund AD, Askeland RB, Hegemann L, Jensen P, Pettersen JH, Rayner C, Ayorech Z, Bakken NR, Wootton RE, Havdahl A. phenotools: An R package to facilitate efficient and reproducible use of phenotypic data from MoBa and linked registry sources in the TSD environment. 2023. <https://doi.org/10.17605/OSF.IO/6G8BJ>
45. Surén P, Saasen-Havdahl A, Bresnahan M, Hirtz D, Hornig M, Lord C, Reichborn-Kjennerud T, Schjølberg S, Øyen A-S, Magnus P, Susser E, Lipkin WI, Stoltenberg C. Sensitivity and specificity of early screening for autism. *BJPsych Open*. 2019;5(3): e41. <https://doi.org/10.1192/bjo.2019.34>.
46. Surén P, Havdahl A, Øyen A-S, Schjølberg S, Reichborn-Kjennerud T, Magnus P, Bakken IJL, Stoltenberg C. Diagnosing autism spectrum disorder among children in Norway. *J Norwegian Med Assoc*. 2019. <https://doi.org/10.4045/tidsskr.18.0960>.
47. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. *GigaScience*. 2019;8(7):giz082. <https://doi.org/10.1093/gigascience/giz082>.
48. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, Bryois J, Chen C-Y, Dennison CA, Hall LS, Lam M, Watanabe K, Frei O, Ge T, Harwood JC, Koopmans F, Magnusson S, Richards AL, Sidorenko J, et al. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*. 2022;604(7906):502–508. <https://doi.org/10.1038/s41586-022-04434-5>
49. Insel TR. Rethinking schizophrenia. *Nature*. 2010;468(7321):187–93. <https://doi.org/10.1038/nature09552>.
50. Morris-Rosendahl DJ, Crocq M-A. Neurodevelopmental disorders—the history and future of a diagnostic concept. *Dialogues Clin Neurosci*. 2020;22(1):65–72. <https://doi.org/10.31887/DCNS.2020.22.1/macrocq>.
51. Coombes BJ, Ploner A, Bergen SE, Biernacka JM. A principal component approach to improve association testing with polygenic risk scores. *Genet Epidemiol*. 2020;44(7):676–86. <https://doi.org/10.1002/gepi.22339>.
52. Wu H, Estabrook R. Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika*. 2016;81(4):1014–45. <https://doi.org/10.1007/s11336-016-9506-0>.
53. Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012. <https://doi.org/10.18637/jss.v048.i02>.
54. Jorgensen TD, Pornprasertmanit S, Schoemann AM, Rosseel Y. semTools: Useful tools for structural equation modeling. 2022. <https://CRAN.R-project.org/package=semTools>
55. Mbatchou J, Barnard L, Backman J, Marcketta A, Kosmicki JA, Ziyatdinov A, Benner C, O'Dushlaine C, Barber M, Boutkov B, Habegger L, Ferreira M, Baras A, Reid J, Abecasis G, Maxwell E, Marchini J. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet*. 2021;53(7):1097–103. <https://doi.org/10.1038/s41588-021-00870-7>.
56. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, Duncan L, Perry JRB, Patterson N, Robinson EB, Daly MJ, Price AL, Neale BM. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236–41. <https://doi.org/10.1038/ng.3406>.
57. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017. <https://doi.org/10.1038/s41467-017-01261-5>.
58. R Core Team. R: A Language and Environment for Statistical Computing [Computer software]. R Foundation for Statistical Computing. 2021. <https://www.R-project.org/>
59. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, Ip HF, Marioni RE, McIntosh AM, Deary IJ, Koellinger PD, Harden KP, Nivard MG, Tucker-Drob EM. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav*. 2019;3(5):513–25. <https://doi.org/10.1038/s41562-019-0566-x>.
60. Gillberg C. The ESSENCE in child psychiatry: early symptomatic syndromes eliciting neurodevelopmental clinical examinations. *Res Dev Disabil*. 2010;31(6):1543–51. <https://doi.org/10.1016/j.ridd.2010.06.002>.
61. Kaplan BJ, Dewey DM, Crawford SG, Wilson BN. The term comorbidity is of questionable value in reference to developmental disorders: data and theory. *J Learn Disabil*. 2001;34(6):555–65. <https://doi.org/10.1177/002221940103400608>.
62. Marvin AR, Marvin DJ, Lipkin PH, Law JK. Analysis of Social Communication Questionnaire (SCQ) screening for children less than age 4. *Curr Dev Disord Rep*. 2017;4(4):137–44. <https://doi.org/10.1007/s40474-017-0122-1>.
63. Riglin L, Tobarra-Sanchez E, Stergiakouli E, Havdahl A, Tilling K, O'Donovan M, Nigg J, Langley K, Thapar A. Early manifestations of genetic liability for ADHD, autism and schizophrenia at ages 18 and 24 months. *JCPP Adv*. 2022;2(3):e12093. <https://doi.org/10.1002/jcv.212093>.
64. Schlag F, Allegrini AG, Buitelaar J, Verhoef E, van Donkelaar M, Plomin R, Rimfeld K, Fisher SE, St Pourcain B. Polygenic risk for mental disorder reveals distinct association profiles across social behaviour in the general population. *Mol Psychiatry*. 2022;27(3):1588–98. <https://doi.org/10.1038/s41380-021-01419-0>.
65. Ronald A, de Bode N, Polderman TJC. Systematic review: How the attention-deficit/hyperactivity disorder polygenic risk score adds to our understanding of ADHD and associated traits. *J Am Acad Child Adolesc Psychiatry*. 2021;60(10):1234–77. <https://doi.org/10.1016/j.jaac.2021.01.019>.
66. Mostafavi H, Harpak A, Agarwal I, Conley D, Pritchard JK, Przeworski M. Variable prediction accuracy of polygenic scores within an ancestry group. *Elife*. 2020;9:e48376. <https://doi.org/10.7554/eLife.48376>.
67. St Pourcain B, Robinson EB, Anttila V, Sullivan BB, Maller J, Golding J, Skuse D, Ring S, Evans DM, Zammit S, Fisher SE, Neale BM, Anney RJL, Ripke S, Hollegaard MV, Werge T, iPSYCH-SSI-Broad Autism Group, Ronald A, Grove J, et al. ASD and schizophrenia show distinct developmental profiles in common genetic overlap with population-based social communication difficulties. *Mol Psychiatry*. 2018;23(2):263–270. <https://doi.org/10.1038/mp.2016.198>
68. Akingbuwa WA, Hammerschlag AR, Bartels M, Middeldorp CM. Systematic review: molecular studies of common genetic variation in child and adolescent psychiatric disorders. *J Am Acad Child Adolesc Psychiatry*. 2022;61(2):227–42. <https://doi.org/10.1016/j.jaac.2021.03.020>.
69. Nagel M, Watanabe K, Stringer S, Posthuma D, van der Sluis S. Item-level analyses reveal genetic heterogeneity in neuroticism. *Nat Commun*. 2018. <https://doi.org/10.1038/s41467-018-03242-8>.
70. Mallard TT, Tubbs JD, Jennings M, Zhang Y, Gustavson DE, Grotzinger AD, Westwater ML, Williams CM, Fortgang RG, 23andMe Research Team, Elson SL, Fontanillas P, Davis LK, Raznahan A, Tucker-Drob EM, Choi KW, Ge T, Smoller JW, Palmer AA, Sanchez-Roige S. The pleiotropic architecture of human impulsivity across biological scales. *medRxiv*. 2023. <https://doi.org/10.1101/2023.11.28.23299133>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.