Polygenic Score of Intelligence is More Predictive of Crystallized than Fluid Performance Among Children

Robert J. Loughnan[1], Clare E. Palmer[2], Wesley K. Thompson[3], Anders M. Dale[1,4,5,6,7], Terry L. Jernigan[1,2,5], Chun Chieh Fan[4]

1. Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
2. Center for Human Development, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92161, USA
3. Division of Biostatistics, Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, CA, USA
4. Center for Multimodal Imaging and Genetics, University of California, San Diego School of Medicine, 9444 Medical Center Dr, La Jolla, CA 92037, USA
5. Department of Radiology, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA
6. Department of Neuroscience, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA
7. Department of Psychiatry, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA

Abstract

Scores on intelligence tests have been reported to correlate significantly with educational, occupational and health outcomes. Twin and genome wide association studies in adults have revealed that intelligence scores are moderately heritable. We aimed to better understand the relationship between genetic variation and intelligence in the context of the developing brain. Specifically, we questioned if a genetic predictor of intelligence derived from a large GWAS dataset a) loaded on specific factors of cognition (i.e. fluid vs. crystallized) and b) were related to differences in cortical brain morphology measured using MRI scans. To do this we calculated a genome-wide polygenic score of intelligence (I-GPS) for the Adolescent Brain Cognitive Development (ABCD) baseline data, which consists of 11,875 nine- and ten- year old children across the US. We found that the I-GPS was a highly significant predictor of estimates of both fluid ($t=7.1$, $p=1.2\times10^{-12}$, 0.6% variance explained) and crystallized ($t=15.0$, $p=3.5\times10^{-50}$, 2.4% variance explained) cognition, with greater predictive power for crystallized than fluid ($t=4.9$, $p=8.7\times10^{-7}$). This indicates a stronger loading of I-GPS on crystallized cognition. I-GPS was significantly related to total cortical surface area ($t=5.3$, $p=1.4\times10^{-7}$, 0.3% variance explained), but not mean thickness ($t=0.25$, $p=0.8$). Vertex-wise analyses showed that the surface area association is largely global across the cortex. The stronger association of I-GPS with crystallized compared to fluid measures is consistent with recent results that more culturally dependent measures of cognition are more heritable. These findings in children provide new evidence relevant to the developmental origins of previously observed cognitive loadings and brain morphology patterns associated with polygenic predictors of intelligence.

Introduction

Intelligence is an important indicator of health and societally defined measures of success[1–3] that has been shown to be moderately heritable at around 50%[4]. In intelligence research two latent factors are often distinguished: crystallized and fluid[5]. Crystallized intelligence is related to aspects of cognition that are developed through experience, such as vocabulary, academic skills, and general knowledge. Conversely, fluid intelligence is related to an individual's ability to perform well cognitively in novel situations. Traditional views of these factors predicted that crystallized intelligence would be less influenced by genetics as it was thought to be more impacted by experience and environment[6]. However, recent evidence has shown that this is not the case in adults. Twin studies in adults have demonstrated that more culturally dependent measures of cognition are more heritable[7]. Kan et al.[7] speculated that these results may reflect the presence of gene-environment correlation (rGE). In this case rGE might reflect the fact that individuals with genotypes that initially bias them toward higher cognitive performance are more likely to end up in environments, or have experiences, that further develop these functions. This could occur, for example as a result of streaming students into classes by aptitude. rGE can thus increase heritability estimates. It has been argued that rGE more strongly impacts culturally-dependent measures of intelligence, as society more readily creates environments that facilitate rGE for crystalized intelligence[7]. Higher heritability for more culturally dependent measures of intelligence has been shown for adults, but not for children[7]. As rGE is presumed to accumulate over time[8] we hypothesize that this differentiation in heritability between fluid and crystallized intelligence might develop across childhood. We thus aimed to investigate the relationship between genetic variation and factors of intelligence in the early adolescent brain.

A recent genome wide association study (GWAS) in 269,867 adults associated 205 genomic loci and 1,016 genes to variability in intelligence[9]. By generating a genome-wide polygenic score (GPS) they explained up to 5.2% of the variability in intelligence in independent samples. They found that associated genes were strongly expressed in the brain, and specifically associated with hippocampal pyramidal neurons and striatal medium spiny neurons. Additionally, studies have found that total brain volume and intelligence are correlated at 0.24-0.33[10,11], with both gray and white matter volume contributing to this association[12]. This correlation between intelligence and both gray and white matter volume has been shown to be largely determined by genetics[13,14]. For adults, thicker cortex has sometimes been associated with greater intelligence[13–15]. A recent study, however, in children reported that at age 9 there was no significant relationship between intelligence and cortical thickness, but at age 12 a negative correlation between intelligence and thickness across the cortex was observed[16]. Conversely, cortical area has been shown to be positively associated with intelligence scores in adolescents[17]. Both thickness and area have been shown to be genetically correlated with intelligence in children and adolescents[16,17]. These findings suggest that brain morphology is related to intelligence and that the two share a common genetic basis.

We aimed to further disentangle the associations between genetics, brain morphometry and intelligence in a large cohort (N= 9,511 individuals) of 9- and 10-year-old children obtained from the Adolescent Brain Cognitive Development (ABCD) study. To investigate these associations, we generated an genome-wide polygenic score for intelligence (I-GPS) for each individual in the ABCD dataset using summary statistics from a GWAS of intelligence on 269,867 individuals[9]. After controlling for socioeconomic and demographic differences, we

predicted that the I-GPS would: 1) significantly predict cognitive performance in the ABCD sample; 2) be more associated with crystallized than with fluid intelligence; and, 3) be associated with cortical morphology.

Methods and Data

2.1 ABCD data

The ABCD study (http://abcdstudy.org) consists of N=11,875 individuals aged 9/10 years old at baseline[18]. This longitudinal study was designed to follow the development of children at 21 sites across the US for ten years. The cohort exhibits a large degree of socio-economic and demographic diversity. Exclusion criteria were limited to: 1) lack of English proficiency; 2) the presence of severe sensory, neurological, medical or intellectual issues that would inhibit the child's ability to comply with the protocol; and, 3) an inability to complete an MRI scan at baseline.

Here, we utilized baseline data from ABCD release 2.0 (DOI: 10.15154/1503209). A wide range of measurements were collected for each individual. In addition to demographic and socio-economic variables, for the current study we utilized three data sources: 1) cognitive assessments from the NIH Toolbox[19]; 2) whole-genome genotyping data[20]; and, 3) magnetic resonance imaging[21,22]. Each of these data types will briefly be described below.

2.1.1 NIH Toolbox Cognitive Assessment:

The NIH Toolbox® Cognition Battery (http://www.nihtoolbox.org)[23], herein referred to as 'the Toolbox', consists of seven different tasks that test executive function, working memory, episodic memory, attention, processing speed and language ability. The Toolbox® was normed on individuals between 3 and 85 years old. The total time to complete the battery is approximately 35 minutes. The ABCD study administers the Toolbox in English[24], as eligibility criteria requires that youth participants are fluent in English.

The Toolbox Reading Recognition Task® is a test in which individuals pronounce single words. The Toolbox Picture Vocabulary Task®[25] tests participants vocabulary by asking them to match spoken words to pictures. The Toolbox Pattern Comparison Processing Speed Test®[26] measures processing speed by asking them to identify if two side by side pictures are the same or different as rapidly as possible. The Toolbox List Sorting Working Memory Test ® tests participants working memory by requiring them to order presented objects in size order. The Toolbox Picture Sequence Memory Test® assesses episodic memory by asking participants to reproduce a sequence of items in the correct order[27]. The Toolbox Flanker Task®, a variant of the Eriksen Flanker task[28], is designed to measure cognitive control by requiring individuals to identify the direction of a central arrow that is flanked by either congruent or incongruent arrows. The Toolbox Dimensional Change Card Sort Task® is designed to measure cognitive flexibility[29]. All tasks provide raw scores, uncorrected standard scores, and age-corrected standard scores[24]. Uncorrected task scores were used for all our analyses.

Two summary scores also provided are the Crystallized Composite and Fluid Composite. The Crystallized Composite score is derived from performance on the Reading Recognition and the Picture Vocabulary tasks, the Fluid Composite score from performance on the five remaining measures. These composite scales have been shown to have high convergent validity with 'gold standard' measures of fluid and crystallized intelligence in both adults[30] and children[31].

2.1.2 Genetic Data

Saliva samples were collected at the baseline visit and sent to Rutgers University Cell and DNA Repository for storage and DNA isolation. Genotyping was performed using the Smokescreen array[32], consisting of 733,293 genetic variants. Quality controls (QC) on the genotyping were performed to ensure each genetic variant has been successfully called in more than 95 percent of the sample. After QC 573,845 SNPs remained. Based on genotyped data, we derived genetic ancestry using fastStructure[33] with four ancestry groups. Genetic relatedness was calculated using PLINK. We then performed imputation using the Michigan Imputation Server[34] using hrc.r1.1.2016 reference panel, Eagle v2.3 phasing and multiethnic imputation process. PLINK[35] was used to convert dosage files to plink files using a best guess threshold of 0.9 for each loci.

2.1.3 Neuroimaging Data

The imaging component of the ABCD study was developed by the ABCD Data Analysis and Informatics Center (DAIC) and the ABCD Image Acquisition Workgroup. Imaging methods were developed and optimized to be harmonized across all 21 sites and 3 scanner platforms: Siemens Prisma, General Electric 750 and Phillips. Details of these data collection methods and scanning protocols can be found at [22]. Image postprocessing was conducted by the ABCD DAIC[21]. For each subject, a 3D model of their cortical surface was reconstructed using Freesurfer (http://surfer.nmr.mgh.harvard.edu/). Vertex-wise cortical thickness was estimated after defining the cortical surface and underlying white/gray matter boundary. Vertex estimates of cortical area were computed by calculating the area of elements of the standardized tessellation mapped to each subject's native space. Details of this procedure can be found at [36–40].

2.2 Methods

2.2.1 Computing the Polygenic Score

Polygenic scores aggregate the effects of individual SNPs estimated from a previous GWAS discovery analysis, to produce a single score for each individual. The discovery dataset was computed on 269,867 individuals by Savage et al, using a meta-analysis in which neurocognitive tests primarily gauged fluid cognitive performance[9]. The summary statistics from this analysis were downloaded from (https://ctg.cncr.nl/software/summary_statistics). As nearby SNPs are correlated with one another these are removed before polygenic scoring; this process is known as clumping and pruning. After imputation was performed for the ABCD sample we performed clumping and pruning of SNPs using PRSice[41] with a clumping window of 250 kb, clumping r^2 of 0.1 and no thresholding of significance on the summary statistics. SNPs from the major histone compatibility complex were also removed from the analysis. The polygenic score for each individual was then computed as a sum of their SNPs, with each SNP being weighted by the effect in the discovery sample.

2.2.2 Statistical Model for Behavioral Tasks

To assess the association between the I-GPS and cognitive performance in ABCD, we fit Generalized Linear Mixed-Effect Models (GLMMs). Each model had a different task or composite score from the NIH Toolbox as the dependent variable. In addition to the I-GPS, all models included the fixed effects of sex at birth, parental marital status, age, education level of parent/caregiver, household income and estimated genetic ancestry factors (AMR: American, EAS: East Asian, AFR: African, referenced to EUR: European). Data collection site and family

were input as random effects. Continuous variables were z-scored before model fitting to allow coefficients to be interpreted as standardized effect sizes. GLMMs were implemented using the R gamm4 package[42]. In order to assess the increased predictive power of the I-GPS beyond the covariates alone, we calculated the change in variance explained between the null model (just covariates) and the full model (covariates + I-GPS) and performed a likelihood ratio test to provide a level of significance. To test if standardized regression coefficients differed between analogous regressions we performed a z-test on the difference between coefficients, based on the propagated standard error for the two regression coefficients.

2.2.3 Neuroimaging Analysis

In order to test the association between I-GPS and overall measures of brain morphology, we used the same GLMMs described for predicting Toolbox measures but instead predicted total cortical surface area and overall mean thickness. To explore regional brain morphology features associated with individuals' I-GPS, we fit univariate general linear models to predict vertex-wise area and thickness from I-GPS. The fixed effects were the same as those used for behavioral data. We used scanner ID instead of study site as a covariate, as this is more relevant for imaging measures. All covariates were treated as fixed effects due to the large computational burden of fitting vertex-wise mixed models. Family was excluded as a covariate as treating it as a fixed effect would have drastically increased the number of estimated parameters. Once again predictors and responses were z-scored to allow coefficients to be mapped and interpreted as standardized effect sizes. False discovery rate (FDR) corrected p-values were calculated for each vertex, these were also plotted as maps and used to threshold the maps at an FDR-corrected p value of 0.05.

Results

Behavioral Results

Due to missing demographic information and/or Toolbox scores 1,308 individuals were removed, with 1,018 of those being due to missing declared household income. Failure of individuals' genetic data to pass QC metrics (high calling quality, missing rate lower than 20 percent, and expected genetic relatedness given outbred samples) resulted in a further 1,424 individuals being removed. Table 1 shows behavioral and demographic statistics for the remaining individuals used in this analysis. Note: self-declared race is in this table for the readers' information, however for statistical models estimates of genetic ancestry were used as covariates (see methods).

	Total Analyzed Sample
Total N	9143
	Mean(SD)
Toolbox Fluid Composite Score	92.11 (10.47)
Toolbox Crystallized Composite Score	86.84 (6.92)
Age - months	119.04 (7.47)
Gender	**N(%)**
F	4367 (47.8)

M	4776 (52.2)
Parent Married = Yes	6435 (70.4)
Parental Education	
< HS Diploma	339 (3.7)
HS Diploma/GED	704 (7.7)
Some College	2302 (25.2)
Bachelor	2472 (27.0)
Post Graduate Degree	3326 (36.4)
Household Income	
[<50K]	2554 (27.9)
[>=50K & <100K]	2627 (28.7)
[>=100K]	3962 (43.3)
Race Ethnicity	
White	5103 (55.9)
Hispanic	1745 (19.1)
Black	1144 (12.5)
Asian	195 (2.1)
Other	946 (10.4)

Table 1: Summary of demographics and composite toolbox scores for individuals with full data used in behavioral analysis. (Self declared race is reported here,, however continuous estimates of genetic ancestry were used as covariates instead)

Fitting GLMMs showed that the I-GPS was a significant predictor of both composite scores of the Toolbox (fluid: t=7.1, $p=1.2x10^{-12}$ and crystallized: t=15.0, $p=3.5x10^{-50}$). The change in percent variance explained between the base (covariates only) and full model (adding in I-GPS) was 0.6% for fluid ($\chi^2(1)=50.6$, $p=1.1x10^{-12}$) and 2.4% for crystallized ($\chi^2(1)=222.2$, $p=3.0x10^{-50}$). Due to z-scoring of variables, regression coefficients can be interpreted as standardized effect sizes. Comparing the effect size of I-GPS between the two composite scales, we see that the effect size on crystallized was significantly greater than that for fluid (z=4.9, $p=8.7x10^{-7}$). A full table of outputs from these two regressions can be found in Supplementary Table 1 and 2. We also compared the results with a regression excluding parental education and income as covariates which resulted in a larger percent variance explained by I-GPS, see Supp. Figure 1.

Fitting separate regression models for each individual test of the Toolbox, we found that I-GPS was a significant predictor for each cognitive measure (all p values$<10^{-3}$), except the pattern comparison processing speed task (t=2.51, p=0.01) which did not survive a Bonferroni-corrected significance threshold of 0.05/9=0.006. Standardized coefficients for I-GPS predicting each cognitive measure are displayed in Figure 1 (fluid components – blue, crystallized components – red), alongside the composite score regression coefficients. In this figure there is a clear separation where the cognitive measures used to produce the fluid composite score have consistently lower I-GPS standardized regression weights than the two measures used to produce the crystallized composite score.

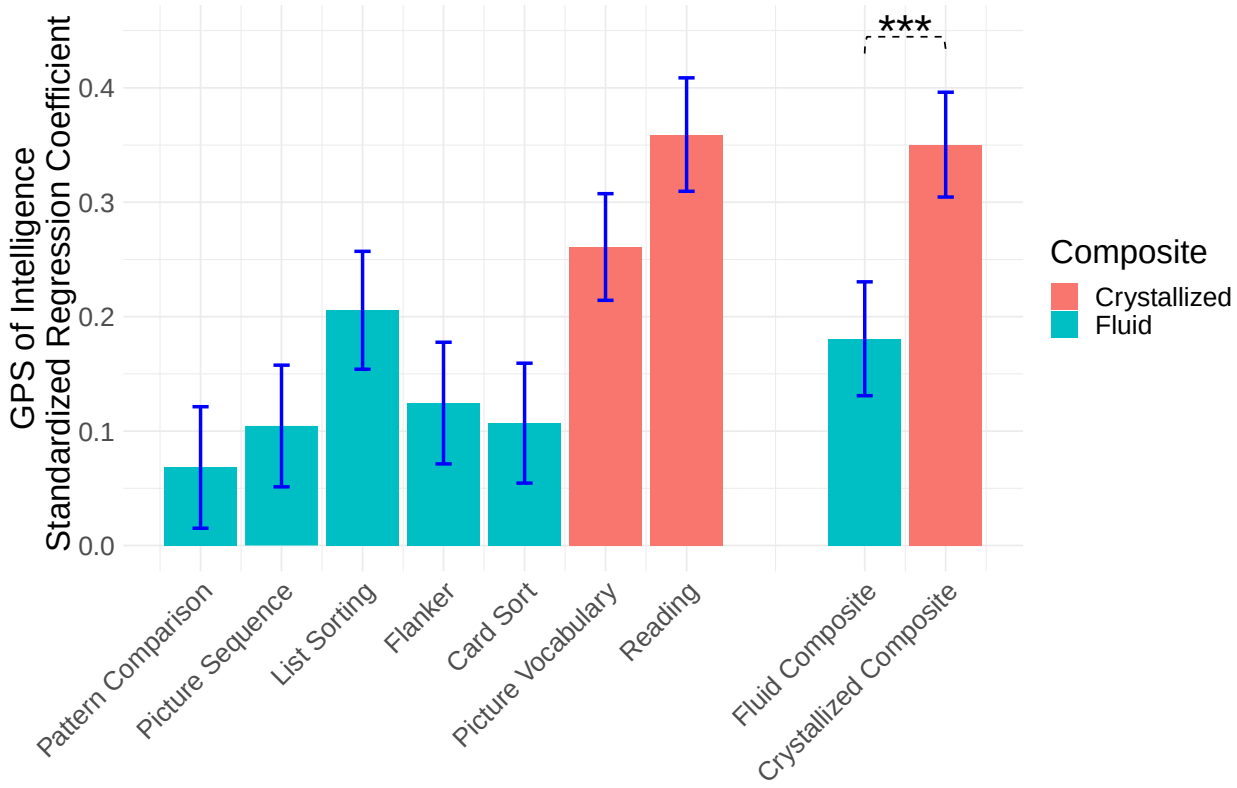

Figure 1: Standardized regression coefficeints of Intelligence GPS for fitting linear mixed models to each Toolbox measure and the two composite scales (fluid and crystallized). Units making up the fluid subscale have consistantly lower regression coefficients than those making up the crystallized subscale.

Neuroimaging Results

For the neuroimaging analyses an additional 261 individuals were excluded due to missing or failed QC of MRI scans. At the level of the whole brain, I-GPS was significantly associated with larger total cortical surface area ($t=5.3$, $p=1.4\times10^{-7}$) explaining 0.3% of the variance in cortical surface area above and beyond the socioeconomic and demographic covariates. I-GPS was not associated with mean thickness ($t=0.25$, $p=0.8$). Figure 2 shows the regional pattern of cortical area associations with higher I-GPS. The maps are of a) standardized regression coefficients (as in Figure 1) and b) FDR-corrected p-values. Both maps are thresholded at an FDR-corrected p-value of 0.05. They suggest a distributed and global cortical area phenotype associated with high I-GPS, characterized by slightly larger associations in bilateral parietal and left lateral pre-frontal regions.

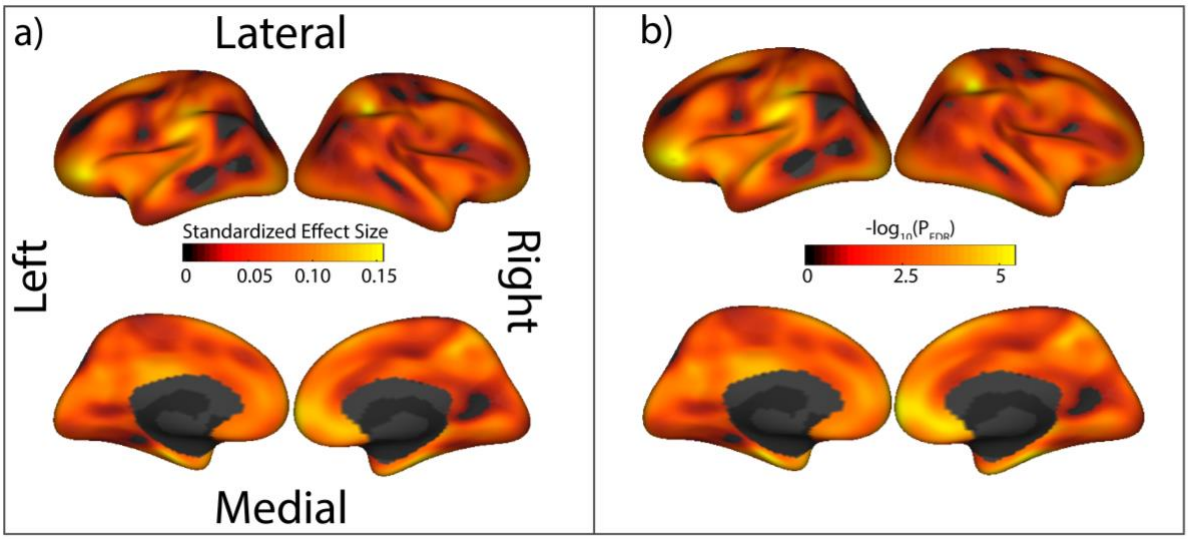

Figure 2 Vertex-wise associations between I-GPS and area. a) Standardized Effect Sizes (predictors and response variables z scored – i.e. units of standard deviation) b) -log10(PFDR). Both maps are thresholded at 0.05 FDR corrected p value.

Discussion

Results reveal that a GPS of intelligence is more predictive of crystallized than fluid cognitive performance in a large sample of 9- and 10-year-old children, despite the fact that the discovery GWAS was trained predominantly on fluid dimensions of cognition[9]. Conventional theories of general intelligence would predict that more culturally-dependent cognition should be more impacted by ones environment and therefore less heritable[6]. We show here, however, that this is not the case, similar to heritability estimates from a prior twin study[7]. It is in fact the more culturally-mediated measures of crystallized intelligence that are more strongly predicted by genetics. A plausible explanation that has been suggested for this unexpected result attributes the effect to gene-environment correlation[7]. For example, individuals with an initial slight bias toward higher cognitive performance may be more likely to end up in environments or having experiences (e.g., reading more or taking more advanced classes in school) that are likely to exaggerate the effect of this initial genetic predisposition. It is argued that the reason for this effect being stronger for culturally-loaded factors of intelligence is that these factors represent societal demands[43]. As such, society creates environments that facilitate gene-environment correlations (rGE) for culture-mediated factors, in a way that it does not for culture-reduced factors. If this argument holds, we expect that as participants in ABCD get older the effects of gene-environment correlation will become greater and the association presented here should become larger (i.e. a larger difference in predictive power of I-GPS between fluid and crystallized factors). In a recent study, Beam and Turkheimer modeled the effects of increasing rGE and showed that it could explain often observed increases in the heritability of measures of cognitive function between childhood and adolescence[8]. We anticipate testing this hypothesis in later time points of this longitudinal study.

In our sample we also found that total cortical area was associated with higher I-GPS. This is consistent with the findings in adults that total brain volume is positively correlated with intelligence[10,11] and that they share a common genetic basis[13,14]. Vertex-wise analysis showed that the pattern of cortical area associated with higher I-GPS was global across the cortex with a few bilateral parietal and left prefrontal regions showing slightly higher

associations. Neither mean nor vertex-wise cortical thickness were found to be significantly associated with I-GPS. This is consistent with a recent study's finding that in 9-year-old children that there was no relationship between cortical thickness and intelligence[16]. The same study showed the emergence of a negative correlation between thickness and measures of intelligence at 12 years of age. We may therefore find that I-GPS is negatively correlated with cortical thickness for future time points as brain development proceeds in the ABCD sample.

A note of caution should be added when interpreting the I-GPS: it should not simply be thought of as a proxy for genetics or 'nature'. Each individual in this study inherited half of their genome from each parent and so these genetic effects can also have indirect influences on their cognitive performance through the cognitively enriching environments that parents provide. Indeed a recent study demonstrated that up to 30% of a polygenic score based on individuals can be explained through a score based on non-transmitted alleles of parents[44]. We find that the association of I-GPS on cognitive performance is attenuated after accounting for parental education and income, see Supp. Figure 1, which suggests that there is shared predictive variance between the I-GPS and parental education and income. Furthermore, it should be emphasized that in addition to one's DNA sequence, epigenetic effects of chromatin and histone modifications as well as DNA methylation are also biological factors that have been shown to impact cognition[45]. These are biological mechanisms that can be impacted by one's environment and influence one's cognitive function and brain structure dynamically over the lifespan.

A limitation of the current study is that the polygenic score discovery dataset was trained only on individuals of European ethnic ancestry[9], which we have deployed in the ABCD dataset – a highly admixed population. Despite controlling for genetic ancestry in our analysis, training and validating polygenic scores on different ancestry groups can cause issues. These issues can result from alleles being present for a certain ancestry group, but being not represented in the mono-ethnic discovery GWAS. Indeed, training polygenic scores on one ancestry group and deploying them in different ancestry groups can decrease predictive performance[46,47]. New methods are being developed to generalize polygenic scores across ethnic groups[48] and these will be important tools to try to ensure that technologies and findings from genetic studies are not limited to overrepresented groups[49].

Although the association between I-GPS and cognitive performance is highly significant, the effect is a moderate one (fluid:0.18σ and crystallized: 0.35σ). We expect that these effects will become larger for later time points collected in the ABCD study. This expectation is based on the robust finding that heritability of intelligence increases over age[50–52] and studies finding that a GPS based on educational attainment (the number of years completed in education) has stronger correlations with school performance of older children[53,54]. The predictor with the largest effect size in our analysis was parental education, with children of highly educated parents (post graduate) on average having crystallized scores 0.89 standard deviations higher than those of low educated parents (<high school diploma) – see Supp. Table 2. Parental education is an important socioeconomic measure that is partially a proxy for material resources. However, it is also confounded by genetics: highly educated individuals are likely to possess genotypes that are advantageous for performing better in school and this in turn will be passed on to their children. It will be important to leverage the wealth of data available in ABCD and other studies to develop new methods that can partial socioeconomic and environmental effects from genetic ones. More precisely characterizing these components will enable us to inform societal policies that can maximize the cognitive potential of individuals.

ABCD Acknowledgement

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners *under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, and U24DA041147*. A full list of supporters is available at https://abcdstudy.org/nih-collaborators. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/principal-investigators.html. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

The ABCD data repository grows and changes over time. The ABCD data used in this report came from [NIMH Data Archive Digital Object Identifier (10.15154/1503209)].

Acknowledgement

We would like to thank Rachel N. Koester for her assistance in developing a theoretical understanding of computational pipelines for general linear mixed effect models.

Bibliography

1. Deary, I. J. & Gottfredson, L. S. Intelligence Predicts Health and Longevity, but Why? *Curr. Dir. Psychol. Sci.* **13**, 1–4 (2004).
2. Kuh, D., Richards, M., Hardy, R., Butterworth, S. & Wadsworth, M. E. J. Childhood cognitive ability and deaths up until middle age: A post-war birth cohort study. *Int. J. Epidemiol.* **33**, 408–413 (2004).
3. Kell, H. J., Lubinski, D. & Benbow, C. P. Who Rises to the Top? Early Indicators. *Psychol. Sci.* **24**, 648–659 (2013).
4. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
5. Cattell, R. B. Theory of Fluid and Crystalized Intelligence: A Critical Experiment. *J. Educ. Psychol.* **54**, 1–22 (1963).
6. Cattell, R. B. *Intelligence: Its structure, growth and action*. *Advances in Psychology, No. 35.* (1987).
7. Kan, K. J., Wicherts, J. M., Dolan, C. V. & van der Maas, H. L. J. On the Nature and Nurture of Intelligence and Specific Cognitive Abilities: The More Heritable, the More Culture Dependent. *Psychol. Sci.* **24**, 2420–2428 (2013).
8. Burgess, A., Shah, K., Hough, O. & Hynynen, K. Phenotype–environment correlations in longitudinal twin models. *Dev Psychopathol.* **25**, 7–16 (2013).
9. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
10. McDaniel, M. A. Big-brained people are smarter: A meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence* **33**, 337–346 (2005).
11. Pietschnig, J., Penke, L., Wicherts, J. M., Zeiler, M. & Voracek, M. Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean? *Neurosci. Biobehav. Rev.* **57**, 411–432 (2015).
12. Thompson, P. M. *et al.* Genetic influences on brain structure. *Nat. Neurosci.* **4**, 1253–1258 (2001).
13. Posthuma, D. *et al.* The association between brain volume and intelligence is of genetic origin. *Nat. Neurosci.* **5**, 83–84 (2002).
14. Hulshoff Pol, H. E. *et al.* Genetic Contributions to Human Brain Morphology and Intelligence. *J. Neurosci.* **26**, 10235–10242 (2006).
15. Haier, R. J., Jung, R. E., Yeo, R. A., Head, K. & Alkire, M. T. Structural brain variation and general intelligence. *Neuroimage* **23**, 425–433 (2004).
16. Brouwer, R. M. *et al.* Genetic associations between intelligence and cortical thickness emerge at the start of puberty. *Hum. Brain Mapp.* **35**, 3760–3773 (2014).
17. Schmitt, J. E. *et al.* A Comprehensive Quantitative Genetic Analysis of Cerebral Surface Area in Youth. *J. Neurosci.* **39**, 2248–18 (2019).
18. Morgan, G. D. *et al.* The conception of the ABCD study: From substance use to a broad NIH collaboration. *Dev. Cogn. Neurosci.* **32**, 4–7 (2017).
19. Bleck, T. P., Nowinski, C. J., Gershon, R. & Koroshetz, W. J. What is the NIH Toolbox, and what will it mean to neurology? *Neurology* **80**, 874–875 (2013).

20. Uban, K. A. *et al.* Biospecimens and the ABCD study: Rationale, methods of collection, measurement and early data. *Dev. Cogn. Neurosci.* **32**, 97–106 (2018).
21. Hagler, D. J. *et al.* Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *bioRxiv* (2018). doi:10.1101/457739
22. Watts, R. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).
23. Weintraub, S. *et al.* Cognition assessment using the NIH Toolbox Cognition assessment using the NIH Toolbox. (2013). doi:10.1212/WNL.0b013e3182872ded
24. Casaletto, K. B. *et al.* Demographically Corrected Normative Standards for the English Version of the NIH Toolbox Cognition Battery. *J Int Neuropsychol Soc.* **21**, 378–391 (2015).
25. Gershon, R. C. *et al.* NIH toolbox cognition battery (CB): Measuring language (vocabulary comprehension and reading decoding). *Monogr. Soc. Res. Child Dev.* **78**, 49–69 (2013).
26. Carlozzi, N. E., Beaumont, J. L., Tulsky, D. S. & Gershon, R. C. The NIH Toolbox Pattern Comparison Processing Speed Test: Normative data. *Arch. Clin. Neuropsychol.* **30**, 359–368 (2015).
27. Dikmen, S. S. *et al.* Measuring Episodic Memory Across the Lifespan: NIH Toolbox Picture Sequence Memory Test Sureyya. *J Int Neuropsychol Soc* **20**, 1–22 (2014).
28. Eriksen, B. & Eriksen, C. Effects of noise letters upon the identification of a target letter in a nonsearch task. *J. Child Psychol. Psychiatry Allied Discip.* **16**, 143–149
29. Zelazo, P. D. *et al.* NIH Toolbox Cognition Battery (CB): validation of executive function measures in adults. *J. Int. Neuropsychol. Soc.* **20**, 620–629 (2014).
30. Heaton, R. K. *et al.* Reliability and validity of composite scores from the NIH toolbox cognition battery in adults. *J. Int. Neuropsychol. Soc.* **20**, 588–598 (2014).
31. Akshoomoff, N., Beaumont, J. L., Bauer, P. J. & Dikmen, S. NIH Toolbox Cognitive Function Battery (CFB): Composite Scores of Crystallized, Fluid, and Overall Cognition. *Monogr Soc Res Child Dev.* **78**, 119–132 (2013).
32. Baurley, J. W., Edlund, C. K., Pardamean, C. I., Conti, D. V. & Bergen, A. W. Smokescreen: A targeted genotyping array for addiction research. *BMC Genomics* **17**, 1–12 (2016).
33. Raj, A., Stephens, M. & Pritchard, J. K. FastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics* **197**, 573–589 (2014).
34. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284 (2016).
35. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, (2015).
36. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *Neuroimage* **9**, 179–194 (1999).
37. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *Neuroimage* **9**, 195–207 (1999).
38. Fischl, B. *et al.* Sequence-independent segmentation of magnetic resonance images. *Neuroimage* **23**, S69–S84 (2004).
39. Dale, A. M. *et al.* Automatically Parcellating the Human Cerebral Cortex. *Cereb. Cortex* **14**, 11–22 (2004).
40. Fischl, B. & Dale, A. M. Measuring the thickness of the human cerebral cortex from

magnetic resonance images. *Proc. Natl. Acad. Sci.* **97**, 11050–11055 (2000).
41. Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466–1468 (2015).
42. Wood, S. & Scheipl, F. gamm4: Generalized additive mixed models using mgcv and lme4. R package version 0.2-3. (2014).
43. Dickens, W. T. What is g? Paper presented at the annual meeting of the Behavior Genetics Association. in (2008).
44. Kong, A. *et al.* The nature of nurture: Effects of parental genotypes. *Science (80-.).* **359**, 424–428 (2018).
45. Day, J. J. & Sweatt, J. D. Epigenetic mechanisms in cognition Jeremy. *Neuron* **70**, 813–829 (2011).
46. Vilhjálmsson, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).
47. Martin, A. R. *et al.* Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
48. Márquez-Luna, C., Loh, P. R. & Price, A. L. Multi-ethnic polygenic risk scores improve risk prediction in diverse populations. *Genet. Epidemiol.* **41**, 811–823 (2017).
49. Petrovski, S. & Goldstein, D. B. Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine. *Genome Biol.* **17**, 16–18 (2016).
50. McClearn, G. E. *et al.* Substantial Genetic Influence on Cognitive Abilities in Twins 80 or More Years Old. *Science (80-.).* **276**, 1560–1563 (1997).
51. Trzaskowski, M., Yang, J., Visscher, P. M. & Plomin, R. DNA evidence for strong genetic stability and increasing heritability of intelligence from age 7 to 12. *Mol. Psychiatry* **19**, 380–384 (2014).
52. Haworth, C. M. A. *et al.* The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* **15**, 1112–1120 (2011).
53. Krapohl, E. *et al.* The stability of educational achievement across school years is largely explained by genetic factors. *npj Sci. Learn.* **3**, 1–10 (2018).
54. Selzam, S. *et al.* Predicting educational achievement from DNA. *Mol. Psychiatry* **22**, 267–272 (2017).