Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genetic Architecture of Educational Attainment is Represented by Executive Function in Children

Joanne S. M. Chia
Agency for Science, Technology and Research (A*STAR)
Evelyn C. Law
Agency for Science, Technology and Research (A*STAR)
Chong Yap Seng
Agency for Science, Technology and Research (A*STAR)
Peter D. Gluckman
Agency for Science, Technology and Research (A*STAR)
Johan G. Eriksson
Agency for Science, Technology and Research (A*STAR)
Lourdes M. Daniel
KK Women's and Children's Hospital
Patrícia P. Silveira
McGill University
Michael J. Meaney
McGill University
Michelle Z. L. Kee
Michelle_kee@a-star.edu.sg

Agency for Science, Technology and Research (A*STAR)

Article

Keywords: Educational Status, Child Development, Executive Function, Genetics, Polygenic Score

Posted Date: March 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8863842/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

The genetics of educational attainment is well characterized in adults, but little is known about how it operates during childhood, when cognitive development underlying academic achievement is developing. In this study, we investigated the associations of the genetic architecture of educational attainment with childhood cognitive function. We used summary statistics from a published genome-wide association study for educational attainment and its cognitive and non-cognitive components to create polygenic scores ($PGS_{EA,}$ $PGS_{Cog,}$ PGS_{nonCog}) in a multi-ethnic, longitudinal, and prospective birth cohort in Singapore. Validation with the adults in our sample revealed that the PRSs were significantly associated with university completion rates, explaining up to 6.3% of the variance (*n* = 1,688). Additionally, PGS_{EA} was notably lower in children who required additional learning support at age 7 (*n* = 887). Children's PGS_{Cog} and PGS_{nonCog} were significantly associated with cognitive and emotional regulation of executive function respectively (*n* = 110). Genes derived from PGS_{Cog} were largely brain-specific and related to cognition and psychiatric disorders. In contrast, genes from PGS_{nonCog} were less represented in the brain and linked to risk-taking and broader health outcomes. Our findings suggest that the genetics of educational attainment associate with childhood cognitive-emotional phenotypes that strongly influence academic performance, mirroring previous findings in adults.

INTRODUCTION

Educational attainment has long been examined in adulthood both as a function of socio-economic status and as a robust predictor of health outcomes[1,2]. Beyond its social determinants, educational attainment is also a heritable trait that is highly polygenic and represents more than intelligence[3]. As such, genetic approaches offer a powerful avenue to uncover the biological mechanisms underlying cognitive development that are beneficial to later success in life[4]. A polygenic score (PGS) derived from a genome-wide association study (GWAS) for educational attainment with 1.1 million participants of European ancestry accounted for 11–13% of phenotypic variance[5], revealing an increase in predictive power compared to earlier studies. Notable, the PGS from a meta-analysis of cognitive performance represented 7–10% of the variance in educational attainment, suggesting contribution from an additional, yet meaningful non-cognitive component. Demange et al. (2021)[6] addressed this issue using a "GWAS-by-subtraction" method to isolate the non-cognitive genetic construct, removing genetic variants related to cognitive ability from the genetic educational attainment construct. The cognitive construct had a higher genetic correlation with academic performance, indicated by self-reported math ability and childhood IQ. Conversely, the non-cognitive construct correlated with greater general risk-taking, lower health-risk behaviors such as smoking, and personality traits such as agreeableness, conscientiousness, extraversion, and openness.

Childhood is a dynamic period for developing cognitive skills required for educational attainment. Key among these cognitive skills is executive functions, which are higher-order processes guiding goal-directed behaviors[7,8]. Executive functions emerge primarily from 3–7 years of age and continue into late

childhood, adolescence, and even adulthood[9,10]. There are three core executive functions: (1) inhibition, the ability to restrain prepotent responses; (2) cognitive flexibility, the capacity to adapt in response to changing situations; and (3) working memory, the ability to hold and manipulate information. These skills support abilities such as planning and organizing, or self-monitoring of emotions and behavior. Executive functions have greater importance than intelligence (such as IQ metrics) for school readiness[11]. Numerous studies indicate the involvement of various facets of executive functions on academic outcomes, with attention supporting literacy and working memory underpinning math abilities[12–15].

The existing GWASs for educational attainment provide compelling evidence for the influence of heritable genetic variations on educational attainment[3,5].These GWASs focus on adults, as educational attainment can only be fully assessed when educational history is complete. However, since the outcome of educational attainment is closely linked to performance in earlier phases of schooling, it is critical to understand how the genetic architecture of educational attainment is represented in school-aged children. We addressed this question by examining measures of executive function using the teacher-reported Behavioral Rating Inventory of Executive Function 2 (T-BRIEF2). Designed to assess executive functions in school-aged children within the school environment[16], the T-BRIEF2 is ideally positioned for studies of skills relevant to educational attainment.

The T-BRIEF2 comprises three indices: the behavioral regulation index (BRI), which measures behavioral inhibition and self-monitoring; the cognitive regulation index (CRI), which includes items such as working memory and task monitoring that are essential in completing daily tasks such as schoolwork; and the emotional regulation index (ERI) which measures the ability to adapt to changing environments and to regulate and manage one's emotions. Some subscales, such as Working Memory, reflect core constructs measured through performance-based tasks. In contrast, other subscales, such as Organization of Materials and Tasks Monitoring, do not have a direct performance-task equivalent, making them valuable for detecting difficulties in achieving daily tasks that require executive functions[17].

In this study, we used European population-based GWAS summary statistics for educational attainment[5] and the derived cognitive and non-cognitive components[6] to create PGSs for educational attainment (PGS_{EA}), cognitive (PGS_{Cog}), and non-cognitive components (PGS_{nonCog}) in a longitudinal, multi-ethnic Asian birth cohort study, the Growing Up in Singapore Towards healthy Outcomes (GUSTO). We aim to: (1) evaluate the validity and predictive utility of European-derived PGSs in an adult Asian population; (2) assess the relevance of child PGSs in educational settings by examining if the PGSs are distinct between children enrolled in learning support programs and their peers; (3) examine how the cognitive and non-cognitive genetic components of educational attainment relate to childhood executive functions; and (4) explore the biological processes through which these genetic influences may shape early child executive function abilities. We hypothesized that all three PGSs would predict adult educational attainment in our Asian cohort and distinguish children receiving learning support from those who are not. We also hypothesized that the PGS_{EA} would associate with childhood executive functions, with PGS_{Cog} more

strongly related with cognitive regulation, and PGS_{nonCog} aligning with the cognitive regulation and emotional/behavioral regulation domains, respectively.

METHODS

Participants and measures

Healthy pregnant women aged ≥ 18 years were recruited into the GUSTO longitudinal cohort study from two public hospital maternity units in Singapore between 2009 and 2010[18]. All women and their spouses were of homogenous parental Chinese, Indian, or Malay ethnicity. Pregnant women who were receiving chemotherapy, had chronic diseases or were taking psychotropic drugs, decided to deliver outside the study hospital or leave Singapore within the next 5 years, were past 14 weeks of gestation, had a miscarriage, planned to terminate the pregnancy, or had in vitro fertilization for multiple pregnancies were excluded from the study. Mothers self-reported their age, ethnicity, monthly household income, and highest education level attained upon recruitment. Fathers reported their highest education level attained when the child turned 2–3. Parental education levels were categorized as: (1) up to secondary school education, (2) pre-tertiary education, and (3) university degree and above.

Upon entering primary school (approximately 7 years old), students were screened for English and Mathematics competency. Those requiring additional learning support were enrolled in centrally developed early intervention programs for literacy and/or numeracy. Children (n = 887) were categorized into those who did or did not require learning support. These data were obtained from the Ministry of Education in Singapore with parental consent. As a follow-up measure of academic achievement, self-reported Primary School Leaving Examination scores (n = 469) were used as an outcome measure. This is a nationwide standardized examination which children in Singapore take at 12 years old, where lower scores means better academic achievement.

Teachers completed the T-BRIEF2[16] to rate the behavioral features of executive functions in real-world settings (i.e., educational environments) when the child was 8.5 years old (n = 110). The T-BRIEF2 consists of 63 items in nine scales, which are combined to form three composite indices: the BRI comprised the subscales Inhibit and Self-Monitor; the ERI comprised the subscales Shift and Emotional Control; and the CRI comprised the subscales Initiate, Working Memory, Plan/Organize, Task Monitor and Organization of Materials. Together, these three indices form the Global Executive Composite (GEC). The T-BRIEF2 T-scores were used for analyses, with higher scores indicating poorer executive functioning.

Written informed consent was obtained from all participants, with mothers consenting on behalf of their children. The study was approved by the National Health Care Group Domain Specific Review Board (D/09/02) and SingHealth Centralized Institutional Review Board (2009/280/D). All research was conducted in accordance with the relevant institutional, national, and international guidelines and regulations, including the Declaration of Helsinki.

A flow chart representing the number of adults and children participants from the GUSTO cohort, along with both genetic and academic measures used for analyses, is presented in Supplementary Figure S1.

Genotyping

The GUSTO cohort was genotyped using the Infinium OmniExpressExome array and split by ethnicity for quality checks. Single nucleotide polymorphisms (SNPs) with call rates < 95%, minor allele frequencies (MAF) < 0.05 and $p \leq 10^{-6}$ for Hardy-Weinberg equilibrium (HWE) were removed. Allele frequencies were compared with the 1000G reference panels[19] (East Asian panel for Chinese and Malay; South Asian panel for Indian). SNPs with allele frequency differences > 0.2 for Chinese and Indian, and > 0.3 for Malay were removed. Data were pre-phased using SHAPEIT v2.837[20] with family trio information, then imputed with 1000G Phase 3 as the reference panel and choosing "with PBWT, no pre-phasing" pipeline using the Sanger Imputation Service (https://www.sanger.ac.uk/tool/sanger-imputation-service/). Imputed data that were non-monomorphic (MAF > 0), biallelic SNPs with HWE $p > 10^{-6}$, INFO score > 0.80, MAF > 0.05, and were common across all three ethnicities were retained.

Linkage disequilibrium pruning was applied to remove SNPs with r^2 <0.2 within a window size of 50 SNPs at each step size of 5 using PLINK v1.9 (https://www.cog-genomics.org/plink/1.9). Principal component analysis (PCA) was conducted on the pruned data set using EIGENSTRAT[21] (https://github.com/DReichLab/EIG/archive/v7.2.1.tar.gz) to assess population structure. The first three principal components (PCs) were considered the most informative of the population structure in the GUSTO cohort and included in all analyses.

Polygenic scores (PGSs)

PGSs[22] for parents and children in the GUSTO cohort were calculated using the GWAS summary statistics for educational attainment[5] and its cognitive and non-cognitive components [6]. We applied both LDpred2[23] (see Supplementary Information) and PRSice v2.3.1.e (PRSice2[24]). LDpred2, a Bayesian approach that includes information from linkage disequilibrium when approximating effect sizes, offers improved accuracy compared to the clumping and thresholding approach used in PRSice2. The LDpred2 auto model was used, which estimated parameters such as heritability and proportion of SNPs directly from the data without requiring a validation sample. For PRSice2, variants with MAF < 0.05 were excluded and default clumping parameters were used. The PGSs were calculated at 19 *p*-value thresholds (pT), from 0.01 to 0.1 in increments of 0.01, and from 0.1 to 1 in increments of 0.1, and standardized as z-scores. The total number of SNPs used to calculate PGS_{EA}, PGS_{Cog} and PGS_{nonCog} were 91,311, 86,851 and 86,698, respectively (Supplementary Figures S2-S3).

Statistical analyses

Polygenic prediction of educational attainment. We applied logistic regression to examine if the PGSs predicted adult educational attainment in our Asian population, by using GUSTO parents' university-completion status as an outcome. We selected participants above the minimum age for university completion in Singapore, specifically > 23 years for mothers (n = 1,005) and > 25 years for fathers (n = 683). We included sex and the first 3 PCs of the maternal and paternal genetics to control for population stratification. The PGS with the highest incremental R^2, calculated as the difference between the Nagelkerke R^2 of the full model compared to the model with only the covariates, was used in subsequent analyses (Supplementary Figure S4). One-way ANOVA with Tukey post-hoc test was used to examine the mean PGS difference between the GUSTO parents in the three groups of education levels attained, i.e.: university (n = 370 mothers; n = 280 fathers), pre-tertiary (n = 354 mothers; n = 245 fathers) and secondary and lower (n = 281 mothers; n = 158 fathers).

To assess the use of the PGSs with children, we used the learning support status of the GUSTO child as a proxy for academic ability to model educational attainment. Children were divided into those who enrolled in learning support programs in English and/or Mathematics at 7 years of age (n = 98) and those who did not (n = 789). The child PGSs used for analyses were at the same pT as the parental PGSs. A two-sided Welch two-sample t-test was used to examine the mean difference in scores of those requiring learning support and those who did not.

Polygenic prediction of child executive functions. Data from the four indices (GEC, CRI, ERI, and BRI) of the T-BRIEF2 scales were regressed onto the GUSTO children's PGS_{EA}, PGS_{Cog} and PGS_{nonCog}. The first three child genetic PCs and child sex were included as covariates. The PGSs with the largest variance explained (R^2) across T-BRIEF2 indices were selected for further analyses (Supplementary Figures S5-S7). Regression analyses using the T-BRIEF2 subscales were restricted to the indices that exhibited a significant association with their respective PGS. Two-tailed significance tests were applied, with $p < .05$ set as the threshold for statistical significance. Multiple testing corrections were done using the Benjamini-Hochberg method[25] on the indices and subscales separately. All regression analyses were done in R v4.2.1[26].

Biological process network, transcription factor and tissue specificity analyses.

All SNPs (pT = 1) for each PGS were mapped onto unique genes (PGS_{EA}: 91,311 SNPs to 10,219 genes; PGS_{Cog}: 86,851 SNPs to 9,935 genes; PGS_{nonCog}: 86,698 SNPs to 9,917 genes) using biomaRt v2.40.4[27,28] and used as references for the rest of the bioinformatic analyses. The SNPs of PGS_{EA}, PGS_{Cog} and PGS_{nonCog} at pTs that accounted for the greatest variance in T-BRIEF2 indices were mapped onto genes (PGS_{EA} pT = 0.03: 25,284 SNPs to 5,540 genes; PGS_{Cog} pT = 0.03: 19,934 SNPs to 4,703 genes; PGS_{nonCog} pT = 0.3: 56,138 SNPs to 8,294 genes). These gene sets were compared to the reference genes and analyzed for process network and transcription factor enrichment using MetaCore™ (v23.1.71200, Clarivate Analytics). P-values for transcription factor analyses were determined using hypergeometric intersection. We used FUMA v1.3.7[29] to examine whether the gene sets overlapped with

other GWAS traits and to investigate the up-regulated expression of these genes compared to tissue-specific gene expression data from the publicly available Genotype-Tissue Expression project (GTex v8 RNA-seq data) using Bonferroni-corrected *p*-values[30]. A control gene set was also created for each PGS, by randomly selecting an equivalent number of genes at the most significant PGS pT from the full set of GWAS genes. Bioinformatic analyses were repeated with these control gene sets.

RESULTS

Study characteristics

The GUSTO parental demographic characteristics are summarized in Table 1. The parents were on average 32.7 years of age at recruitment, were 59.5% mothers, comprised of 59.5% Chinese, 16.5% Indian, and 24.1% Malay ethnicity, and 38.5% university completion. 10.9% of the cohort fell below the nominal poverty level (< SGD$2,000 a month per family), which maps onto the population. We note that because of recruitment limited to public birthing hospitals and the inclusion criteria, the GUSTO sample should not be considered as representative of the Singapore population.

Table 1
Study characteristics of parents

Characteristic	N = 1,688[1]
Sex	
Female	1,005 (59.5%)
Male	683 (40.5%)
Missing information	0
Ethnicity	
Chinese	1,003 (59.5%)
Indian	278 (16.5%)
Malay	406 (24.1%)
Missing information	1
Age at recruitment	32.7 (5.2)
Missing information	0
Household income (SGD)	
< $2000	170 (10.9%)
$2000 - $3999	429 (27.4%)
$4000 - $5999	408 (26.1%)
> $6000	557 (35.6%)
Missing information	124
Education level	
Secondary and lower	439 (26.0%)
Pre-tertiary	599 (35.5%)
Tertiary	650 (38.5%)
Missing information	0
[1]n (%); Mean (SD)	

PGS PGS and PGS predict education levels in mothers and fathers

The GWASs for educational attainment, cognitive and non-cognitive abilities were based on European populations[5,6].We used the weighted sum of genetic variants from these GWASs (Supplementary

Figures S2-S3) to create PGSs in our Asian cohort. As a validation, we regressed the university completion status onto PGS_{EA}, PGS_{Cog} and PGS_{nonCog} of the GUSTO parents. The variance explained in adult educational attainment was better with LDpred2-derived PGS_{EA} and PGS_{Cog} compared with PRSice2 (PGS_{EA}: 10.8% with LDpred2 vs 6.1% with PRSice2; PGS_{Cog}: 3.8% with LDpred2 vs 1.7% for PRSice2) but was similar in both methods for PGS_{nonCog} (6.1% with LDpred2 vs 6.3% with PRSice2; Supplementary Table S1). Overall, PGS_{EA} and PGS_{nonCog} explained university completion better than PGS_{Cog}. To understand the biological pathways associated with educational attainment, we continued with PRSice2's clumping and thresholding method, selecting the PGSs with the highest incremental R^2 for subsequent analyses (i.e. PGS_{EA} and PGS_{nonCog}: pT = 0.01; PGS_{nonCog}: pT = 0.02; Supplementary Figure S4).

We next assessed the percentage of university completion in PGS quintiles (Figs. 1a, d, g). As expected, higher PGS_{EA}, PGS_{Cog} and PGS_{nonCog} predicted a higher percentage of university completion. A one-way ANOVA revealed significant differences in mean PGSs among adults categorized into different educational attainment levels (PGS_{EA}: $F(2) = 40.6$, $p < .001$, Fig. 1b; PGS_{Cog}: $F(2) = 12.1$, $p < .001$, Fig. 1e; PGS_{nonCog}: $F(2) = 44.4$, $p < .001$, Fig. 1h). Post-hoc multiple comparisons tests showed a significant difference in PGSs between the university and pre-tertiary levels for PGS_{EA} ($p < .001$, 95% CI [0.26, 0.51]; Fig. 1c), PGS_{Cog} ($p = .002$, 95% CI [0.06, 0.32]; Fig. 1f) and PGS_{nonCog} ($p < .001$, 95% CI [0.26, 0.52]; Fig. 1i). PGSs were also significantly different between adults who obtained university education and those with secondary and lower levels of education in PGS_{EA} ($p < .001$, 95% CI [0.34, 0.62]; Fig. 1c), PGS_{Cog} ($p < .001$, 95% CI [0.14, 0.43]; Fig. 1f) and PGS_{nonCog} ($p < .001$, 95% CI [0.38, 0.67]; Fig. 1i). However, no significant differences were observed between pre-tertiary and secondary and lower levels for all three PGSs. These results suggest that the PGS_{EA}, PGS_{Cog} and PGS_{nonCog} distinguish university education from lower levels of education in our adult Asian sample.

PGS, PGS and PGS predict child academic performance

We investigated if PGSs that best predicted the GUSTO parental educational attainment would predict the GUSTO children in learning support programs at age 7. The lowest tertile of all child PGSs had the highest percentage of children who required learning support (Figs. 2a, g, d). The largest percentage difference between the highest and lowest tertiles was observed in PGS_{nonCog} (9.6%, Fig. 2g), followed by PGS_{EA} (8.4%, Fig. 2a), and PGS_{Cog} (6.4%, Fig. 2d).

Children in learning support programs had lower mean PGSs than their peers, with significant differences for PGS_{EA} ($t(115) = 4.32$, $p < .001$, 95% CI [0.29, 0.77]; Figs. 2b, c), PGS_{Cog} ($t(119) = 2.32$, $p = .02$, 95% CI [0.04, 0.48]; Figs. 2e, f), and PGS_{nonCog} ($t(122) = 4.66$, $p < .001$, 95% CI [0.29, 0.71]; Figs. 2h, i). Interestingly, we observed a wide inter-individual variation of PGS_{EA}, PGS_{Cog} and PGS_{nonCog} (Figs. 2j, k, l). As a follow-up, we associated the PGSs with self-reported scores of a nationwide standardized examination administered at age 12 (Supplementary Figure S8). Children in different tertiles of PGSs had

significantly different mean examination scores (PGS_{EA}: $F(2) = 75.0$, $p < .001$; PGS_{Cog}: $F(2) = 23.9$, $p < .001$; PGS_{nonCog}: $F(2) = 84.3$, $p < .001$). The difference between the highest and lowest tertile was greater with PGS_{EA} and PGS_{nonCog} compared with PGS_{Cog} (PGS_{EA}: $p < .001$, 95% CI [-5.00, -1.72]; PGS_{Cog}: $p = .012$, 95% CI [-3.79, -0.37]; PGS_{nonCog}: $p = .005$, 95% CI [-4.04, -0.60]). Overall, these analyses with child PGSs recapitulate the results from the parents that PGS_{EA} and PGS_{nonCog} were better in distinguishing academic capabilities compared to PGS_{Cog}.

PGS_{EA}, PGS_{Cog} and PGS_{nonCog} associate with measures of executive function

Executive functions are critical for academic achievement[11,31,32]. We used the PGS_{EA}, PGS_{Cog} and PGS_{nonCog} to predict child executive functioning in the 'real world' classroom setting using the four main indices of T-BRIEF2 (i.e. CRI, ERI, BRI, and the GEC) when the child was 8.5 years old. We selected the PGSs that explained the most variance in T-BRIEF2 indices for subsequent analyses, specifically PGS_{EA} and PGS_{Cog} at pT = 0.03, and PGS_{nonCog} at pT = 0.3 (Supplementary Figures S5-S7).

After Benjamini-Hochberg multiple testing corrections, PGS_{EA} was significantly associated with the CRI, GEC, and ERI. PGS_{Cog} was significantly associated with the CRI, GEC, and BRI, In contrast, PGS_{nonCog} was associated only with the ERI, with multiple ERI subscales also significantly associated (Fig. 3, see Supplementary Table S2 for more details). Higher PGS_{EA} was associated not only with better cognitive regulation in working memory, task monitoring, and information organizational skills, but also with a greater ability to consciously modulate an emotional state. Higher PGS_{Cog} expectedly associated with better cognitive regulation across all subscales, and fewer behavioral monitoring problems. Conversely, higher PGS_{nonCog} associated with increased emotional flexibility in response to changing situations. These results suggest that the separation of the genetic architecture into cognitive and non-cognitive aspects could distinguish between the cognitive and emotional functions as indexed by the T-BRIEF2.

Biological annotation of SNPs in PGSs that associate with executive functions

To investigate the genetic networks underlying cognitive and non-cognitive aspects of educational attainment, we mapped the SNPs from the previously selected PGSs that explained the greatest variance in T-BRIEF2 indices onto genes.

Analyses from the FUMA software[29] revealed that the genes from PGS_{EA} significantly overlapped with those from GWASs for various traits from the database (Supplementary Figure S9). Genetically associated traits could be categorized into chronotype ($p = 1.75 \times 10^{-21}$), cognition (intelligence: $p = 1.01 \times 10^{-19}$; and general cognitive ability: $p = 8.45 \times 10^{-13}$), psychiatric diseases (schizophrenia: $p = 1.44 \times 10^{-15}$; attention-deficit/hyperactive disorder $p = 2.42 \times 10^{-5}$; bipolar disorder $p = 2.46 \times 10^{-6}$; and major

depressive disorder $p = 2.53 \times 10^{-5}$), risk-taking behaviors (adventurousness: $p = 3.85 \times 10^{-6}$; and general risk tolerance: $p = 1.50 \times 10^{-11}$), and metabolic health (BMI: $p = 9.88 \times 10^{-8}$; Type II diabetes: $p = 7.21 \times 10^{-8}$; and waist circumference: $p = 4.72 \times 10^{-5}$). Genes from PGS_{Cog} showed similar overlap, with the strongest associations with intelligence ($p = 1.51 \times 10^{-31}$) and general cognitive ability ($p = 1.05 \times 10^{-19}$). Conversely the genes from PGS_{nonCog} overlapped with chronotype ($p = 1.23 \times 10^{-4}$), systemic traits such as systolic blood pressure and pulse pressure ($p = 2.28 \times 10^{-3}$ and 4.34×10^{-3} respectively), general risk tolerance ($p = 4.34 \times 10^{-3}$), smoking status (ever vs never smokers; $p = 7.13 \times 10^{-3}$), cognitive ability ($p = .020$) and intelligence ($p = .045$). Our results are consistent with a similar analysis using genomic correlations from selected traits[6] with the addition of health and cognitive outcomes.

Tissue-specific expression analysis using FUMA showed that all PGS gene sets were significantly upregulated in brain tissue (Supplementary Figure S10). Genes from PGS_{EA} and PGS_{Cog} were largely upregulated in the frontal cortex, anterior cingulate cortex, hippocampus, amygdala, and nucleus accumbens (all Bonferroni-corrected $p < .001$). Conversely, genes from PGS_{nonCog} were only significantly upregulated in the anterior cingulate cortex, hippocampus, and putamen (Bonferroni-corrected *p*s = .008, .010 and .049 respectively). This brain-specific expression pattern was not observed in control analyses using randomly selected genes (Supplementary Figure S11).

We identified the biological process networks represented by the gene sets related to executive function. Significant process networks in both PGS_{EA} and PGS_{Cog} that passed multiple testing corrections (Supplementary Tables S3-S4) were related to neuronal development, such as cell adhesion in synaptic contact (PGS_{EA}: $p = 1.73 \times 10^{-8}$; PGS_{Cog}: $p = 7.42 \times 10^{-4}$), and cadherins (PGS_{EA}: $p = 9.14 \times 10^{-5}$; PGS_{Cog}: $p = 3.57 \times 10^{-3}$), synaptogenesis (PGS_{EA}: $p = 1.33 \times 10^{-4}$; PGS_{Cog}: $p = 3.57 \times 10^{-3}$) and axonal guidance (PGS_{EA}: $p = 1.33 \times 10^{-4}$; PGS_{Cog}: $p = 3.57 \times 10^{-3}$). Cell adhesion for synaptic contact ($p = .03$) was the only significant process network for the genes in PGS_{nonCog} (Supplementary Table S5). These associations were all non-significant in a control gene set (Supplementary Tables S6-S8).

Transcription factor enrichment analyses revealed that all three PGSs had common enriched transcription factors (Supplementary Figures S12 and Supplementary Tables S9-S11) that were not observed in controls (Supplementary Tables S12-S14). The most prominent transcription factors were the Wnt signaling pathway gene *TCF7L2*, the estrogen-related receptor *ESSRG*, and a ubiquitous transcriptional enhancer *TEAD1*. Notably, the glucocorticoid receptor *NR3C1* was highly enriched in both PGS_{EA} and PGS_{nonCog} but not PGS_{Cog}. Interestingly, the circadian regulator *BMAL1* was enriched in PGS_{EA}, PGS_{Cog} and PGS_{nonCog}.

Discussion

We used published GWASs to create PGSs for educational attainment, together with the cognitive and non-cognitive components of educational attainment. The PGS_{EA}, although derived from a GWAS study

of Caucasians, was significantly associated with educational attainment in an adult Asian population of mixed East and South Asian ancestry. We then examined the associations of the PGS in childhood, a period of life when the origin of academic success is established. The PGS_{EA} differentiated children with limited English literacy/numeracy skills at entry to primary school from those with proficient early academic skills. These findings provide validation for the use of the PGS_{EA} in the research analyses of biological pathways underlying academic achievement in the GUSTO children, as distinct from a tool for screening or intervention.

Our primary aim was to examine the potential influence of the genetic architecture of educational attainment on cognitive-emotional skills critical for success in early schooling, focusing on executive function performance in the classroom. There is extensive evidence for the importance of executive functions for academic achievement[11,31,32] in primary school. Accordingly, children in the learning support group had elevated teacher-reported BRIEF2 scores, indicating worse cognitive and emotional regulation, as well as general executive function. PGS_{EA}, PGS_{Cog} and PGS_{nonCog} were lower in the learning support group compared to the non-learning support group, indicating that these measures are relevant for early academic achievement. While our results show a distinction in PGSs between those in learning support versus those who are not, the wide inter-individual variation in the PGSs also suggests that early learning is subject to various influences beyond genetics and may thus reflect the capacity for intervention to facilitate better learning and therefore lead to attaining higher education levels.

PGS_{EA} and PGS_{Cog} were significantly associated with the cognitive regulation in children, where higher PGSs associated with better working memory, task monitoring and materials organization. PGS_{Cog} was additionally associated with the initiation of planning, organizing information, and executing tasks. In contrast, higher PGS_{EA} and PGS_{nonCog} were associated with better emotional flexibility. In sum, these findings suggest that the genetic architecture of educational attainment defined in adult populations is associated with both 'hot' (or affective) and 'cold' (or cognitive) forms of executive functions essential for success in primary school. For example, in our study, PGS_{Cog} of the child was associated with working memory and self-monitoring, which are known to underlie mathematical ability[11,15]. This finding is consistent with other studies where PGS_{Cog} better predicts mathematical ability compared to PGS_{nonCog}[6]. Interestingly, PGS_{nonCog} better predicted both adult and child academic outcomes. Our results therefore suggest that PGS_{Cog} and PGS_{nonCog} capture complementary pathways, where PGS_{Cog} reflects cognitive skills that support academic domains, while PGS_{nonCog} reflects emotional regulatory skills that promote general academic success.

Twin studies suggest a genetic contribution to inter-individual variation of executive function[33]. However, these findings do not inform on potential biological mechanisms. Our bioinformatic analyses identified pathways related to executive functions and educational attainment. Interestingly, the genes significantly associated with executive functions from PGS_{EA} and PGS_{nonCog} significantly overlapped with those found in chronotype GWAS. Additionally, the transcription factor BMAL1, a key regulator of circadian

rhythms[34], featured prominently within the gene sets. The influence of chronotype on academic achievement in primary school was reported to be mediated by conscientiousness[35], a personality trait genetically correlated with PGS_{nonCog}[6]. The gene set from PGS_{Cog} was also related to psychiatric disorders, where polygenic risk scores for psychiatric disorders have been reported to predict performance in executive function tests[36]. Also, one of the prominent transcription factors in the gene set for PGS_{nonCog} was the glucocorticoid receptor NR3C1, an important moderator of the stress response[37].

The gene sets derived from PGS_{EA} and PGS_{Cog} were highly expressed in the brain, particularly the prefrontal cortex and anterior cingulate cortex, regions prominently involved in executive function and cognitive control[38]. Furthermore, the bioinformatic analyses emphasized neurodevelopmental processes, implying that the genetic architecture of educational attainment might directly influence developmental processes relevant to educational attainment.

A limitation of the study is that the analyses do not include task-based measures commonly associated with executive functioning[13]. The teacher-reported rating in the T-BRIEF2 assesses executive functioning in a classroom setting, which, although seemingly relevant for educational attainment, may not necessarily be directly associated with performance-based measures[39,40]. Compared to performance tasks assessing cognitive efficiency under supervision, the ratings represent the achievement of daily goals in an unsupervised environment[17], which may be more relevant to our genetic scores for educational attainment. In a study conducted in low-income environments, there were high correlations between the cognitive aspects of the T-BRIEF2 and achievement test scores, whereas performance-based measures did not show this association[15]. Also, an executive function rating scale was found to better predict favorable life outcomes, such as social interactions, handling daily responsibilities, and occupational functioning (including current salary and self-rated quality of work), compared to performance tests of executive functioning[41]. Analyses of genotype–outcome associations can be population and context-specific. This study should therefore be considered as exploratory with the requirement for replication across diverse populations, including additional Asian contexts.

In sum, our findings underscore the importance of both cognitive and emotional executive functions in children's academic achievements. We demonstrate that the genetic architecture of educational attainment, when delineated into cognitive and non-cognitive components, is most prominently represented by working memory and emotional flexibility respectively, and offer insights into biological mechanisms such as brain regions, processes, and transcription factors supporting these functions.

Declarations

COMPETING INTERESTS

All authors have reported no biomedical financial interests or potential conflicts of interest.

Author Contribution

JSMC, MJM and MZLK contributed equally to the study conception and design, analyses and interpretation of the data, primary drafting and revision of the manuscript. JSMC prepared all figures and tables. ECL curated the data related to executive function. PPS advised on the genetics analyses and contributed to the draft of the manuscript. CYS, PDG, JGE, LMD and MJM contributed equally to the funding acquisition, investigation of the cohort, and project administration. All authors reviewed the manuscript.

Acknowledgement

The authors thank all children and families who participated in the study, as well as research, administrative, and technical staff who supported this work in every way. The GUSTO study is funded by the Singapore National Research Foundation (NRF) under its Open Fund-Large Collaborative Grant (OF-LCG; MOH-00504) administered by the Singapore Ministry of Health's National Medical Research Council (NMRC), and the Agency for Science Technology and Research (A*STAR), Singapore. In RIE2025, GUSTO is supported by funding from the NRF's Human Health and Potential (HHP) Domain, under the Human Potential Programme. This work is also funded by the Toxic Stress Network of the JPB Foundation (MJM), and the Sackler Foundation (PPS and MJM).

Data Availability

The data used are described in [https://gustodatavault.sg](https:/gustodatavault.sg) . Due to multi-site partnership agreements, the GUSTO data are not deposited into a public repository. However, the GUSTO data are routinely made available upon reasonable request, via submission and approval by the cohort executive committee of a data access form, which can be found in the GUSTO Form Application System (https://fas.sicsapps.com/site/login).

References

1. Conti, G., Heckman, J. & Urzua, S. The Education-Health Gradient. *American Economic Review* 100, 234–238 (2010).
2. Ross, C. E. & Wu, C.-L. The Links Between Education and Health. *Source: American Sociological Review* 60, 719–745 (1995).
3. Krapohl, E. *et al.* The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences* 111, 15273–15278 (2014).

4. Bueno, D. Genetics and Learning: How the Genes Influence Educational Attainment. *Front. Psychol.* 10, (2019).
5. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* 50, 1112–1121 (2018).
6. Demange, P. A. *et al.* Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat. Genet.* 53, 35–44 (2021).
7. Diamond, A. Executive Functions. *Annu. Rev. Psychol.* 64, 135–168 (2013).
8. Miyake, A. *et al.* The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cogn. Psychol.* 41, 49–100 (2000).
9. Ferguson, H. J., Brunsdon, V. E. A. & Bradford, E. E. F. The developmental trajectories of executive function from adolescence to old age. *Sci. Rep.* 11, (2021).
10. Garon, N., Bryson, S. E. & Smith, I. M. Executive Function in Preschoolers: A Review Using an Integrative Framework. *Psychol. Bull.* 134, 31–60 (2008).
11. Blair, C. & Razza, R. P. Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. *Child Dev.* 78, 647–663 (2007).
12. Lan, X., Legare, C. H., Ponitz, C. C., Li, S. & Morrison, F. J. Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *J. Exp. Child Psychol.* 108, 677–692 (2011).
13. McClelland, M. M. *et al.* Links Between Behavioral Regulation and Preschoolers' Literacy, Vocabulary, and Math Skills. *Dev. Psychol.* 43, 947–959 (2007).
14. Sims, D. M. & Lonigan, C. J. Inattention, Hyperactivity, and Emergent Literacy: Different Facets of Inattention Relate Uniquely to Preschoolers' Reading-Related Skills. *Journal of Clinical Child and Adolescent Psychology* 42, 208–219 (2013).
15. Waber, D. P., Gerber, E. B., Turcios, V. Y., Wagner, E. R. & Forbes, P. W. Executive functions and performance on high-stakes testing in children from urban schools. *Dev. Neuropsychol.* 29, 459–477 (2006).
16. Gioia, G. A., Isquith, P. K., Guy, S. C. & Kenworthy, L. Behaviour rating inventory of executive function (2nd ed.). Preprint at (2015).
17. Toplak, M. E., West, R. F. & Stanovich, K. E. Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry and Allied Disciplines* vol. 54 131–143 Preprint at https://doi.org/10.1111/jcpp.12001 (2013).
18. Soh S, Tint M, Gluckman P & Godfrey K. Cohot Profile: Growing Up in Singapore Towards healthy Outcomes (supp). (2014).
19. Auton, A. *et al.* A global reference for human genetic variation. *Nature* vol. 526 68–74 Preprint at https://doi.org/10.1038/nature15393 (2015).

20. Delaneau, O. *et al.* Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat. Commun.* 5, (2014).
21. Price, A. L., Zaitlen, N. A., Reich, D. & Patterson, N. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* 11, 459–463 (2010).
22. Wray, N. R., Goddard, M. E. & Visscher, P. M. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res.* 17, 1520–1528 (2007).
23. Privé, F., Arbel, J. & Vilhjálmsson, B. J. LDpred2: Better, faster, stronger. *Bioinformatics* 36, 5424–5431 (2020).
24. Choi, S. W. & O'Reilly, P. F. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience* 8, 1–6 (2019).
25. Benjamini, Y. & Hochberg, Y. *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing*. *Source: Journal of the Royal Statistical Society. Series B (Methodological)* vol. 57 (1995).
26. R core team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2021).
27. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/ Bioconductor package biomaRt. *Nat. Protoc.* 4, 1184–1191 (2009).
28. Durinck, S. *et al.* BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* 21, 3439–3440 (2005).
29. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* 8, 1–10 (2017).
30. Ardlie, K. G. *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science (1979)*. 348, 648–660 (2015).
31. Clark, C. A. C., Pritchard, V. E. & Woodward, L. J. Preschool Executive Functioning Abilities Predict Early Mathematics Achievement. *Dev. Psychol.* 46, 1176–1191 (2010).
32. Fuhs, M. W., Farran, D. C. & Nesbitt, K. T. Prekindergarten children's executive functioning skills and achievement gains: The utility of direct assessments and teacher ratings. *J. Educ. Psychol.* 107, 207–221 (2015).
33. Friedman, N. P. *et al.* Individual Differences in Executive Functions Are Almost Entirely Genetic in Origin. *J. Exp. Psychol. Gen.* 137, 201–225 (2008).
34. Zheng, Y. *et al.* Neural function of Bmal1: an overview. *Cell and Bioscience* vol. 13 Preprint at https://doi.org/10.1186/s13578-022-00947-8 (2023).
35. Arbabi, T., Vollmer, C., Dörfler, T. & Randler, C. The influence of chronotype and intelligence on academic achievement in primary school is mediated by conscientiousness, midpoint of sleep and motivation. *Chronobiol. Int.* 32, 349–357 (2015).
36. Schork, A. J. *et al.* Polygenic risk for psychiatric disorders correlates with executive function in typical development. *Genes Brain Behav.* 18, (2019).

37. Turecki, G. & Meaney, M. J. Effects of the Social Environment and Stress on Glucocorticoid Receptor Gene Methylation: A Systematic Review. *Biological Psychiatry* vol. 79 87–96 Preprint at https://doi.org/10.1016/j.biopsych.2014.11.022 (2016).
38. Friedman, N. P. & Robbins, T. W. The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology* vol. 47 72–89 Preprint at https://doi.org/10.1038/s41386-021-01132-0 (2022).
39. Mcauley, T., Chen, S., Goos, L., Schachar, R. & Crosbie, J. Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society* 16, 495–505 (2010).
40. Vriezen, E. R. & Pigott, S. E. The relationship between parental report on the BRIEF and performance-based measures of executive function in children with moderate to severe traumatic brain injury. *Child Neuropsychology* 8, 296–303 (2002).
41. Barkley, R. A. & Fischer, M. Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported Executive Function (EF) deficits versus EF tests. *Dev. Neuropsychol.* 36, 137–161 (2011).

Figures

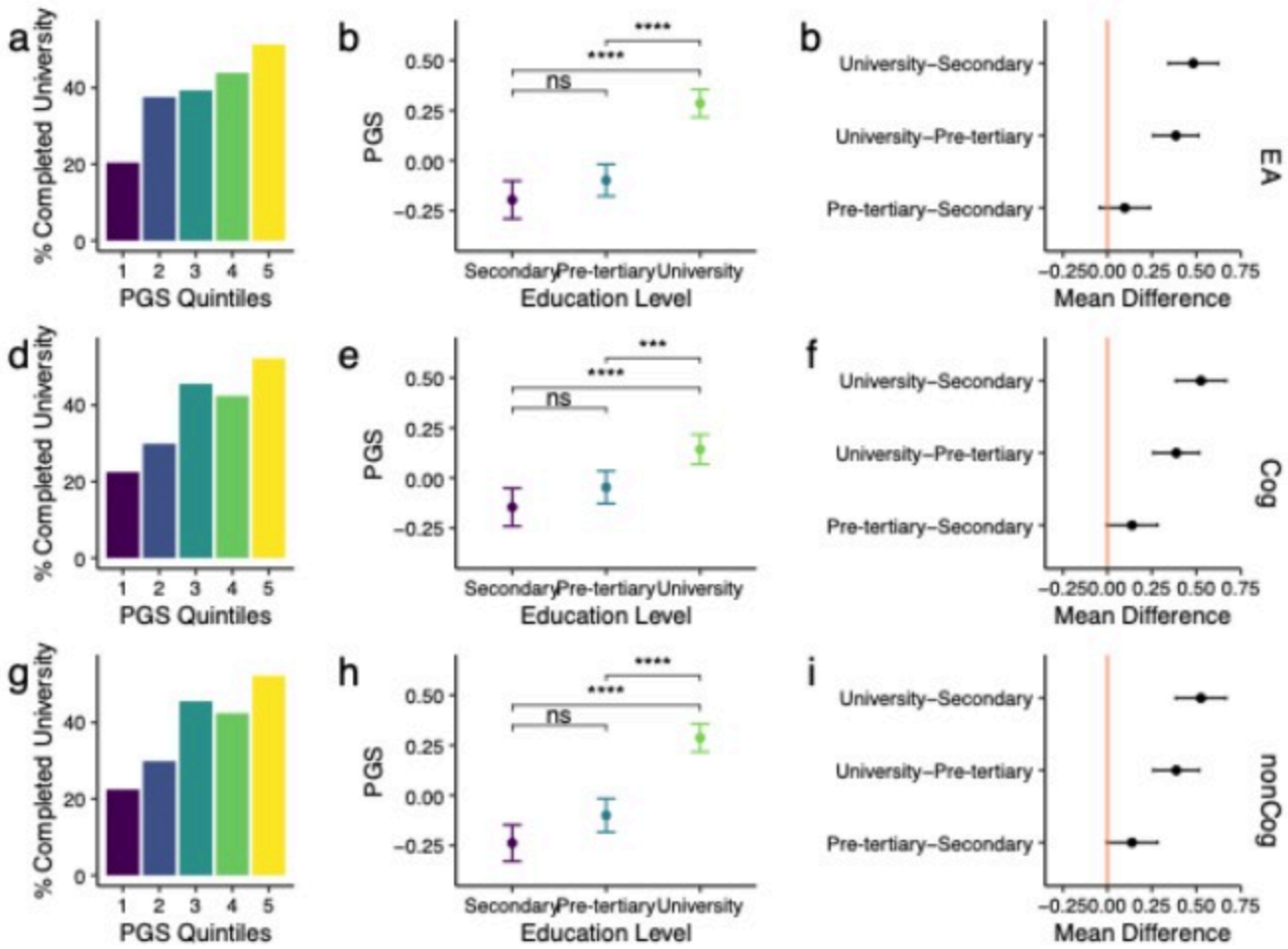


Figure 1

PGSs and educational attainment in parents.PGSs were divided into quintiles to assess the percentage of individuals completing university within each quintile (a, d, g). Mean differences in PGSs across the educational attainment categories (i.e, secondary and lower, pre-tertiary, and university) were analysed using one-way analysis of variance with Tukey post-hoc tests (b, e, h). The points represent the mean PGS and the error bars represent the 95% confidence interval (CI). The mean difference in PGS between pairs of educational attainment categories was plotted (c, f, i), with points representing the mean and the error bars representing the 95% CI. Group means are considered statistically significant if the error bar intervals do not include zero (red line). ns: $p > .05$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, *** $p \leq .0001$.

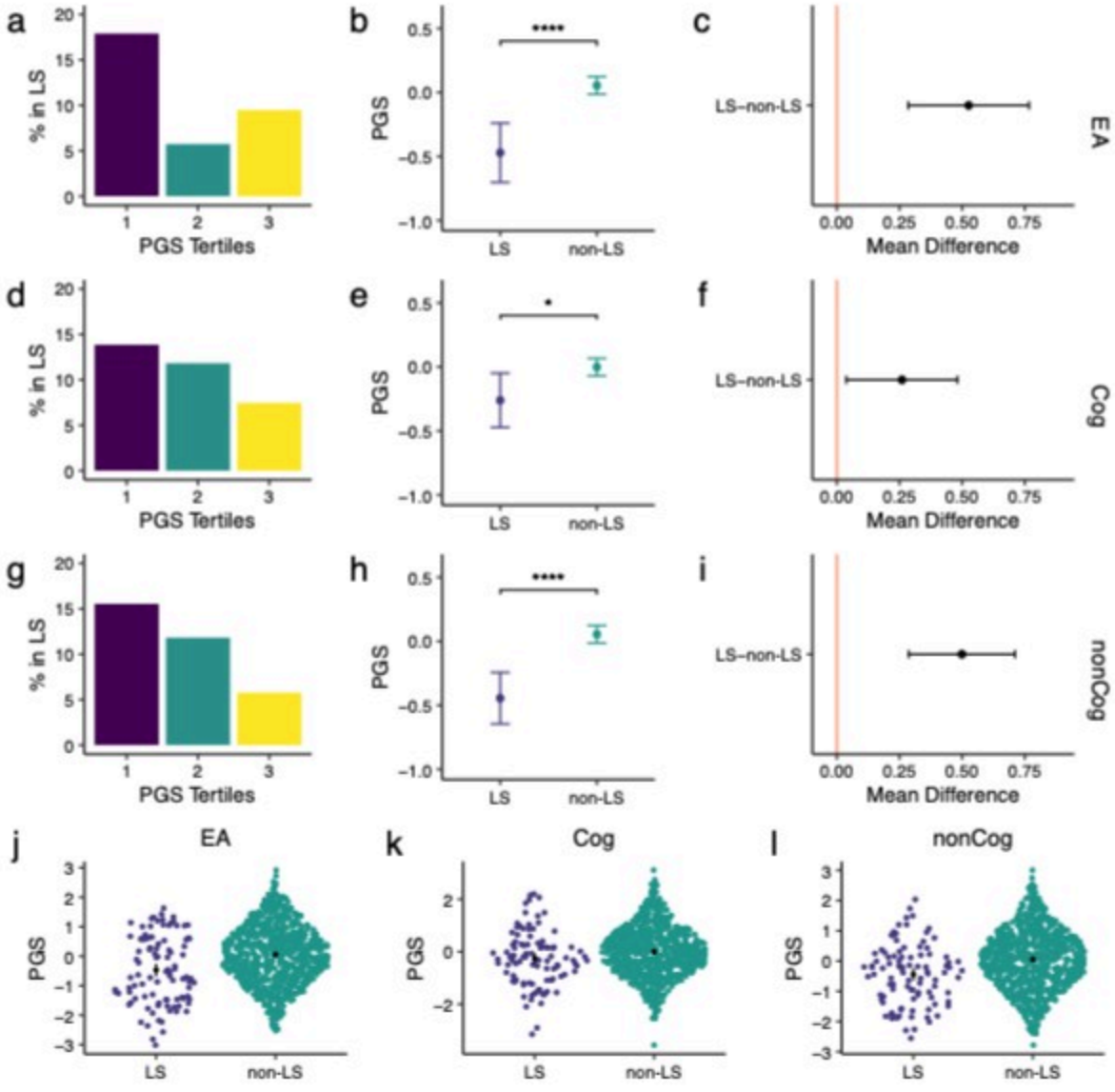


Figure 2

PGSs and learning support in children.PGSs were divided into tertiles and the percentage of children who required learning support was assessed (a, d, g). T-tests were used to determine if there were significant

differences between children who required learning support and those who did not (b, e, h). The points represent the mean PGS and the error bars represent the 95% confidence interval (CI). The mean difference in PGS between pairs of educational attainment categories was plotted (c, f, i), with points representing the mean and the error bars representing the 95% CI. The distribution of PGSs for the two groups was plotted (j, k , l) with each point representing an individual, where the black points represent the group means and the error bars represent the 95% CI. ns: $p > .05$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, *** $p \leq .0001$.

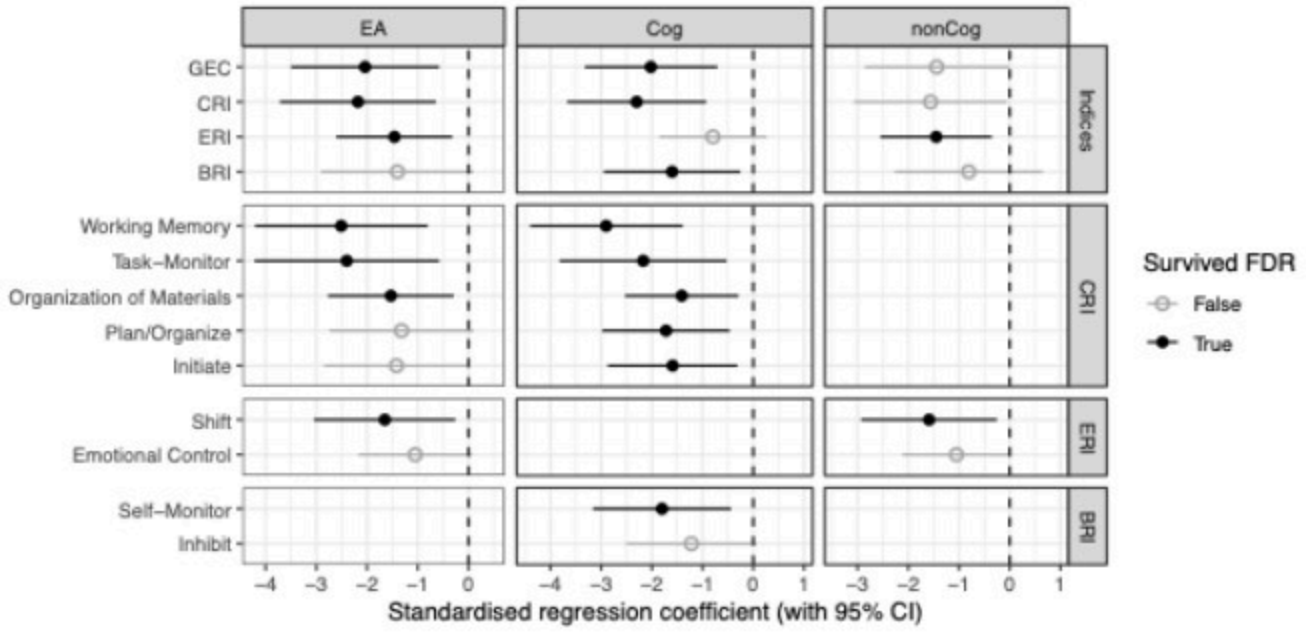


Figure 3

Association of PGS_{EA}, PGS_{Cog} and PGS_{nonCog} with subscales of the T-BRIEF2. T-BRIEF2 items were regressed onto the PGSs with the first three genetic principal components and sex as covariates. The indices were regressed first (top row). Only items of the indices that passed multiple testing corrections (black, filled circles) were regressed onto the PGSs (second, third, and fourth rows). Points represent the standardised regression coefficients of each regression and error bars represent the 95% confidence intervals. Associations that did not pass the false discovery rate (FDR) level of 0.05 were plotted as grey unfilled circles.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- npjSoLChildrenPGSEAandExecutiveFunctionFeb2026Supplementary.pdf