Supplemental File 1: Supplemental Material for Global Ancestry and Cognitive Ability in a Sample of American Youths: Multi-Group Confirmatory Factor Analysis

#### 1. Data Preparation

The following variables were used to perform a multiple-group confirmatory factor analysis (MGCFA) of the cognitive tests used in the Adolescent Brain Cognitive Development cohort. Additional details about these tests and their battery can be found at the website for the study. The tests used were:

- Picture Vocabulary
- Flanker
- List Sorting
- Card Sorting
- Pattern Comparison
- Picture Sequence Memory
- Oral Reading Recognition
- The WISC Matrix Test
- The Little Man Test
- The Rey Auditory Verbal Learning Test (RAVLT): Immediate
- RAVLT: Delayed

To delineate SIRE groups, we used ABCD's race\_ethnicity variable which organizes individuals into 5 mutually exclusive categories (single race, non-Hispanic White, "White"; single race, non-Hispanic Black, "Black"; Hispanic of any race, "Hispanic" (this group is known to be very heterogeneous in self-description and ancestry as will be shown with ternary plots and other methods); single race, non-Hispanic Asian; "Asian"; and a residual non-Hispanic other groups, "Other"). These classifications were based on parental responses to 18 questions asking about the child's race and one question asking about ethnicity. Non-Hispanic children who were reported to belong to two or more races were classified as Other (a heterogeneous collection of different groups, primarily composed of individuals who identify as more than one race. The Asian group was not included due to its minute sample size coupled with its considerable ethnic heterogeneity (as it includes both South and East Asians). In addition to removing the Asian category, we removed any individual who was identified as Asian with the multiple choice SIRE questions. This includes Asians classified as Hispanic and also multi-racial individuals, classified as Other, who were identified as being part Asian. For these MGCFA analyses, we included individuals with missing admixture and genetic-based scores. However we also verified that the results held for the subsets not missing these data.

Outlier detection was the first step of data preparation. Rosner tests were run to reduce the possibility that observations near outliers would be masked (see Rosner, 1983). These were conducted at the level of the individual test and indicated eighteen outliers in the dataset of, at this point, n = 11,124. Eleven of these outliers were from the List Sorting test, while four of them came from the Picture Vocabulary test, one of them came from the Flanker test, and one of them was from the Little Man test. Removing only the outlier test scores for these individuals and imputing their artificially missing scores did not affect the results<sup>1</sup> because they made such a minor contribution to the aggregate sample. Additionally, retaining these outliers did not affect the results of any analyses, only affecting the assessment of differences between linear and nonlinear age-score relationships. As such, primary analyses do not involve these eighteen observations since we did not want to deal with assumptions about the reasons for their outlying score. To some, the inclusion of these people would taint subsequent analyses even though our results were robust to their imputation and inclusion.

The second step of data preparation was assessment of missingness and subsequent imputation of missing data. The largest amount of missingness was observed for the Little Man Test, with 2.84% of cells missing. This was followed by the Delayed RAVLT, which had 2.28% of observations missing, then by the Matrix Test, which was missing 2.11% of observations, and then by the Immediate RAVLT, which was missing 1.80% of observations. List Sorting was missing 1.72%, followed by Pattern Comparison which was missing 1.49%, then Oral Reading Recognition with 1.41%, Picture Sequence Memory with 1.40%, Flanker and Card Sorting with the same 1.34%, and finally Picture Vocabulary, with 1.30%. Before imputation was conducted, we assessed the possibility of score and demography-related patterns of missingness, observing, firstly, that there were no differences in missingness for certain tests based on scores on other tests, nor differences by age (average = 118.96 months, SD = 7.48, range = 107-133 months), broad race/ethnicity (see the description of the variable "race ethnicity" for more detail) or sex (n female = 5,299 and male = 5,807). With no pattern to demographic or scorebased missingness, we conclude that imputation is viable since it does not appear that there is systematic missingness by any variable relevant to our focal analyses. For completeness' sake, we ran our analyses with the removal of all cells with missingness and there were no differences. In virtually every case, there were no differences in results to three decimal places, excepting those for  $\chi$ 2, which did not differ enough to affect the interpretation of our results. After finding that missingness did not present any immediately discernible rhyme nor reason, we utilized Iterative Robust Model-Based Imputation or IRMI (Templ et al., 2011) with our convergence threshold set to 5, the number of multiple imputations set to 1, and our maximum number of iterations set to 100 using the R package VIM (Kowarik & Templ, 2016). Two observations had to be removed to make this possible, since they included no responses to any cognitive test.

Subsequent aspects of data preparation involved adjustment for criteria like age, sex, assessment site, and family ID, but all analyses were run with all combinations of these adjustments and the lack thereof as well as adjustments on a per-group rather than an

<sup>&</sup>lt;sup>1</sup> This sort of remark refers to the results of supplementary analyses throughout this supplement.

aggregated basis, with no alteration of our ultimate results. Therefore, we consider these results to be robust to these corrections, even if the results presented in this supplement are concerned primarily with fully aggregately adjusted data.

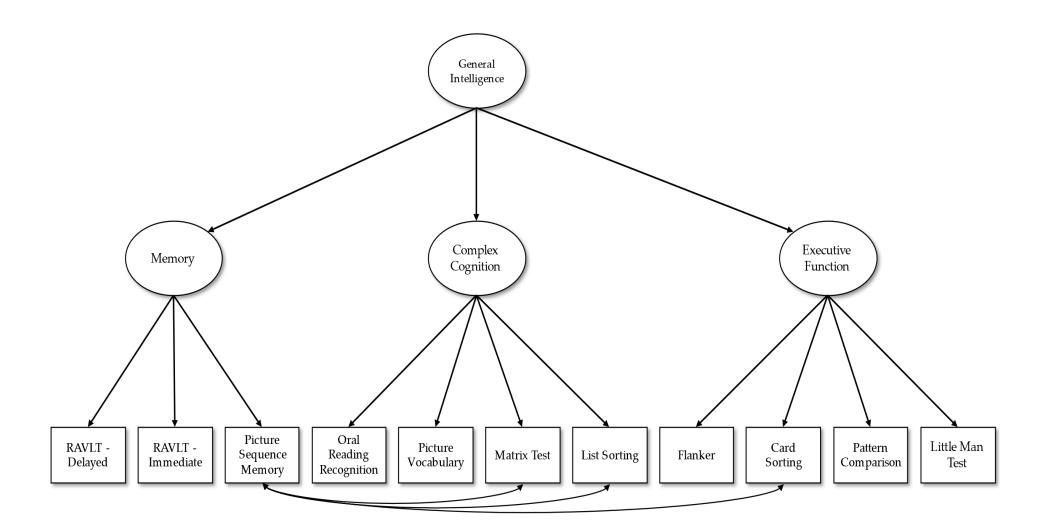
Our first adjustment involved age. We investigated the possibility of nonlinear effects of age by comparing linear regressions to Savitzsy-Golay filter (i.e., LOESS) results for the age-test score relationship at the level of each individual test. There were no meaningful differences between LOESS and linear regression. Next, we observed that all residuals were near-zero for the linear regressions of age on test scores and that they were all nearly normal, though there were clear ceiling and floor effects in the areas where observations were scant on many of the tests. Trimming to remove those effects on a per-test basis where the ceiling or floor appear to begin did not affect the results; this was rerun with trimming at plus and minus 0.1 standard deviations from that point to no effect.

We assessed the same results for restricted cubic splines (RCS; with three to nine splines) and generalized additive models (GAM) with little difference. We then compared the  $\chi 2$  and AIC values of these models, using a p value of 0.05 and a p value – adjusted to be comparable to 0.05 at our large sample size – of 0.000007 (see Naaman, 2016). When presented, p values are not rounded based on the next significant digit to avoid improper rounding issues; when they are highly significant or insignificant, however, they are presented based on a boundary *p* value like 0.05. We focus on the scaled results, since those were more likely to be accurate given our large sample. We reran models where possible if p values <0.05 were indicated and this did not change our results. In terms of  $\chi$ 2, the RCS and GAM models did not fit better except for the Matrix Test: the RCS fit best for the Matrix Test but adjusting for it did not affect our results, so we residualized for it like the rest of the tests anyway. In terms of AIC, the GAM was the best fit for Picture Vocabulary, the Matrix Test, and the Delayed RAVLT. The RCS fit best in terms of AIC for Flanker. Adjusting for age based on GAMs or RCS did not change our ultimate results. This may be because of the effect of sample size on AIC, where the differences of between six and forty-six points ought to have been considered negligible. We did not consider the GAM a better fit when indicated by AIC when the degrees of freedom of the GAM were also negative, since this represented an invalid model. However, adjusting for those invalid GAMs also did not affect results.

Regardless of the reasons for the lack of effect, it is beyond this paper to ascertain them. What is certain, however, is that it ultimately did not affect results following adjustment. Breusch-Pagan tests were insignificant for all regressions except for Picture Vocabulary and the Little Man test. The same pattern was observed for non-constant variance score tests with the addition of Flanker. Despite this, our MGCFA results did not differ whether they were typical or robust, perhaps due to the small number of affected tests. The results also did not differ

when the three tests were excluded in pairs. This was a possibility for testing because they were loaded on factors with three indicators without them, albeit with other biased indicators forced to be included (i.e., Flanker plus Picture Vocabulary or Little Man plus Picture Vocabulary removed). With the model refit with equal factor loadings for the two remaining subtests with all of them excluded, results, surprisingly, differed only marginally. Finally, we fit a local structural equation model (Hildebrandt et al., 2009, 2016) across our range of ages (in months, with a bandwidth of two). RMSEA and CFI were not appreciably better or worse across the range of ages and there was no change in BIC from the youngest to the oldest ages.

Our second adjustment involved correcting for sex in the outlier-pruned, age-adjusted data. To qualify adjustment by sex, we tested an MGCFA by sex. We fit our model using Pornpattananangkul et al.'s (2021) model, for this same sample, as guidance. The group factors modeled were dubbed Complex Cognition (COC), Memory (MEM), and Executive Function (EF); these are pictured below. Because the number of factors was three, there were no differences in fit between a higher-order factor model in which *g* sat atop the group factors and one in which a model with correlated group factors was fit; a bifactor model did fit better, but we elected not to pursue testing with this model because it is not acceptable on theoretical grounds (Hood, 2010), although this can be subject to change given certain results not presently found in the literature on intelligence (e.g., common pathway models supporting a bifactor model over a higher-order one). The model required three residual covariances, between Picture Sequence Memory and the Matrix Test, List Sorting, and Card Sorting. Model fits are provided in Table S1.



**Table S1.** Model Fit Statistics for the ABCD Sex MGCFA

Model	Description	Comparison	χ2	df	CFI	RMSEA	BIC
MGCFM	First-Order						
B1	Configural	NA	1086.32	76	0.970	0.049	310166
B2	Metric	B1	1110.99	84	0.969	0.047	310116
В3	Scalar	B2	1424.88	92	0.960	0.051	310355
B3A	Partial Scalar*	B2	1144.36	89	0.968	0.046	310303
B4	Strict	B3A	1253.09	100	0.966	0.046	310109
B5	Latent Variances	B4	1257.70	103	0.966	0.045	310085
B6	Latent Means	B5	1485.99	106	0.959	0.048	310286
MGCFA	Higher-Order						
M1	Configural	B1	1086.32	76	0.970	0.049	310166
M2	Metric	M1	1113.05	86	0.969	0.046	310099
M3	Scalar	M2	1426.96	93	0.960	0.051	310348
M3A	Partial Scalar*	M2	1146.41	90	0.968	0.046	310095
M4	Strict	M3A	1255.35	101	0.965	0.045	310102
M5	Latent Variances	M4	1259.91	104	0.965	0.045	310078
M6	Latent Means	M5	1488.11	108	0.958	0.048	310269
M6A	Latent Means Group Factors	M5	1477.42	107	0.959	0.048	310268
M6B	Latent Means MEM and EXE**	M5	1260.47	106	0.965	0.044	310060
M6C	Latent Means MEM, EXE, and g	M5	1384.24	107	0.962	0.046	310175

<sup>\*</sup> The intercepts for the Little Man Test, Matrix Test and Flanker were freed. \*\* We tested among models of all possible individual group factor constraints and used BIC to decide among them.

In the higher-order model prior to any mean constraints, differences in g amounted to an insignificant (p = 0.015) 0.053 g female advantage, a 0.165 g advantage in MEM and a 0.149 g advantage in EXE with a 0.118 g deficit in COC. These came with Z values of 2.444, 9.365, 7.037 and 16.231, respectively. In a model without differences in g, the group differences in MEM, COC, and EXE, respectively, were 0.198, -0.070 (not significant, p = 0.003), and 0.183 in favor of the female group (Zs = 9.692, 2.994 and 7.374). The only major difference between the male and female groups was in the variances of their factors. For example, the standard deviations for the factor scores for g, COC, MEM, and EXE were 0.922, 2.314, 1.220, and 1.263 for males versus 0.837, 2.134, 1.176, and 1.098 for females. SDI<sub>2</sub>, an effect size for invariance violations proposed by Gunn et al., (2020), yielded values of 0.140, -0.154 and -0.242 for the Matrix Test, Flanker, and Little Man Test (positive = favors the female group and vice-versa). Thus, all the violations of invariance observed had small-to-moderate effects. Since we aimed to use factor scores, which are unaffected by this, and all groups had very similar sex ratios, we corrected for sex. We also

corrected for assessment site and family ID. Invariance by site could not be reasonably assessed due to the small samples found in some sites. To the extent this was the case, family ID was worse. Adjusting or not adjusting, the results were the same; adjustment was still done to obviate concern about the results.

#### 2. Multiple Group Confirmatory Factor Analysis for Blacks, Whites, Hispanics, and Other

After preparing the cognitive data for analysis by self-described race or ethnicity, we performed an MGCFA with the same model used to assess invariance for sex. There were 6,176 participants in the White group, 1,780 in the Black group, 2,318 in the Hispanic group, and 830 in the "Other" group. With the full sample of 11,104 being used, the cutoff Z-value used was 4.34. Our MGCFA model fit results are as follows in Table S2. Table S5 contains the means and standard deviations for the resulting factor scores with the latter in parentheses while Tables S3-S4 contain the means from the MGCFA model in units of Hedge's g. The Hispanic group's means were set to 0. The other groups are compared relative to them. The gaps from the best-fitting mean model (i.e., the strong form of Spearman's hypothesis where only g causes differences; all results are in units of Hedge's g) were -0.584 for Blacks, 0.022 for "Other", and 0.523 for Whites. Both the strong form of Spearman's hypothesis and the model with both the MEM and EXE factors constrained (M6B) had nearly equivalent fits. CFI was lower for the former.

Table S2. Model Fit Statistics for the ABCD Race and Ethnicity MGCFA

Model	Description	Comparison	χ2	df	CFI	RMSE A	BIC
MGCFM	First-Order						
B1	Configural	NA	1102.19	152	0.967	0.047	306263
B2	Metric	B1	1262.14	176	0.962	0.047	306200
B3	Scalar	B2	1592.26	200	0.951	0.050	306306

B3A	Partial Scalar*	B2	1302.09	194	0.961	0.045	306072
B4	Strict	B3A	1977.72	227	0.939	0.053	306440
B4A	Partial Strict**	B3A	1489.84	218	0.956	0.046	306036
B5	Latent Variances	B4A	1500.89	227	0.956	0.045	305963
B6	Latent Means	B5	2538.39	236	0.920	0.059	306917
MGCFA	Higher-Order						
M1	Configural	B1	1102.20	152	0.967	0.047	306263
M2	Metric	M1	1277.08	182	0.962	0.047	306159
M3	Scalar	M2	1607.68	203	0.951	0.050	306294
M3A	Partial Scalar*	M2	1317.33	197	0.961	0.045	306059
M4	Strict	M3A	1997.06	230	0.938	0.053	306432
M4A	Partial Strict**	M3A	1505.48	221	0.955	0.046	306024
M5	Latent Variances	M4	1516.54	230	0.955	0.045	305951
M6	Latent Means	M5	2553.83	242	0.919	0.059	306877
M6A	<b>Latent Means Group Factors</b>	<b>M</b> 5	1545.10	239	0.954	0.044	305896
M6B	Latent Means MEM and EXE	M5	1519.38	236	0.955	0.044	305898
M6C	Latent Means MEM, EXE, and g	M5	2125.01	239	0.934	0.053	306476

<sup>\*</sup> The intercepts for the Picture Vocabulary and Picture Sequence Memory Tests were freed. \*\* The variances for Flanker, Card Sorting and Pattern Comparison were freed.

Our criteria for *not* moving to partial invariance was stricter than what is typical in the literature on MGCFA, but we believe it is justified to understand which tests might be biased for users of the ABCD data. The page below contains a plot of factor scores for *g* by group.

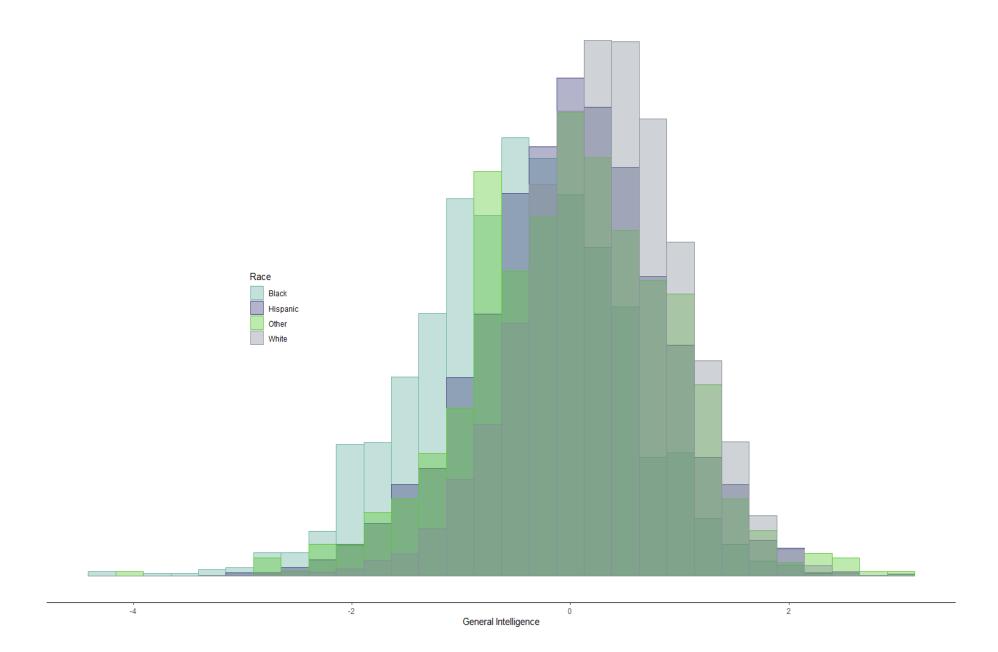


Table S3. Group Differences in the ABCD (MGCFA; Latent Variance Model)

Group	General Intelligence	Memory	Executive Function	Complex Cognition
Black	-0.450	-0.138	-0.117	-0.086
White	0.339	0.006	0.066	0.247
Other	0.016	-0.059	-0.005	0.044

The Hispanic group is the comparison group, whose means are set to 0.

Table S4. Group Differences in the ABCD (MGCFA; Spearman's Weak Hypothesis Model)

Group	General Intelligence	Memory	Executive Function	Complex Cognition
Black	-0.661	0	0	0.109
White	0.394	0	0	0.197
Other	-0.044	0	0	0.097

The Hispanic group is the comparison group, whose means are set to 0.

**Table S5.** Group Differences in the ABCD (Factor Scores from Latent Variance Model)

Group	General Intelligence	Memory	Executive Function	Complex Cognition
White	0.306 (0.724)	0.235 (1.149)	0.312 (0.995)	1.090 (1.664)
Black	-0.481 (0.897)	-0.537 (1.217)	-0.519 (1.057)	-1.140 (1.985)
Hispanic	0 (0.819)	0 (1.159)	0 (1.049)	0 (1.903)
Other	0.017 (0.912)	-0.065 (1.246)	0.007 (1.104)	0.140 (2.117)

*Note*: Standard deviations (*SDs*) are in parentheses. Note, within SIRE groups the SDs are not 1.

## References

- Gunn, H. J., Grimm, K. J., & Edwards, M. C. (2020). Evaluation of Six Effect Size Measures of Measurement Non-Invariance for Continuous Outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(4), 503–514. https://doi.org/10.1080/10705511.2019.1689507
- Hildebrandt, A., Lüdtke, O., Robitzsch, A., Sommer, C., & Wilhelm, O. (2016). Exploring Factor
   Model Parameters across Continuous Variables with Local Structural Equation Models.
   Multivariate Behavioral Research, 51(2–3), 257–258.
   https://doi.org/10.1080/00273171.2016.1142856
- Hildebrandt, A., Wilhelm, O., & Robitzsch, A. (2009). Complementary and competing factor analytic approaches for the investigation of measurement invariance. *Review of Psychology*, 16(2), 87–102.
- Hood, S. B. (2010). *Latent Variable Realism in Psychometrics*. https://scholarworks.iu.edu/dspace/handle/2022/8120
- Kowarik, A., & Templ, M. (2016). Imputation with the R Package VIM. *Journal of Statistical Software*, 74(1), 1–16. https://doi.org/10.18637/jss.v074.i07
- Naaman, M. (2016). Almost sure hypothesis testing and a resolution of the Jeffreys-Lindley paradox. *Electronic Journal of Statistics*, 10(1), 1526–1550. https://doi.org/10.1214/16-EJS1146

- Pornpattananangkul, N., Riglin, L., Anney, R., Wang, Y., Barch, D. M., Thapar, A., & Stringaris, A. (2021). Motivation and Cognitive Abilities as Mediators between Polygenic Scores and Psychopathology in Children. medRxiv, 2020-06.
- Rosner, B. (1983). Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics*, 25(2), 165–172. https://doi.org/10.2307/1268549
- Templ, M., Kowarik, A., & Filzmoser, P. (2011). Iterative stepwise regression imputation using standard and robust methods. *Computational Statistics & Data Analysis*, 55(10), 2793–2806. https://doi.org/10.1016/j.csda.2011.04.012

#### 1. Data

Seven indicators were used to compute socioeconomic status: financial adversity, area deprivation index, neighborhood safety protocol, parental education, parental income, parental marital status, and parental employment status. These are detailed below:

- 1. Financial Adversity: The seven item Financial Adversity Questionnaire (PRFQ) was administered to parents. They were asked: "In the past 12 months, has there been a time when you and your immediate family experienced any of the following:
- (1) "Needed food but could not afford to buy it or could not afford to go out to get it?" (1 = "yes", 0 = "no"),
- (2) "Were without telephone service because you could not afford it?" (1 = "yes", 0 = "no"),
- (3) "Did not pay the full amount of the rent or mortgage because you could not afford it?" (1 = "yes", 0 = "no"),
- (4) "Were evicted from your home for not paying the rent or mortgage?" (1 = "yes", 0 = "no"),
- (5) "Had services turned off by the gas or electric company, or the oil company would not deliver oil because payments were not made?" (1 = "yes", 0 = "no"),
- (6) "Had someone who needed to see a doctor or go to the hospital but did not go because you could not afford it?" (1 = "yes", 0 = "no"), and
- (7) "Had someone who needed a dentist but could not go because you could not afford it?" (1 = "yes", 0 = "no")

We summed responses (maximum: 7; minimum: 0) and reverse coded the variable so that higher scores indicated less financial adversity, The results were then standardized.

- 2. Area Deprivation Index (ADI): Parents completed a residential history questionnaire including residential addresses and the number of full years they lived at each residence. ABCD computed Area Deprivation Index (ADI) for each residential address based on the following variables:
  - 1. "Percentage of occupied housing units without complete plumbing (log)"

- 2. "Percentage of occupied housing units without a telephone"
- 3. "Percentage of occupied housing units without a motor vehicle"
- 4. "Percentage of single"
- 5. "Percentage of population below 138% of the poverty threshold"
- 6. "Percentage of families below the poverty level"
- 7. "Percentage of civilian labor force population aged >=16 y unemployed (unemployment rate)"
- 8. "Percentage of occupied housing units with >1 person per room (crowding)"
- 9. "Percentage of owner"
- 10. "Median monthly mortgage"
- 11. "Median gross rent"
- 12. "Median home value"
- 13. "Income disparity defined by Singh as the log of 100 x ratio of the number of households with <10000 annual income to the number of households with >50000 annual income"
- 14. "Median family income"
- 15. "Percentage of population aged >=25 y with at least a high school diploma"
- 16. "Percentage of population aged >=25 y with <9 y of education"

Scores were provided in terms of national percentiles. We used scores for the most recent residence (variable: reshist\_addr1\_adi\_perc). The resultant values were reverse coded to make higher values indicate better neighborhoods, and then standardized.

- 3. Neighborhood Safety Protocol: Parents were asked three Likert scale questions about neighborhood safety: "I feel safe walking in my neighborhood, day or night," "Violence is not a problem in my neighborhood," and "My neighborhood is safe from crime" (1 = strongly disagree; 5 = strongly agree). ABCD pre-computed means scores based on these three questions (Minimum =1; Maximum =5) (variable: nsc\_p\_ss\_mean\_3\_items). We standardized these scores.
- 4. Education: Parents were asked about educational attainment: "What is the highest grade or level of school you have completed or the highest degree you have received." We recoded

responses to create interval scores (ranging from 0 to 18): Never attended/Kindergarten only = 0, 1st grade = 1, 2nd grade = 2, 3rd grade = 3, 4th grade = 4, 5th grade = 5, 6th grade = 6, 7th grade = 7, 8th grade = 8, 9th grade = 9, 10th grade = 10, 11th grade = 11, 12th grade = 12, High school graduate = 12, GED or equivalent Diploma General = 12, Associate degree: Occupational Program = 14, Associate degree: Academic Program = 14, Bachelor's degree = 16, Master's degree = 18, Professional school = 18, Doctoral degree = 18. We standardized the recoded educational scores for each parent, averaged the standardized scores, and then standardized the average.

- 5. Income: Parents were asked about total family income in the past 12 months. We recoded the variable to a dollar amount scale: 1.00 = less than \$5000 (recode: 4,500); 2.00 = \$5000 to 11,999 (recode: 5,000); 3.00 = \$12,000 to 15,999 (recode: 12,000); 4.00 = \$16,000 to 24,999 (recode: 16,000); 5.00 = \$25,000 to 34,999 (recode: 25,000); 6.00 = \$35,000 to 49,999 (recode: 35,000); 7.00 = \$50,000 to 74,999 (recode: 50,000); 8.00 = \$75,000, to 99,999 (recode: 75,000); 9.00 = \$100,000 to 199,999 (recode: 100,000); 10.00 = \$200,000 and greater (recode: 200,000). The recoded variable was standardized.
- 6. Marital Status: The responding parent was asked about their relationship status. Parental marital status was coded as 1 if married and 0 for any other arrangement (widowed, divorced, separated, never married, living with partner, or refused to answer)

7. Employment Status. The responding parent was asked about their and their partner's employment status. Parental employment was coded as 1 if at least one parent was working now either full or part time and 0 for all other cases.

### 2. Analysis

SES: We imputed missing data for the 7 SES indicators using the mice package (df, m=5, maxit = 50, method = 'pmm', seed = 500). We then standardized the five continuous variables (i.e., everything except marital and employment status). After, we submitted the variables to Principal Component Analysis (PCA), using the R package PCAmixdata to handle mixed categorical and continuous data (Chavent, Kuentz-Simonet, & Saracco, 2014). The first unrotated component explained 42% of the variance in the full sample. The loadings are shown in Table S1 below. This summary SES score correlated at r = .38 with g in the full sample. The correlation matrix for the full sample is shown in Table S1.

Table S1. Principal Component Loadings for the Seven Socioeconomic Indicators.

	PC_1	PC_2	PC_3	PC_4	PC_5
Financial_Adversity	0.31	0.05	0.53	0.07	0.04
ADI	0.49	0.08	0.04	0.02	0.00
Neighborhood_Safety	0.31	0.12	0.15	0.36	0.01

Education	0.53	0.00	0.01	0.11	0.09
Income	0.66	0.00	0.00	0.08	0.00
Marital_Status	0.42	0.09	0.03	0.00	0.44
Employment_Status	0.21	0.58	0.01	0.09	0.08

We also checked the congruent coefficients for the SIRE group PC\_loadings. These were .97 or greater indicating the same structures across SIRE groups.

# References

Marie Chavent, Vanessa Kuentz-Simonet, Amaury Labenne, Jérôme Saracco. Multivariate Analysis of Mixed Data: The R Package PCAmixdata. 2017. 〈 hal-01662595〉

### Supplemental File 3 for Global Ancestry and Cognitive Ability in a Sample of American Youths: Supporting Data

#### **Table of Contents**

- 1. Tab S1. Descriptive Statistics for the ABCD sample and subsamples
- 2. Tab S2. Means, standard deviations, and correlations with confidence intervals for the SES indicators and g
- 3. Tab S3: Subsample and sample correlation matices
- 4. Tab S4: Black subgroup performance
- 5. Tab S5: Admixture-regression results for White Americans
- 6. Tab S6: eduPGS predictive validities
- 7. Tab S7: Regression Results for the Effect of eduPGS and Ancestry on Cognitive Ability in the SIRE subsamples.
- 8. Tab S8: PseudoPGS results
- 9. Tab S9: dge Matrices for Admixture-regression results with eduPGS

Tab S1. Descriptive Statistics for the ABCD sample and subsamples

Table S1. Descriptive Statistics for the ABCD sample and subsamples

Table 31. Descriptive Stat	Total sample	•	Black		Hispanic		Other		White	
	М	SD	M	SD	М	SDS	М	SD	М	SD
Age (in Months)	118.98	7.49	118.90	7.28	118.59	7.58	118.70	7.40	119.17	7.52
g	0.00		-0.69	1.07	-0.10		-0.09	1.09	0.24	0.86
SES	0.00		-0.98	0.93	-0.36		-0.40		0.45	0.75
frac White SIRE	0.73		0.00	0.00	0.67		0.39		1.00	0.03
frac Black SIRE	0.20		1.00	0.04	0.07		0.29		0.00	0.00
frac Native American	0.02		0.00	0.00	0.03		0.18		0.00	0.00
frac NOC SIRE	0.06	0.23	0.00	0.04	0.23		0.13		0.00	0.03
European_ancestry	0.75		0.16	0.11	0.60	0.21	0.62	0.25	0.98	0.05
African_ancestry	0.18	0.31	0.82	0.11	0.10		0.32	0.26	0.01	0.02
Amerindian_ancestry	0.06	0.14	0.01	0.02	0.28	0.19	0.04	0.09	0.01	0.03
South_Asian_ancestry	0.00	0.02	0.00	0.01	0.01	0.01	0.01	0.05	0.00	0.01
East_Asian_ancestry	0.01	0.03	0.01	0.02	0.01	0.02	0.01	0.07	0.00	0.02
State_racism	0.00	1.00	0.44	0.93	-0.35	0.93	0.26	0.99	-0.04	0.99
Discrim_fact	0.00	1.00	0.47	1.24	0.09	1.04	0.19	1.07	-0.19	0.83
Ethnic_attachment	0.00	1.00	0.38	1.03	0.23	1.01	0.05	1.02	-0.19	0.94
Skin_color	0.00	1.00	1.32	0.42	0.58	0.80	0.36	0.91	-0.62	0.60
P_Brown_Eye	0.56	0.41	0.97	0.10	0.83	0.28	0.73	0.35	0.33	0.35
P_Intermediate_Eye	0.08	0.07	0.02	0.03	0.06	0.06	0.07	0.06	0.10	0.07
P_Blue_Eye	0.36	0.41	0.02	0.08	0.11	0.25	0.20	0.33	0.57	0.39
P_Black_Hair	0.23	0.24	0.55	0.16	0.38	0.23	0.28	0.20	0.09	0.10
P_Brown Hair	0.46	0.18	0.43	0.13	0.49	0.16	0.54	0.15	0.45	0.20
P_Red_or_Blond_Hair	0.30	0.29	0.02	0.05	0.13	0.18	0.18	0.21	0.46	0.26
Child_USA_Born	0.98	0.15	0.98	0.15	0.94	0.24	0.98	0.14	0.99	0.12
Family_USA_Born	0.28	0.45	0.14	0.35	0.73	0.44	0.20	0.40	0.18	0.38
Puerto_Rican					0.10	0.31				
Mexican					0.51	0.50				
Cuban					0.09	0.28				
eduPGS	0.00	1.00	-1.33	0.57	-0.21	0.76	-0.36	0.85	0.50	0.77
N	10370		1690		2021		748		5911	

Tab S2. Means, standard deviations, and correlations with confidence intervals for the SES indicators and  ${\it g}$ 

Table S2. Means, standard deviations, and correlations with confidence intervals for the SES indicators and g (N = 10370).

Variable	M	SD 1	2	3	4	5	6	7	8	9	10	11	12	13
1. Financial_ Adversity	0.00	1.00												
2. ADI	0.00	1.00 .29** [.27, .31]												
3. Neighborhood_ Safety	0.00	1.00 .25** [.23, .27]	.35**											
4. Education	0.00	1.00 .29** [.27, .31]	.42** [.40, .43]	.29** [.28, .31]										
5. Income	0.00	1.00 .35** [.33, .36]	.51** [.50, .53]	.32** [.31, .34]	.57** [.56, .59]									
6. Marital_ Status	0.66	0.47 .27**	.31**	.24**	.34**	.47** [.45, .48]								
7. Employment_ Status	0.91	0.29 .15**	.18**	.15**	.27**	.27**	.29**							
8. SES	0.00	1.00 .56** [.54, .57]	.70** [.69, .71]	.56** [.55, .57]	.73** [.72, .74]	.81** [.80, .82]	.65** [.64, .66]	.46** [.45, .48]						
9. European	0.75	0.33 .27** [.25, .29]	.42** [.40, .43]	.34** [.32, .36]	.44** [.42, .45]	.45** [.43, .46]	.42** [.40, .43]	.23** [.21, .24]	.57** [.56, .59]					
10. African	0.18	0.3125** [27, -	44** [45, -	30** [31, -	29** [31, -	36** [38, -	39** [41, -	20** [44, -	50** [51, -	89** [90, -				
11. Amerindian	0.06	0.1407**	-0.01 [03, .01]	13** [15, -	38** [39, -	23** [25, -	10** [12, -	08** [10, -	23** [45, -	31** [33, -	13** [15, -			
12. South_Asian	0.00		0 .03** ] [.01, .05]	-0.01 [02, .01]						03** [05, -	04** [06, -	.04** [.02, .06]		
13. East_Asian	0.01	0.03 -0.0 [02, .01	1 .03** ] [.01, .05]	-0.02 [04, .00]	-0.02 [04, .00]	!02* [04, -	-0.02 [04, .00]		-0.01 [03, .00]	12** [13, -		.07** [.05, .09]	.05** [.03, .07]	
14. g	0.00	1.00 .18** [.17, .20]	.26** [.24, .28]	.16** [.14, .18]	.36** [.34, .38]	.32** [.30, .34]	.25** [.23, .27]	.17** [.15, .19]	.38** [.37, .40]	.36** [.34, .37]	33** [35, -	10** [12, -	.03** [.01, .05]	0 [02, .02]

Note. M and SD are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). \*indicates p < 05. \*\*indicates p < 0.

Table S3. Means, standard deviation	s, and correla	tions	vith (	onfide	ence is	nterva	ls for th	e Full :	ample	(N = 10	)370)								-										
Variable	M	SD		1		2	3		1	5	6	7	8	9	10	- 11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. age	118.98	7.	49																										
2. g	0.00	1.		-0.0 03, .01																									
3. SES	0.00	1.			.384																								
4. frac White SIRE	0.73	0.					.51** [.50, .52	)																					
5. frac Black SIRE	0.20	0.	38 [-	-0.0 03, .01	131	**	46** [4/, -	82  82	**																				
6. frac Native American SIRE	0.02	0.	11 -	03*	02 [04	*	08** [10, -	18 [20	**	05**  0/, -																			
7. frac NOC SIRE	0.06	0.	23	02, .02	007	** *, -	15** [1/, -	41 [43	**	13**	04** [06, -																		
8. European	0.75	0.	33	0.0 01, .03	1 .36	.37]	.57** [.56, .59	.86*	* .86]	83**  84, -	03** [uɔ, -	19** [21, -																	
9. African	0.18	0.	31	-0.0 03, .01	133	**	50** [51, -	83 [84	**	.95** [.95, .96]	-0.01 [03, .01]	03** [uɔ, -	89** [90, -																
10. Amerindian	0.06	0.	14	-0.0 03, .01	110	** :, -	23** [25, -	13 [15	**	17**  19, -	.08**	.49** [.48, .50]	31** [33, -	13** [12, -															
11. South Asian	0.00	0.	02	0.0	1 .03*	.05]	[02, .02	0 -	0.02	04**  06, -	.05**	.08**	03**	04**	.04**														
12. East Asian	0.01	0.	03		0	0	-0.0	107	**		0.02	.13**	12**	(	07**	.05**													
13. State racism	0.00	1.	00	02, .02	0 -	-0.01	19** [21, -	=.10 [=.12	**	.20**	.06**	17** [18, -	10** [12, -	.21**	-,22** [24, -	03** [05, -	05**												
14. discrim fact	0.00	1.	00 I-	03**	20  22	**	24**	23  20	**	.23**	.04**	.03**	25**	.24**	.05**	-0.01	[02, .02]	0.06**											
15. Ethnic attachment	0.00	1.	00	0.0	103	**	09** [11, -	20 	**	.18**	.04**	.06**	23** [25,-	.19**	.09**	.02*	.03**	0.0	1 .07**										
16. Skin color	0.00						47** [48, -	- 68	**	63**	03**	20**	- 80**	68**	32**	.05**	08**	05**	20**	.19**									
17. P Brown Eye	0.56				222		39** [40, -									.06**					.74** [.73, .75]								
18. P Intermediate Eye	0.08	0.	07	02, .02	0 .17	.19]	.28**									03** [04, -													
19. PBlueEye	0.36	0.														05** [07, -				16** [18, -		98** [98, -	-0.01 [03, .01]						
20. P Black Hair	0.23	0.	24	-0.0 04, .00	229	je:e i, -	48** [49, -	=.69 [70	**										.20**	.20**	.79** [.79, .80]	.73**	45**  46, -	66** [b/, -					
21. P Brown Hair	0.46	0.																					.32** [.30, .33]		09** [11, -				
22. P Red Blond Hair	0.30	0.																					.17**			56** [-3/,-			
23. Child US Born	0.98															03** [05, -							.04**	.06**	08**	0.0	1 .06**		
24. Immierant Family	0.28	0.																					09** [10, -	19** [21, -	.20**	.03**	18** [20,-	*.21** [22, -	
25. eduPGS	0.00	1.	00		0 .38*			.62*	*	62**	04**	09**	.70**	67**	15**	0.01	03**	13**	24**	12**	56**	45**	33**	.40**	56**	.03**	.45**		-0.0

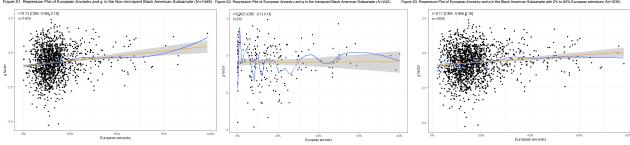
Table S4. Means, standard deviations, and correlations with confidence intervals for the Black subsample (N = 1690)

Variable	М	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. age	118.90	7.28																								
2. g	-0.69	1.07	-0.01																							
3. SES	-0.98		)5* 10, .10]	.23**																						
4. frac White SIRE	0.00	0.00 N	IA NA, NA]	NA [NA, NA]	NA [NA, NA	1																				
5. frac Black SIRE	1.00				[05, .05	0 NA ] [NA, NA]																				
6. frac Native American SIRE	0.00	0.00 N	IA ia, naj	NA [na, na]	NA [NA, NA	NA   [NA, NA]	NA [NA, NA	1																		
7. frac NOC SIRE	0.00	0.04	-0.01 .06, .03]	0.03	[05, .05	0 NA 1 [NA, NA]	-1.00**	NA [NA, NA	ı																	
8. European	0.16	0.11			.06* [.01, .10]	NA [NA, NA]		0 NA ] [NA, NA	( [05, .05]																	
9. African	0.82	0.11	05, .05]	09** [14, -	06** [11, -	NA [NA, NA]	[05, .05	0 NA 1 INA, NA	( [05, .05]	96** [9/, -																
10. Amerindian	0.01	0.02	0.02	05* [10, -		1 NA ] [NA, NA]		1 NA ] [NA, NA	-0.01 [06, .04]		)19**  23, -															
11. South Asian	0.00	0.01	0.01	-0.01 [06, .04]	0.0	5 NA ] [NA, NA]	0.0	1 NA 1 [NA, NA	-0.01	.18**	25** [29, -	[05, .05]	1													
12. East Asian	0.01	0.02	-0.02 .07, .03]	0.02	.06* [.01, .10]	NA [NA, NA]	0.0	1 NA ] [NA, NA	-0.01 [05, .04]	=0.01 [06, .03]	117** [21, -	0.01	0.05													
13. State racism	0.44	0.93	0.02	0.02	15** [20, -	NA [NA, NA]	-0.0 [06, .04	1 NA 1 [NA, NA	0.01	10** [15, -	.11**	-0.01 [06, .03]	06** [11, -	07** [12, -												
14. discrim fact	0.47	1.24	-0.03 .08, .01]	07**  12, -	-0.0 [08, .02	3 NA ] [NA, NA]	0.0	1 NA ] [NA, NA	-0.01 [06, .04]	-,09** [14, -	.07**	0.04 [00, .09]	05* [10, -	.05* [.00, .10]	-0.05 [09, .00]											
15. Ethnic attachment	0.38				.14**			5 NA   [NA, NA				3 0.02														
16. Skin color	1.32	0.42	0 .05, .05]	-0.01 [06, .03]	07** [11, -	NA [NA, NA]	0.0	1 NA ] [NA, NA	-0.01 [06, .04]	-,33** [5/, -	.32** [.27, .36]	0.01	-0.01	0 [05, .04]	-0.01 [05, .04]	0.04	-0.01 [05, .04]									
17. P Brown Eye	0.97	0.10	0.02	-0.05 [09, .00]	[05, .05	0 NA   [NA, NA]	-0.0 [05, .04	1 NA ] [NA, NA	0.01	33** [38, -	.32**	-0.01 [06, .04]	[04, .05]	0.02	0.02	0 [05, .05]	[05, .05]	.36** [.32, .40]								
18. P Intermediate Eye	0.02	0.03	-0.01 .05, .04]	0.05	-0.0 [07, .03	2 NA ] [NA, NA]	[05, .05	0 NA ] [NA, NA	( [05, .05]	.40** [.36, .44]	38** [42, -	[05, .05]	0.01	-0.02 [06, .03]	-0.02 [07, .02]	-0.01 [06, .03]	0 [04, .05]	43** [4/, -	77** [/y, -							
19. PBlueEye	0.02	0.08	-0.02 .07, .03]	0.04	0.0	1 NA   [NA, NA]	0.0	1 NA 1 [NA, NA	-0.01 [05, .04]	.28** [.23, .32]	26** [31, -	0.01	-0.01	-0.02 [06, .03]	-0.02 [07, .03]	0 [04, .05]	[05, .05]	30** [54, -		.62** [.59, .64]						
20. P Black Hair	0.55	0.16	-0.02 .07, .03]	-0.03 [08, .02]	[03, .06	1 NA   [NA, NA]	-0.0 [07, .02	3 NA ] [NA, NA	0.03	39** [43, -	.37** [.33, .41]	-0.01 [05, .04]	-0.01	0.03	-0.02 [07, .03]	.05* [.01, .10]	0.04	.45** [.41, .49]	.49** [.45, .52]	67**	38** [42, -					
21. P Brown Hair	0.43	0.13	0.02	0.01	-0.0 [08, .01	4 NA   [NA, NA]	0.0	3 NA 1 [NA, NA	-0.03	.32** [.27, .36]	30** [34, -	0.01	[04, .05]	-0.03 [08, .02]	0.03	06* [10, -	-0.05 [09, .00]	35** [39, -	34** [58, -	.56** [.52, .59]	.24** [.19, .28]	96** [56, -				
22. P Red Blond Hair	0.02	0.05	05, .05]	.06* [.01, .11]	.05* [.01, .10]	NA [NA, NA]	0.0	1 NA 1 [NA, NA	-0.01 [05, .04]	.41** [.37, .45]	39** [43, -	[05, .04]	0.01	[05, .05]	-0.02 [07, .03]	-0.02 [07, .03]	-0.02 [07, .03]	51** [55, -	65** [b/, -	.65** [.62, .68]	.58** [.55, .61]	64** [b/, -	.40** [.36, .44]			
23. Child US Born	0.98	0.15	0.03	0 [05, .05]	-,14** [-,18, -	NA [NA, NA]	-0.0 [05, .04	1 NA ] [NA, NA	0.01	[05, .05]	0.0	1 .05* ] [.00, .10]	12** [16, -	07** [12, -	.11** [.06, .15]	-0.04 [09, .01]	0.03	-0.01 [06, .03]	[05, .05]	[01, .09]	-0.02 [06, .03]	09** [14, -	.11** [.06, .16]	-0.01 [05, .04]		
24. Immigrant Family	0.14	0.35 .0	)6* 01, .10]	.08** [.04, .13]	.22**	NA [NA, NA]	0.0	2 NA ] [NA, NA	-0.02 [06, .03]	-0.03	3 0.0 [03, .07	2 -0.01	1 .10** [.05, .14]	0.04	-20**	0.03	.09** [.04, .13]	0.05	0.02	-0.02 [07, .03]	-0.01 [06, .03]	.08** [.04, .13]	09** [1.5, -	-0.04 [08, .01]	27** [51, -	
25. eduPGS	-1.33	0.57	.05, .04]	.11** [.06, .16]	.12**	NA [NA, NA]	-0.0 [05, .04	1 NA ] [NA, NA	0.01	.38** [.34, .42]	37** [41, -	-0.04 [09, .00]	1 .13** [.08, .17]	.05* [.00, .10]	-0.04 [08, .01]	08** [13, -	.07** [.02, .12]	19** [24, -	11** [16, -	.13** [.08, .18]	.09** [.05, .14]	13** [18, -	.10** [.05, .15]	.16** [.12, .21]	-0.04 [08, .01]	0.04

Tab S4: Black subgroup performance
Table S8: Descriptive statistics for Black subsamples by mierant status

	N	Q		SES		African a	ncestry	European	ancestry	Amerindia	an ancestry	
		м	SD	M	SD	M	SD	M	SD	M	SD	
Black Family USA Born	1475	-0.74	1.07	-1.06	0.89	0.82	0.10	0.17	0.10	0.01	0.02	
Black Family Immigrant	215	-0.36	1.02	-0.38	0.99	0.82	0.17	0.16	0.16	0.01	0.03	
Black African born parents	60	-0.28	0.94	-0.04	0.85	0.87	0.22	0.13	0.21	0.00	0.00	
Black African born parents	43	-0.23	0.97	-0.01	0.94	0.99	0.02	0.01	0.02	0.00	0.00	
and > 90% African ancestry												
Black West Indian born parents	44	-0.39	1.05	-0.68	0.96	0.86	0.08	0.12	0.08	0.01	0.02	





Tab S5: Admixture-regression results for Whites

Table S9. Regression Results for the Effect of Genetic Ancestry on g among White Americans (*N*=5911). Figure S4. Regression Plot of European Ancestry and *g* in the White American Subsample (N=5911).

	Mo	odel 1	Mo	odel 2	Mo	del 3	Mo	del 4
Predictors	b	p	b	p	b	p	b	р
(Intercept)	0.13 (0.11)	0.252	0.13 (0.11)	0.235	0.13 (0.11)	0.261	-0.03 (0.11)	0.778
Amerindian_ancestry	-1.84 (0.38)	<0.001	-1.83 (0.38)	<0.001	-1.63 (0.39)	<0.001	-0.96 (0.38)	0.012
African_ancestry	-1.71 (0.55)	0.002	-1.70 (0.55)	0.002	-1.59 (0.55)	0.004	-0.85 (0.53)	0.110
East_Asian_ancestry	-0.16 (0.56)	0.772	-0.16 (0.56)	0.776	-0.17 (0.56)	0.759	0.05 (0.54)	0.921
South_Asian_ancestry	0.54 (1.04)	0.604	0.51 (1.04)	0.623	0.47 (1.04)	0.654	0.06 (1.01)	0.950
Child_US_Born	0.21 (0.10)	0.034	0.21 (0.10)	0.035	0.19 (0.10)	0.055	0.21 (0.10)	0.028
Immigrant_Family	0.09 (0.03)	0.004	0.09 (0.03)	0.005	0.09 (0.03)	0.006	0.05 (0.03)	0.134
Ethnic_attachment			0.02 (0.01)	0.151	0.02 (0.01)	0.134	0.01 (0.01)	0.212
State_racism					0.03 (0.04)	0.437	0.09 (0.05)	0.058
discrim_fact					-0.13 (0.01)	<0.001	-0.09 (0.01)	<0.001
Skin_color					-0.00 (0.02)	0.882	-0.00 (0.02)	0.980
P_Brown_Eye					0.04 (0.04)	0.350	0.01 (0.04)	0.730
P_Intermediate_Eye					-0.20 (0.17)	0.227	-0.14 (0.16)	0.373
P_Black_Hair					-0.16 (0.15)	0.283	-0.15 (0.14)	0.285
P_Brown_Hair					0.04 (0.07)	0.517	0.06 (0.07)	0.344
SES							0.32 (0.02)	<0.001
Random Effects								
$\sigma^2$	0.36		0.36		0.36		0.36	
τ <sub>00</sub>	0.35 site_id_ 0.04 site_id		0.35 site_id 0.04 site_id	_l:rel_family_id	0.33 site_id		0.29 site_id_ 0.06 site_id_	
ICC	0.52	-	0.52	-	0.51	-	0.49	-
N	22 site_id_1		22 site_id_1		22 site_id_1		22 site_id_1	
	4835 rel_far	mily_id	4835 <sub>rel_far</sub>	mily_id	4835 <sub>rel_far</sub>	nily_id	4835 <sub>rel_far</sub>	nily_id
Observations	5911		5911		5911		5911	
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.008 / 0.52	25	0.008 / 0.5	25	0.024 / 0.5	21	0.093 / 0.53	36

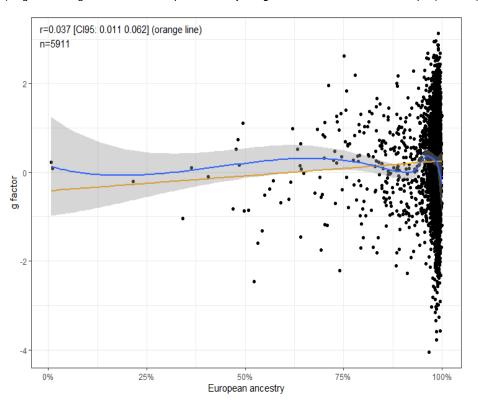


Table Starbe, Valley and solven well-stilled in the Conference SEE group.

PER Controls Complete Starb Memories SEE group.

PER Controls Complete Starb Memories SEE group.

PER Controls Complete Starb Memories Starb Starber Starbe

No.4		Winnels	- 04		White	_		Blok Street		White	Hat Happi Oler Wile	Parkers.	
, ,		,		,			Predictor	1 , 1			Feedor 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	(beauty)	000 150 03 400 03 400 0 030 000 033 00
432 41 615	0 4	13 6040	0.00	0.361	401 0.17 6110		(Amongs)	436 8180 625 4	ME 012 02	421 1302		+600	\$15 \$80 \$21 AND \$21 AND \$2
017 66					121 -080		10/50	020 000 029 d			60°C 60°C 60°C 60°C 60°C 60°C	Alex	AND
940	- 9	10	(0.00)		(0.00)		March .	546 616	(815)	600	Allon 47 488 AN 107 54 109 48 102 870 687 887 887		630 636 636 63
1732 68 (842)	P 10	21 000	(8.73)	109	2024 64E (0836)		PCI	1100 0200 EW 0	141 1536 646	7120 KMB	Ber, Arm 500 670 677 536 686 600 603 603 675 686 675 675	Ear,/one	470 877 420 466 161 468 40 873 686 675 68
-3630 66 -0160	N -17		410		-6.00 0.00 (0.70)		PCT	OH AND -2110 B	MI 210 00	4.79 1207	\$66,000 0.00 0.00 0.00 0.00 0.00 0.00 0.0	Strift_Asia.	339 830 12 400 10 608 61 330 330 330 670 61
1949 42 07475	2 8	6547	29.54		10.75 845		80	411 190 179 0	471 N.O. 68	9 00 000s	Assessine -0.0 800 -0.0 -0.00 -0.0 800 -0.0 800 -0.00 -0.00 -0.00	Approxima	340 MM AN AND A17 MM A
420 62			000		J.D. 419			33.95 01.16	(7.25)	0.50	Radio (Str.)	Factor Effects	
(817)	- 0	(2)	(624)		(8.80)		PCI	417 Edit 177 0	001 004 0M	4 135 1226	Name of State of Stat	2	THE RM LOW CO.
499 18	9 1	(9 6394 (6)	679	5.128	451 880 (0.24)		PCS	279 8990 638 0	OR 1476 600	110 100	10 mai) 10 mai) 10 mai) 10 mai)	100	the six of the
3.6 07 0075	9 4	10 0.00	4.89	0.362	251 948 510		P/A	201 000 071 0			N Basic Basic Basic Basic Waters Physics Waters Waters	Constant	
-179 64		ei 430			AN 185			(5.26) (8.95)			Owner 16 35 16 50 Security (minors) 100 104 100 100 100 100 100 100	Magnal B <sup>2</sup> Coults	mod 2" 525 525 525 539 523 525 530 53
0.76	0	14)	0.10		0.80		PCI	230 03'W 1837 0 0A0 0240	263 200 0.T	0.400 E400			
221 04 090	4 6	10 006	0.46	144	1.04 0.00 6.16		PCI	130 000 -130 0	240 334 08	0.20 EMD			
489 01	F 2	0.00	446	0.00	677 046 670		800	(454) (415) 207 (450 (417 (4					
377 67					140 190		10.5	(4.70) (3.75)	(8.30)	(2.30)			
0.90	0.	10	(495)		(6.76)		PCH	454 6719 122 0 120 0.00					
410 03 920	7 4	62 0.280 96	0.50	0.904	0.00 0.00 0.00		PS II	495 1207 446 9	10 40 00	226 546			
436 66	N 4	67.0	4660	0.268	496 626		FC12	AT 100 AT 0					
.170 64	W 4	E 080			480 026		KII	A75 0300 A73 0	(D.Z0	(8.50) E.358			
0.80	-		(8.14)		(414)		PCU	024 026 026 0 020 020	89 19 18	2 285 1396			
450 67 070	n 4	10 129	0.40	0,002	235 439 GBB		PS14	410 070 417 0	m: 011 0W	1 20 100			
464 63	8 4	34 0.512	3.6	0.402	-236 9.19 0.80			(4.80) (5.90)	(8.26)	(3.80)			
.19 04	0 4	se sen	1.24	9.60	410 100		PCII	490 0128 421 6 CARD (A1D	0.20	0.00 000			
0.10	9.	10	(854)		(2.16)		PCH .	437 466 757 6	MT 131 0.0				
477 97 G80	H (	930	1.36		0.20		KIT	49 MM -1H 0	(33b)	0.50			
CID 04	F 0	989	281	0,376	431 000 6/2)		PCI	(3.86) (9.26)	0.15	(IF)			
010 01					476 070		PCH	436 0879 147 0 040 040	110 100 100	510 BHS			
0.50	- 6	569	(9.00)		(14)		PSW	33 486 IN 9	NO 121 CH	130 6326			
231 03 (24b)	B (4	27 0.334	3.83	0.349	A32 080 6.140			date date					
							PC30	3.0 6.00 3.60 0 9.20 0.20	(1.12)	(1.00) 1.70			
0.98	0.41		141				Rondon Lifera						
65 autota	0.61	exit.)				×		Distance of Control	1.07	0.09			
6.58			3.45		50		100 100	(00 000 mg/s)	508	000 (m)4)			
Baco	27.0	00	Sauce.		100		8	na Past		Feet			
147 of Jank of 197	2101	NUMBER OF	Self-pi_feeb		CT nc bestrue		Orservations	170 1714	307	1674			
180 000 030	200	19.200	1120-1170		HD 430		Named E. Continued E.	SEE   0.001   0.001   0.001	5139 9 182	0.005/0.028			

Tab 7: Regression Results for the Effect of edu/IQPGS and Ancestry on Cognitive Ability in the SIRE subsamples 2

Table S15: Recression Results for the Effect of edu/IOPGS and Ancestry on Coontive Ability in the SIRE Table S16: Effects (b) of Amerindian and African Ancestry on g in multilevel models with environmental controls (Model 4/3) and multilevel models with environmental controls and also IOPGS (Model 5).

	Black		Hispanic		Other		White		Full	
	Model 4	Model 5	Model 4	Model 5	Model 4	Model 5	Model 4	Model 5	Model 4	Model 5
Amerindian	-3.46	-3.15	-0.86	-0.55	-1.09	-0.55	-0.96	-0.65	-0.86	-0.58
African	-0.89	-0.69	-0.58	-0.16	-1.08	-0.48	-0.85	-0.32	-0.79	-0.40

k		panie		her		hite
P	å	P	b		å	P
0.726	-0.15 (0.17)	0.368	-0.16 (0.38)	0.679	-0.08 (0.11)	0.497
0.009	-0.55 (0.19)	0.003	-0.55 (0.51)	0.279	-0.65 (0.37)	0.083
0.023	-0.16 (0.26)	0.535	-0.48 (0.31)	0.118	-0.32 (0.52)	0.538
0.879	-1.16 (0.89)	0.190	0.75 (0.58)	0.192	0.15 (0.53)	0.769
0.089	4.10 (3.29)	0.212	1.45 (0.78)	0.065	0.06 (0.99)	0.949
0.062	0.11 (0.09)	0.231	0.31 (0.29)	0.294	0.17 (0.09)	0.065
0.067	0.11 (0.05)	0.050	-0.02 (0.11)	0.968	(0.03)	0.228
0.002	(0.02)	0.062	(0.04)	0.157	(0.01)	0.400
0.460	(0.05)	0.188	0.09 (0.07)	0.207	(0.05)	0.042
0.036	-0.11 (0.02)	<0.001	-0.08 (0.03)	0.018	-0.08 (0.01)	<0.001
0.087	-0.07 (0.04)	0.055	0.01 (0.06)	0.819	(0.02)	0.759
0.424	0.26 (0.11)	0.019	(0.17)	0.564	(0.04)	0.810
0.686	0.36 (0.42)	0.399	0.70 (0.72)	0.335	-0.13 (0.16)	0.415
0.875	-0.08 (0.18)	0.666	0.26 (0.32)	0.421	-0.18 (0.14)	0.211
0.754	0.09 (0.19)	0.626	-0.21 (0.30)	0.481	(0.05)	0.268
<0.001	(0.03)	<0.001	(0.04)	<0.001	(0.02)	<0.001
0.026	(0.03)	<0.001	0.28 (0.06)	<0.001	(0.01)	<0.001
	-0.10 (0.06)	0.091				
	-0.13 (0.09)	0.130				
	(0.09)	0.002				
	0.21 (0.15)	0.144	0.20 (0.32)	0.523		
	0.06 (0.16)	0.715	-0.21 (0.23)	0.355		
	-0.03 (0.06)	0.620	-0.07 (0.19)	0.694		
	0.40		0.63		0.36	
el_flenity_id	0.40 see_st_	rel_family_id	0.34 sim_id_	l.rel_family_id	0.27 single	tret_femily
	0.03 sine_st_ 0.52		0.08 sinc_id_ 0.40	1	0.06 ma_st	
	44 mm_163		44 mod.)		44 sau (d.)	
y_st	1794 pd_fan	nky_id	Date religions	504	4835 pt_fa	nty_id
	y_14	22 <sub>simulat</sub> ) 1794 <sub>militar</sub> 2021	22 majdj 1794 mljandyjd	22 mejaj 22 mejaj rja 1794 mijambyja 644 mijami 2021 748	22 um_jd_j 22 um_jd_j y_jd 1794 nd_famity_jd 644 nd_famity_jd 2021 748	22 um.jd.j 22 um.jd.j 22 um.jd.j 22 um.jd.j y.jd 1794 nd_famity.jd 644 nd_famity.jd 4835 nd_famity.jd 4835 nd_famity.jd 5911

# Tab S8: PseudoPGS results

Table S17: Regression Results for the Effect of PseudoPGS and Ancestry on Cognitive Ability in the Full sample. □

	Mo	del 1	Mo	del 2
Predictors	ь	p	b	р
(Intercept)	0.17 (0.07)	0.018	-0.02 (0.08)	0.784
frac_Black_SIRE	0.01 (0.09)	0.942	0.04 (0.08)	0.671
frac_Native_American_SIRE	-0.09 (0.09)	0.332	0.00 (0.09)	0.975
frac_NOC_SIRE	-0.08 (0.05)	0.110		
Hispanic	0.02 (0.04)	0.584	0.06 (0.04)	0.122
Child_US_Born	0.13 (0.06)	0.029	0.17 (0.06)	0.004
Immigrant_Family	0.15 (0.03)	<0.001	0.09 (0.02)	<0.001
Amerindian_ancestry	-1.55 (0.12)	<0.001	-0.84 (0.13)	<0.001
African_ancestry	-1.28 (0.11)	<0.001	-0.76 (0.11)	<0.001
East_Asian_ancestry	0.14 (0.33)	0.665	0.23 (0.32)	0.477
South_Asian_ancestry	0.39 (0.56)	0.485	0.80 (0.54)	0.139
PseudoPGS_1_z	0.00 (0.01)	0.958	-0.00 (0.01)	0.915
PseudoPGS_2_z	0.00 (0.01)	0.751	-0.00 (0.01)	0.818
PseudoPGS_3_z	0.01 (0.01)	0.198	0.01 (0.01)	0.227
PseudoPGS_4_z	0.02 (0.01)	0.075	0.01 (0.01)	0.160
PseudoPGS_5_z	-0.01 (0.01)	0.538	-0.00 (0.01)	0.604
PseudoPGS_6_z	-0.02 (0.01)	0.061	-0.02 (0.01)	0.074
PseudoPGS_7_z	-0.01 (0.01)	0.203	-0.01 (0.01)	0.166
PseudoPGS_8_z	-0.01 (0.01)	0.381	-0.01 (0.01)	0.310

PseudoPGS_9_z	0.03 (0.01)	0.005	0.03 (0.01)	0.002
PseudoPGS_10_z	-0.01 ( (0.01)	0.403	-0.00 (0.01)	0.599
frac_Other_SIRE			-0.01 (0.05)	0.863
State_racism			0.08 (0.04)	0.045
discrim_fact			-0.09 (0.01)	<0.001
Skin_color			-0.01 (0.02)	0.557
P_Brown_Eye			0.03 (0.04)	0.387
P_Intermediate_Eye			-0.07 (0.15)	0.662
P_Black_Hair			-0.05 (0.08)	0.535
P_Brown_Hair			0.03 (0.06)	0.626
SES			0.31 (0.01)	<0.001
Random Effects				
$\sigma^2$	0.41		0.42	
τ <sub>00</sub>	0.43 site_id_l:rel_i	family_id	0.35 site_id_1	rel_family_id
	0.03 site_id_1		0.05 site_id_1	
ICC	0.53		0.49	
N	22 site_id_1		22 site_id_1	
	8672 rel_family_i	d	8672 rel_fam	ily_id
Observations	10370		10370	
	0.179 / 0.610		0.247 / 0.616	

#### 1. African Americans

```
Linear mixed model fit by REML ['ImerMod']
Formula: g ~ Amerindian_ancestry + African_ancestry + East_Asian_ancestry + South_Asian_ancestry + Child_US_Born + Immigrant_Family + Ethnic_attachment + State_racism + discrim_fact + Skin_color + P_Brown_Eye + P_Intermediate_Eye + P_Black_Hair + P_Brown_Hair +
     SES + eduPGS + (1 | site_id_l) + (1 | site_id_l:rel_family_id)
    Data: merged_df_NH_B
REML criterion at convergence: 4809.8
Scaled residuals:
Min 1Q Median 3Q Max
-3.09497 -0.48359 0.02515 0.51826 2.93191
Random effects:
                               Name
                                                  Variance Std.Dev.
 Groups
 site_id_l:rel_family_id (Intercept) 0.45944 0.6778
site_id_l (Intercept) 0.03674 0.1917
                                             0.56367 0.7508
Number of obs: 1690, groups: site_id_l:rel_family_id, 1437; site_id_l, 22
Fixed effects:
                      Estimate Std. Error t value
(Intercept) 0.25416 0.72437 0.351
Amerindian_ancestry -3.14677 1.20512 -2.611
African_ancestry -0.68527 0.30081 -2.278
East_Asian_ancestry -0.22810 1.50213 -0.152
South_Asian_ancestry -5.64035 3.31188 -1.703
Child_US_Born
Immigrant_Family
                                eduPGS
                               0.10766 0.04835 2.227
 Correlation matrix not shown by default, as p = 17 > 12.
Use print(x, correlation=TRUE) or
                        if you need it
    vcov(x)
> coeffvec <- coef(fit.p1)
> varcov <- vcov(fit.p1, full=FALSE)
> varcov
17 x 17 Matrix of class "dgeMatrix"
 [.1] [.2] [.3] [.4] [.5] [.6] [.7] [.8] [.9] [.10] [.11] [.11] [.12] [.3.47150e-01 -0.0483488673 -0.0417765059 -5.867230e-02 -0.171332977 -2.244147e-02 -2.093430e-03 4.959454e-04 2.666032e-04 -2.920766e-05 8.444476e-03 [2.] -4.834887e-02 1.4523191351 0.0971993257 3.725762e-02 0.209445280 -9.945977e-03 -2.153429e-03 -1.247248e-03 3.547406e-04 -1.105404e-03 -2.424662e-03 [3.] -4.177651e-02 0.0971993257 0.0904847465 7.245479e-02 0.244631921 1.163761e-04 -1.268606e-03 -4.059461e-04 -1.165533e-03 -2.759315e-04 -1.750503e-03
 [4] -5.867230e-02 0.0372576174 0.0724547882 2.256401e+00 0.016084862 8.965490e-03 -5.929947e-04 -1.369701e-05 1.130884e-03 -1.571809e-03 8.599023e-06 [5.] -1.713330e-01 0.2094452799 0.2446319205 1.608486e-02 10.968577029 5.827275e-02 -1.431241e-02 -1.023744e-03 2.146097e-03 1.830553e-03 -6.240272e-03
 4.959454e-04 -0.0012472478 -0.0004059461 -1.369701e-05 -0.001023744 -2.382607e-04 -1.141254e-04 6.492771e-04 -1.329850e-05 -1.891710e-05 3.580453e-05
[8], 1.9594546-04 -0.00124/24/6 -0.0004059461 -1.3597016-05 -0.001025744 -2.3520076-04 -1.1412946-04 -0.352076-04 -1.328506-05 2.8722326-03 1.8951006-05 7.4438336-05 [10], 2-2.9207666-05 -0.0011054041 -0.0002759315 -1.5718096-03 0.001830553 1.0204196-04 -1.7948906-05 -1.8917106-05 3.9997246-04 -1.3950786-05 [11], 8.4444766-03 -0.0024246618 -0.0017505031 8.5990236-06 -0.006240272 -2.5143436-05 -1.5521556-04 3.5804536-05 7.4438336-05 -1.3950786-05 5.1733316-03 [12], 6.3263396-02 0.0053549917 -0.0007083358 -9.9587996-03 -0.038513668 -1.6923996-03 -8.4638886-05 5.0679846-05 2.5763276-04 7.9381256-05 9.0760196-04
[13,] -6.323981e-01 0.0604644800 0.0483744602 3.078529e-02 0.024603704 -4.840579e-03 -5.053843e-03 -1.483718e-03 1.769622e-03 -5.252444e-04 9.426405e-03
[14,] -4.085245e-01 -0.0310465573 -0.0283470751 -1.912117e-02 -0.085906688 -4.407336e-03 -1.294662e-03 -5.291197e-04 -2.338622e-05 -3.260798e-05 -1.540993e-02 [15,] -4.270531e-01 -0.0251840190 -0.0210169518 -7.811174e-03 -0.073701180 -7.659365e-03 -6.483522e-04 -2.426034e-04 -2.484527e-04 9.268599e-05 -1.366713e-02
[16] -1.320957e-03 0.0009581751 0.0004065839 -1.498902e-03 0.002037185 5.255491e-04 -3.647817e-04 -8.787560e-05 1.246713e-04 1.908334e-05 9.308669e-05 [17,] -1.234413e-03 0.0067678031 0.0045066814 1.240984e-03 -0.003973998 1.212038e-04 -1.078277e-04 -8.021109e-05 -4.719278e-06 5.035844e-05 2.796188e-04
 [.12] [.13] [.14] [.15] [.16] [.17] [1.] -6.326339e-02 -0.6323980722 -4.085245e-01 -4.270531e-01 -1.320957e-03 -1.234413e-03 [2.] 5.354992e-03 0.0604644800 -3.104656e-02 -2.518402e-02 9.581751e-04 6.767803e-03
 [3] -7.083358e-04 0.0483744602 -2.834708e-02 -2.101695e-02 4.065839e-04 4.506681e-03 [4] -9.958799e-03 0.0307852865 -1.912117e-02 -7.811174e-03 -1.498902e-03 1.240984e-03
 [5], -3.851387e-02 0.0246037037 -8.590669e-02 -7.370118e-02 2.037185e-03 -3.973998e-03 [6], -1.692399e-03 -0.0048405790 -4.407336e-03 -7.659365e-03 5.255491e-04 1.212038e-04
        -8.463888e-05 -0.0050538434 -1.294662e-03 -6.483522e-04 -3.647817e-04 -1.078277e-04
[17] -3.405066-05 -0.0014837184 -5.291197e-04 -2.426034e-04 -3.047617e-04 -1.076277e-04 [8.] 5.067984e-05 -0.0014837184 -5.291197e-04 -2.426034e-04 -8.787560e-05 -8.021109e-05 [9.] 2.576327e-04 0.0017696221 -2.338622e-05 -2.484527e-04 1.246713e-04 -4.719278e-06 [10.] 7.938125e-05 -0.0005252444 -3.260798e-05 9.268599e-05 1.908334e-05 5.035844e-05 [11.] 9.076019e-04 0.0094264051 -1.540993e-02 -1.366713e-02 9.308669e-05 2.796188e-04
[12,] 1.698002e-01 0.4259192838 -1.019558e-01 -1.193283e-01 -7.164938e-05 -1.009277e-04 [13,] 4.259193e-01 2.9483784541 2.068591e-01 4.182633e-02 3.252560e-03 1.782964e-03
[14,] -1.019558e-01 0.2068591007 5.621212e-01 5.718523e-01 1.135744e-03 5.101581e-05 [15,] -1.193283e-01 0.0418263262 5.718523e-01 6.362338e-01 1.135786e-03 2.072545e-04 [16,] -7.164938e-05 0.0032525604 1.135744e-03 1.135786e-03 9.615101e-04 -9.291198e-05
2. Hispanics
Linear mixed model fit by REML ['ImerMod']
Formula: g ~ Amerindian_ancestry + African_ancestry + East_Asian_ancestry + South_Asian_ancestry + Child_US_Born + Immigrant_Family +

Mexican + Cuban + Puerto_Rican + frac_Black_SIRE + frac_Native_American_SIRE + frac_NOC_SIRE + Ethnic_attachment + State_racism + discrim_fact +

Skin_color + P_Brown_Eye + P_Intermediate_Eye + P_Black_Hair + P_Brown_Hair + SES + eduPGS + (1 | site_id_I) + (1 | site_id_I:rel_family_id)
    Data: merged_df_H
```

REML criterion at convergence: 5271.5

Scaled residuals

Min 1Q Median 3Q Max -2.82697 -0.45736 0.02203 0.48087 2.60846