Supplementary Information for:

Genetic and environmental contributions to IQ in adoptive and biological families with 30-year-old offspring

Abstract

While adoption studies have provided key insights into the influence of the familial environment on IQ scores of adolescents and children, few have followed adopted offspring long past the time spent living in the family home. To improve confidence about the extent to which shared environment exerts enduring effects on IQ, we estimated genetic and environmental effects on adulthood IQ in a unique sample of 486 biological and adoptive families. These families, tested previously on measures of IQ when offspring averaged age 15, were assessed a second time nearly two decades later (M offspring age = 32 years). We estimated the proportions of the variance in IQ attributable to environmentally mediated effects of parental IQs, sibling-specific shared environment, and gene-environment covariance to be .01 [95% CI .00, .02], .04 [95% CI .00, .15], and .03 [95% CI .00, .07] respectively; these components jointly accounted for 8 percent of the IQ variance in adulthood. The heritability was estimated to be .42 [95% CI .21, .64]. Together, these findings provide further evidence for the predominance of genetic influences on adult intelligence over any other systematic source of variation.

1. Methods

We used observed correlations to estimate the following parameters: a = the effect of the genetic score on the phenotype; q = the variance of the genetic score, i.e., the additive genetic variance; d^2 = dominance genetic variance; s^2 = variance of environmental factors shared by siblings reared together, other than the phenotypes of their parents; m = direct effect of maternal phenotype on offspring phenotype; p = direct effect of paternal phenotype on offspring phenotype; w = covariance between latent additive genetic and family environment factors; x = variance of the shared environment induced by parental phenotypes; μ = correlation between spouses. Because only one parent from each family was assessed on intake measures of IQ, we use a stand-in spousal correlation value derived from pairs of mothers and fathers from the Minnesota Twin Family Study, who were measured on the same version of the same Wechsler scales as our SIBS sample.

To estimate variance components, we adapted the Cascade model (Keller et al., 2009) to include adoptive relationships:

- Mother-father $= \mu$

- Biological mother-child $= \frac{1}{2}a(qa + w) + \frac{1}{2}a(qa + w)\mu + m + p\mu$

- Biological father-child $= \frac{1}{2}a(qa + w) + \frac{1}{2}a(qa + w)\mu + p + m\mu$

- Adoptive-adoptive siblings $= \frac{x+s^2}{1-2aw}$

- Adoptive-biological siblings $= \frac{x+s^2+aw}{\sqrt{1-2aw}}$

- Biological-biological siblings $= a^2(q - \frac{1}{2}) + \frac{1}{4}d^2 + 2aw + x + s^2$

- Adoptive mother-child $= \frac{m+p\mu}{\sqrt{1-2aw}}$

- Adoptive father-child $= \frac{p+m\mu}{\sqrt{1-2aw}}$

Note that if the variance of the phenotype has been set to one in biological individuals, the variance may be less than one in adopted individuals. We accounted for this by including the appropriate rescaling factor in any theoretical correlation involving an adopted individual.

After applying Fisher's z-transformation, we minimized the squared differences between the empirical correlations and the theoretical (model-predicted) correlations by adjusting the parameter estimates. Each term in the sum of squared differences was weighted by the reciprocal of its variance, $N - 3$, where N is the number of pairs in the correlation. To perform statistical inference, we took bootstrap resamples of our families and re-estimated the parameters each time.

We constrained the dominance genetic variance to equal zero. The point estimate of this parameter was zero for most phenotypes, but in bootstrap resampling it occasionally assumed unrealistically large values. Since there is compelling theory and evidence for most genetic variance being additive (Hill et al., 2008; Maki-Tanila & Hill, 2014) and dominance variance being negligible (Hivert et al., 2021; Pazokitoroudi et al., 2021), we decided to constrain this parameter to zero in order to improve statistical inference about other parameters. Note that a negligible value of the dominance genetic variance is perfectly consistent with declines in IQ as a result of inbreeding (Schull & Neel, 1965; Jensen, 1983; Joshi et al., 2015).

Decomposition of variance terms shown in main text Figure 1 and presented in main text Table 2 were therefore computed as follows:

- Heritability (A) $= h^2 = qa^2$

- Parental environment (F) $= m^2 + p^2 + 2mp\mu$

- Sibling environment (S) $= s^2$

- G–E covariance $= 2 \times aw$

- Non-shared environment = Normed as 1.0 – (A + S + F + G-E covariance)

Derived parameters (nonlinear constraints) shown in main text Figure 1 are computed thus:

- Variance of family environment $(x) = m^2\sigma^2 + p^2\sigma^2 + 2m\sigma^2\mu^2\sigma^2 = m^2 + p^2 + 2mp\mu$

- Variance of additive genetic effects $(q) = 1 + \mu(qa + w)^2$

- Gene-environment covariance $(w) = \frac{1}{2}(qa + w)m + \frac{1}{2}(qa + w)p + \frac{1}{2}(qa + w)\mu m + \frac{1}{2}(qa + w)\mu p$

2. Comparison of participating and non-participating offspring at follow-up 3

We conducted a comparison of participants with non-participants in the current wave based on measures taken at intake to evaluate the possibility of attrition effects. A comparison of participants with non-participants in the current wave based on measures taken at intake indicated no selection on family SES (standardized effects size $d < .10$ in absolute value) or parental characteristics, with minimal selection on offspring IQs. The largest difference was for intake IQ, which was about 2.5 points higher in current participants than in non-participants. Comparison of participants and non-participants is shown in Table S1.

3. Scale means

For each demographic criterion and IQ phenotype, means and standard deviations are computed separately for mothers, fathers and offspring and for both adoptive and biological families (Table S2).

4. ICAR-16 sample test

We use the ICAR-16 sample test as a measure of general cognitive ability in our sample. Detailed loadings and item analysis can be found in Condon & Revelle (2014),

Reliability indices. Detailed item analysis for the ICAR-16 sample test is presented in Condon & Revelle (2014; (Condon & Revelle, 2014)). The sample test consists of 16 items taken from the full 60-item ICAR test, each of which comprises one of four item types or subtests. These four subtests are summarized as letter and number sequences (LN), matrix reasoning (MR), 3D rotation (R3D) and verbal reasoning (VR). These comparisons are shown in Table S3.

Confirmatory factor analysis. We used confirmatory factor analysis to test both a single (g) and four-factor model of the ICAR-16. A four-factor model is composed of letter and number sequences, matrix reasoning, 3D rotation and verbal reasoning. This model was fit using the *lavaan* package in R with full-information maximum likelihood (FIML). Model fit was strong, with a Tucker-Lewis Index (TLI) of .978, Comparative Fit Index (CFI) of .982, Standardized Root Mean Square Residual (SRMR) of .024, and Root Mean Square Error of Approximation (RMSEA) of .021 (90% CI: .015, .028). As expected, the indicators all showed significant positive factor loadings, with standardized coefficients ranging from .20 to .73 (Table S4). The single-factor model was also strong, with TLI of .984, CFI of .987, SRMR of .022, and RMSEA of .018 (95% CI: .010, .025).

Additionally, we observed significant positive correlations among all four latent factors (Table S5), indicating that participants who showed high ability in one dimension were more likely to show high ability in the others as well. Taken together, these results are consistent with use of the ICAR-16 as a good short-form measure of cognitive ability, with the advantage of its short administration time outweighing its limitations in the context of this study.

5. Scale and demographic comparisons by ethnicity of adoptee

Majority of offspring in sample ($N = 1232$ total with valid intake IQ scores) report either white ($N = 662$) or Asian ($N = 460$) ethnicity. Mean differences and significance for each demographic criterion (highest degree completed, years of education, age, and family SES), IQ scores, and polygenic scores are reported below for white and Asian offspring (Table S6).

While no significant mean differences were detected for any IQ criterion, Asian offspring had a substantially higher mean PGS than did white offspring (Cohen's $d = .87$). The reason for this difference is unknown; possibilities include population stratification, allele frequency ascertainment bias, and true population difference. Allele frequency ascertainment bias can arise from systematic deviations from the expected theoretical result due to sampling processes of the genotyping chip used to ascertain SNPs and their population-specific allele frequencies; a lower frequency of the minor allele in East Asians can cause the appearance of a higher PGS without underlying predictive significance. Whatever the reason, it is not strictly relevant to our PGS analyses which are conducted separately for white biological and Asian adoptive offspring.

6. Scale and demographic comparisons by adoption status

Mean differences and significance for each demographic criterion (highest degree completed, years of education, age, and family SES), IQ scores, and polygenic scores are reported below for adopted and biological offspring (Table S7).

7. Intercorrelations among scales and demographic statistics

Tables of correlations among political attitude phenotypes, ICAR-16, age at follow-up 3, years of education, highest degree computed, and socioeconomic status (z-scores) were computed for all individuals in aggregate (Table S8), as well as separately for mothers, fathers and adopted and biological offspring. Correlation matrices for both parents (Table S9), and offspring (Table S10) are shown below.

8. Observed family correlations

As a supplement to Table 2 in the main text, we computed 95% confidence intervals for each familial relationship and for each phenotype. These are shown in Table S11.

9. Full parameter estimates

The raw parameter estimates and associated 95% confidence interval for each phenotype are reported in Table S12.

10. Polygenic scores

We use polygenic scores for years of education (PGS_{EA}) derived from the third GWAS of years of education (EA3; downloadable EA3 summary statistics). These PGS are constructed with the LDpred software package, which uses the correlations between SNPs estimated in an external reference panel (MCTFR white parents) to transform the GWAS summary statistics' univariate regression coefficients to, essentially, partial regression coefficients (Vilhjálmsson et al., 2015). Following the EA3 authors (Lee et al., 2018), we set the LDpred shrinkage parameter equal to unity (prior: 1.0), which represents the highest possible value and the one leading to the least shrinkage of the PGS weights. We do not use the optimization recommended by the developers (using a grid of values to choose the one leading to the best prediction accuracy in the validation sample) in order to avoid "double-dipping". Data from the target MCTFR sample were removed from the SSGAC GWAS before weights were derived due to the MCTFR sample being part of EA3.

Table S14 displays valid number of entries and predictive validity (R^2 and associated p-values) for PGS_{EA} on each demographic criterion (Family SES, family income, educational attainment) and all IQ scales and subscales. These predictions for adoptive, biological, Asian and white offspring are shown in separate columns of Table S14.

11. The meta-analysis of Devlin, Daniels, and Roeder (1997)

Some readers may wonder whether our estimates our consistent with those of Devlin et al. (1997), which include a rather prominent role for shared environment (and, in particular, shared prenatal environment). These authors meta-analyzed a dataset including correlations between twins, which are the chief drivers of their distinctive inferences. Lacking twins in SIBS, we cannot directly confirm or rebut their conclusion with respect to prenatal environment. Here we point to other works that tend to cast doubt on that conclusion.

After excluding correlations between adoptive siblings from their analysis because of their heterogeneity, Devlin et al. (1997) found increasing heritability and decreasing impact of shared environment with age to be less well supported than their favored model. There were five such studies of late adolescents and early adults available to these authors; four of the correlations clustered near zero, and the other was .19 (Bouchard Jr., 2009, 2013). Three of the five studies were longitudinal, and all showed the decline in resemblance expected from a decreasing effect of the shared environment. We note that in our own data the correlations between adoptive relatives in Total IQ or ICAR-16 all decreased from intake to follow-up, although not to the point of statistical significance. Since the time of Devlin et al. (1997), many well-powered twin studies including older adolescents have strongly supported age moderating genetic and environmental effects on IQ in exactly the manner tending to rule out prenatal environment as a substantial variance component (Bergen et al., 2007; Haworth et al., 2010; Briley & Tucker-Drob, 2013). Moreover, studies of whether sharing a chorion prenatally tends to make monozygotic twins more similar have found that any such effect is likely to be small (Marceau et al., 2016; van Beijsterveldt et al., 2016). If the resemblance between reared-together relatives becomes increasingly due to shared genes rather than a shared environment, then there is no need to posit prenatal environment as the explanation of the "surplus" resemblance between monozygotic twins. In any event it may be just as reasonable to attribute a surplus resemblance, if any remains to be explained, as due to higher-order epistasis with small coefficient in the correlations between non-twin relatives (Lynch & Walsh, 1998).

Table S1: Comparison of participants and non-participants in current wave relative to intake assessments

Intake measure	Participants			Non-Participants			Cohen's d
	N	M	SD	N	M	SD	
Offspring age at intake	753	15.03	1.97	481	14.76	1.83	.14*
Dad's education	282	16.53	2.39	126	16.24	2.61	.12
Mom's education	404	15.74	2.34	178	15.92	2.25	−.08
Dad's IQ (totall)	56	118.57	14.76	36	115.78	14.62	.19
Mom's IQ (total)	592	111.91	13.78	326	110.52	13.96	.10
Family SES[a]	743	0	1.01	472	0.01	0.98	−.01
Offspring IQ (total)	751	108.23	13.2	477	105.7	14.59	.18**
Offspring PGS_{EA}	659	0.356	1.04	330	0.228	1.13	.12

Note: PGS_{EA} = Years of education polygenic score. * denotes $p < .05$, ** $p < .01$, and *** $p < .001$ of mean difference from t-test.

[a] Family SES has been standardized with mean at 0.

Table S2: Scale means, standard deviations and valid scores for mothers, fathers and offspring in both types of family

Adoptive families

	Offspring			Mothers			Fathers		
	N	M	SD	N	M	SD	N	M	SD
SES	691	0.18	0.92	524	0.18	0.92	48	0.18	0.92
Highest degree	685	4.6	1.44	685	4.72	1.34	685	5.01	1.46
Years of education	557	15.84	2.56	564	16.07	2.3	515	16.69	2.48
PGS_{EA}	557	0.44	1.13	550	0.25	0.94	468	0.17	1.06
Total IQ	690	106.6	14.15	524	113.36	13.66	48	117.21	14.63
Verbal IQ	691	101.94	13.5	522	111.91	13.76	48	113.52	15.75
Performance IQ	691	110.9	16.51	520	112.58	14.94	48	117.67	14.86
Information	690	10.63	2.42	524	11.51	2.25	48	12.56	2.55
Block design	375	12.09	3.17	348	12.16	2.68	257	13.38	2.51
Picture completion	690	11.07	2.62	520	11.43	2.16	48	11.85	2.2
Vocabulary (intake)	691	10.38	2.49	520	12.11	2.29	48	11.58	2.42
Vocabulary (FU3)	413	10.79	2.18	NA	NA	NA	NA	NA	NA
ICAR-16	365	8.87	3.92	345	8.71	3.2	254	8.87	3.69

Biological families

	Offspring			Mothers			Fathers		
	N	M	SD	N	M	SD	N	M	SD
SES	539	–0.24	1.06	400	–0.24	1.06	44	–0.24	1.06
Highest degree	530	4.72	1.24	530	4.25	1.42	530	4.45	1.63
Years of education	432	16.07	2.15	439	15.39	2.27	388	16.04	2.38
PGS_{EA}	431	0.15	0.99	441	0.13	1.02	317	0.18	0.93
Total IQ	538	108.07	13.3	398	108.88	13.71	44	117.77	14.92
Verbal IQ	538	103.75	13.61	400	105.84	14.25	44	115.98	16.2
Performance IQ	538	111.86	14.79	400	111.84	14.17	44	116.05	13.82
Information	538	11.06	2.46	398	10.63	2.39	44	12.66	2.4
Block design	315	12.38	2.98	236	11.84	2.37	155	12.73	2.76
Picture completion	538	11.04	2.34	398	11.53	2.22	44	11.93	1.78
Vocabulary (IN)	538	10.62	2.38	400	11.1	2.49	44	11.86	2.63
Vocabulary (FU3)	340	10.88	2.13	NA	NA	NA	NA	NA	NA
ICAR-16	306	9.93	3.67	233	8.47	3.32	150	9.08	3.54

Note: PGS_{EA} = Years of education polygenic score; IN = intake; FU3 = follow-up 3.

Table S3: Reliability comparisons of ICAR-16 items in current study and Condon & Revelle (2014)

	α		ω_h		ω_t		Items	
	C&R	*W et al.*	*C&R*	*W et al.*	*C&R*	*W et al.*	*C&R*	*W et al.*
ICAR-16	.81	.80	.66	.64	.83	.82	16	16
LN items	.77	.66	.66	.62	.80	.67	9	4
MR items	.68	.50	.58	.49	.71	.54	11	4
R3D items	.93	.73	.78	.72	.94	.77	24	4
VR items	.76	.51	.64	.52	.77	.58	16	4

Note: C&R = Condon & Revelle (2014), *W et al.* = Willoughby et al. (current study), ω_h = omega hierarchical, ω_t = omega total. Values are based on composites of Pearson correlations between items. Total N sampled in Condon & Revelle (2014) was 96,958 individuals while a total of 1,172 had valid ICAR-16 data in our sample.

Table S4: Standardized factor loadings of each ICAR-16 item on subtest and general latent factors in the current sample. All loadings are significant at $p < .001$.

Indicator	β_S	β_g
Letter & Number		
LN.7	.52	.45
LN.33	.60	.52
LN.34	.64	.54
LN.58	.51	.45
Matrix Reasoning		
MR.45	.41	.34
MR.46	.46	.37
MR.47	.49	.44
MR.55	.40	.35
3D Rotation		
R3D.3	.65	.55
R3D.4	.73	.59
R3D.6	.55	.48
R3D.8	.57	.48
Verbal reasoning		
VR.4	.54	.44
VR.16	.22	.20
VR.17	.56	.45
VR.19	.52	.43

Note: β_S refers to the standardized loading of each item on the latent factor representing one of the four subtests to which it belongs (R3D = Three-dimensional Rotation, LN = Letter And Number series, VR = Verbal Reasoning, MR = Matrix Reasoning); β_g refers to each item's standardized loading on the single general latent factor.

Table S5: Latent factor correlations for the ICAR-16 in the current sample.

Factor 1	Factor 2	Correlation	p-value
LN	MR	.69	$< .001$
LN	R3D	.55	$< .001$
LN	VR	.78	$< .001$
MR	R3D	.66	$< .001$
MR	VR	.67	$< .001$
R3D	VR	.54	$< .001$

Note: R3D = Three-dimensional Rotation, LN = Letter And Number series, VR = Verbal Reasoning, MR = Matrix Reasoning.

Table S6: Comparison of means between white and Asian offspring for each demographic and scale criterion.

	t–statistic	White M	Asian M	Cohen's d [95% CI]	p–value
Demographics					
Highest degree	−1.13	4.61	4.74	.09 [−.07, .25]	.26
Years of education	−0.58	15.91	16.03	.05 [−.11, .21]	.56
Age	−4.36	31.47	32.41	.36 [.20, .52]	< .01
Family SES	−7.68	−0.18	0.26	.46 [.34, .58]	< .01
IQ measures					
ICAR-16	1.54	9.63	9.16	−.13 [−.29, .03]	.12
Total IQ	−0.47	107.61	108	.03 [−.09, .15]	.64
Verbal IQ	0.31	103.21	102.97	−.02 [−.14, .10]	.76
Performance IQ	−0.75	111.6	112.33	.05 [−.07, .17]	.45
Vocabulary (IN)	0.21	10.58	10.55	−.01 [−.13, .11]	.84
Vocabulary (FU3)	−1.78	10.77	11.07	.14 [−.02, .29]	.08
Picture completion	−0.39	11.07	11.13	.02 [−.10, .14]	.70
Information	0.95	10.92	10.78	−.06 [−.18, .06]	.34
Block design	−0.86	12.28	12.43	.05 [−.07, .17]	.39
PGS_{EA}	−13.12	0.046	0.868	.87 [.73, 1.01]	< .01

Note: PGS_{EA} = Years of education polygenic score; IN = intake; FU3 = follow-up 3.

Table S7: Comparison of means between adopted and biological offspring for each demographic and scale criterion.

	t-statistic	Adopted M	Biological M	Cohen's d [95% CI]	p-value
Demographics					
Highest degree	−1.21	4.6	4.72	.09 [−.06, .24]	.23
Years of education	−1.30	15.84	16.07	.10 [−.05, .25]	.20
Age	4.29	32.24	31.37	−.33 [−.48 , −.17]	< .01
Family SES	7.28	0.18	−0.24	−.43 [−.54 , −.31]	< .01
IQ measures					
ICAR-16	−3.61	8.87	9.93	.28 [.13, .43]	< .01
Total IQ	−1.86	106.6	108.07	.11 [−.01, .22]	.06
Verbal IQ	−2.32	101.94	103.75	.13 [.02, .25]	.02
Performance IQ	−1.07	110.9	111.86	.06 [−.05, .17]	.29
Vocabulary (IN)	−1.70	10.38	10.62	.10 [−.02, .21]	.09
Vocabulary (FU3)	−0.56	10.79	10.88	.04 [−.10, .18]	.58
Picture completion	0.19	11.07	11.04	−.01 [−.12 , .10]	.85
Information	−3.02	10.63	11.06	.17 [.06, .29]	< .01
Block design	−1.68	12.09	12.38	.10 [−.02, .21]	.09
PGS_{EA}	4.21	0.437	0.154	−.27 [−.39, −.14]	< .01

Note: PGS_{EA} = Years of education polygenic score; IN = intake; FU3 = follow-up 3.

Table S8: Correlation matrix for all valid individual scores of IQ and IQ subscales, years of education, highest degree obtained, EduYears polygenic score, and ICAR-16 score. Only measures shared across parents and offspring are included; i.e., all lQ scores and subscales are those taken at intake, while educational attainment, SES and ICAR-16 measures were evaluated at follow-up 3.

					Full sample						
	1	2	3	4	5	6	7	8	9	10	11
1. Highest degree											
2. Years of education	.81***										
3. SES	.46***	.41***									
4. PGS_{EA}	.28***	.25***	.21***								
5. ICAR-16	.27***	.30***	.19***	.21***							
6. Verbal IQ	.37***	.38***	.25***	.27***	.33***						
7. Performance IQ	.12***	.18***	.12***	.18***	.39***	.42***					
8. Total IQ	.32***	.35***	.23***	.27***	.43***	.86***	.81***				
9. Information	.36***	.37***	.23***	.26***	.36***	.88***	.39***	.78***			
10. Vocabulary	.32***	.33***	.23***	.24***	.29***	.90***	.38***	.77***	.64***		
11. Picture completion	.07*	.11***	.10***	.09***	.20***	.34***	.76***	.63***	.31***	.31***	
12. Block design	.12***	.16***	.10***	.18***	.40***	.33***	.82***	.66***	.33***	.30***	.28***

Note: PGS_{EA} = Years of education polygenic score. * denotes $p < .05$, ** $p < .01$, and *** $p < .001$.

Table S9: Correlation matrix for parents' valid individual scores of IQ and IQ subscales, years of education, highest degree obtained, EduYears polygenic score, and ICAR-16 score. Only measures shared across parents and offspring are included; i.e., all lQ scores and subscales are those taken at intake, while educational attainment, SES and ICAR-16 measures were evaluated at follow-up 3.

Mothers

	1	2	3	4	5	6	7	8	9	10	11
1. Highest degree											
2. Years of education	.91***										
3. SES	.55***	.54***									
4. PGS_{EA}	.31***	.28***	.27***								
5. ICAR-16	.23***	.28***	.23***	.17***							
6. Verbal IQ	.52***	.53***	.38***	.32***	.31***						
7. Performance IQ	.10*	.17***	.10**	.19***	.44***	.41***					
8. Total IQ	.41***	.45***	.30***	.30***	.43***	.87***	.80***				
9. Information	.48***	.50***	.37***	.30***	.33***	.90***	.39***	.80***			
10. Vocabulary	.44***	.45***	.35***	.28***	.30***	.91***	.39***	.81***	.70***		
11. Picture completion	.11*	.13**	.10**	.15***	.27***	.36***	.81***	.66***	.34***	.34***	
12. Block design	.05	.13**	.08*	.15***	.44***	.31***	.80***	.63***	.31***	.30***	.34***

Fathers

	1	2	3	4	5	6	7	8	9	10	11
1. Highest degree											
2. Years of education	.83***										
3. SES	.70***	.62***									
4. PGS_{EA}	.29***	.24***	.22***								
5. ICAR-16	.31***	.28***	.33***	.23***							
6. Verbal IQ	.30*	.25	.60***	.18	.67***						
7. Performance IQ	−.09	−.18	.37***	.10	.49***	.37***					
8. Total IQ	.17	.08	.61***	.15	.65***	.88***	.76***				
9. Information	.25	.20	.56***	.16	.58***	.93***	.39***	.85***			
10. Vocabulary	.39**	.34*	.63***	.18	.65***	.93***	.43***	.87***	.79***		
11. Picture completion	−.02	−.07	.40***	.02	.50***	.38***	.77***	.64***	.40***	.42***	
12. Block design	−.15	−.28	.19	.14	.28	.25*	.83***	.60***	.27*	.29**	.33**

Note: PGS_{EA} = Years of education polygenic score. * denotes $p < .05$, ** $p < .01$, and *** $p < .001$.

Table S10: Correlation matrix for adopted and biological offspring's valid individual scores of IQ and IQ subscales, years of education, highest degree obtained, EduYears polygenic score, and ICAR-16 score. All lQ scores and subscales are those taken at intake, except for vocabulary which was taken at both intake and follow-up 3, while educational attainment, SES and ICAR-16 measures were evaluated at follow-up 3.

Adopted offspring

	1	2	3	4	5	6	7	8	9	10	11	12
1. Highest degree												
2. Years of education	.71***											
3. SES	.23***	.19***										
4. PGS_{EA}	.27***	.30***	.03									
5. ICAR-16	.31***	.36***	.08	.23***								
6. Verbal IQ	.35***	.37***	.06	.28***	.41***							
7. Performance IQ	.18***	.22***	.11**	.21***	.35***	.45***						
8. Total IQ	.32***	.35***	.10**	.29***	.46***	.85***	.85***					
9. Information	.34***	.34***	.07	.24***	.39***	.88***	.40***	.75***				
10. Vocabulary (IN)	.33***	.34***	.01	.26***	.33***	.87***	.39***	.73***	.58***			
11. Picture completion	.05	.12*	.06	.11*	.15**	.31***	.74***	.62***	.29***	.25***		
12. Block design	.22***	.23***	.10**	.22***	.40***	.41***	.84***	.73***	.36***	.37***	.27***	
13. Vocabulary (FU3)	.42***	.44***	.10	.24***	.43***	.54***	.23***	.47***	.48***	.48***	.15**	.22***

Biological offspring

	1	2	3	4	5	6	7	8	9	10	11	12
1. Highest degree												
2. Years of education	.75***											
3. SES	.31***	.21***										
4. PGS_{EA}	.22***	.16**	.27***									
5. ICAR-16	.23***	.30***	.17**	.23***								
6. Verbal IQ	.24***	.27***	.25***	.32***	.36***							
7. Performance IQ	.10	.17**	.15***	.14**	.38***	.39***						
8. Total IQ	.22***	.27***	.24***	.28***	.45***	.85***	.81***					
9. Information	.24***	.27***	.20***	.29***	.38***	.89***	.36***	.77***				
10. Vocabulary (IN)	.18**	.20***	.25***	.27***	.28***	.87***	.33***	.74***	.61***			
11. Picture completion	.06	.13*	.09	.02	.17**	.30***	.70***	.58***	.26***	.29***		
12. Block design	.08	.14*	.15***	.17***	.38***	.30***	.83***	.66***	.29***	.24***	.20***	
13. Vocabulary (FU3)	.28***	.29***	.23***	.25***	.37***	.60***	.25***	.53***	.50***	.58***	.19***	.19***

Note: PGS_{EA} = Years of education polygenic score; IN = intake; FU3 = follow-up 3. * denotes $p < .05$, ** $p < .01$, and *** $p < .001$.

Table S11: Observed correlations and 95% confidence intervals for each phenotype reported in the main text.

	Parent correlations					Sibling correlations		
	Mom/ bio	*Dad/ bio*	*Mom/ adopt*	*Dad/ adopt*	*Dad/ Mom*	*Bio/ bio*	*Adopt/ bio*	*Adopt/ adopt*
Performance IQ								
Observed	.28	.41	.12	−.08	.08	.26	.27	.08
95% CI	[.18, .37]	[.12, .63]	[.04, .21]	[−.35, .21]	[.01, .15]	[.13, .38]	[.10, .43]	[−.04, .19]
Verbal IQ								
Observed	.35	.41	.07	.19	.40	.43	.24	.07
95% CI	[.26, .43]	[.13, .63]	[−.01, .16]	[−.10, .45]	[.34, .45]	[.31, .53]	[.07, .40]	[−.04, .19]
Total IQ								
Observed	.36	.46	.09	.16	.30	.34	.30	.09
95% CI	[.28, .45]	[.19, .67]	[.00, .17]	[−.13, .43]	[.23, .36]	[.21, .46]	[.13, .45]	[−.03, .21]
Block design								
Observed	.32	.41	.14	−.10	.03	.28	.32	.00
95% CI	[.23, .40]	[.13, .63]	[.06, .23]	[−.37, .19]	[−.04, .10]	[.15, .40]	[.15, .47]	[−.12, .11]
Picture completion								
Observed	.09	.30	.04	.01	.13	.07	.19	.02
95% CI	[−.01, .19]	[.00, .55]	[−.05, .12]	[−.28, .29]	[.06, .20]	[−.06, .21]	[.01, .35]	[−.10, .14]
Vocabulary (IN)								
Observed	.39	.44	.07	.26	.37	.39	.24	.04
95% CI	[.31, .47]	[.16, .65]	[−.01, .16]	[−.03, .50]	[.31, .43]	[.26, .50]	[.07, .40]	[−.08, .15]
Information								
Observed	.23	.28	.04	.15	.34	.37	.15	.13
95% CI	[.13, .32]	[−.02, .53]	[−.04, .13]	[−.14, .42]	[.27, .40]	[.25, .48]	[−.03, .32]	[.01, .24]
ICAR-16								
Observed	.27	.31	−.03	.10	.19	.27	.05	.07
95% CI	[.12, .41]	[.11, .48]	[−.17, .10]	[−.06, .25]	[.08, .29]	[.06, .46]	[−.28, .38]	[−.13, .27]
Vocabulary (FU3)								
Observed	.33	.45	.18	.32	.37	.24	.16	.25
95% CI	[.22, .43]	[.09, .70]	[.07, .28]	[−.07, .62]	[.31, .43]	[.05, .41]	[−.14, .43]	[.08, .41]

Note: Observed and model-predicted correlations are reported in Table 2 in the main text. IN = intake; FU3 = follow-up 3.

17

Table S12: Full parameter estimates for each measure of IQ and the ICAR-16. 95% confidence intervals are calculated over 200 bootstrap resamples. Dominance variance (dsq) is constrained to zero and is not reported below.

	a [95% CI]	q [95% CI]	s^2 [95% CI]	m [95% CI]	p [95% CI]	w [95% CI]	x [95% CI]	μ [95% CI]
Total IQ	0.53 [0.42, 0.60]	1.15 [1.10, 1.20]	0.06 [0.00, 0.13]	0.06 [0.00, 0.13]	0.17 [0.00, 0.31]	0.10 [0.04, 0.15]	0.04 [0.10, 0.10]	0.30 [0.30, 0.31]
Verbal IQ	0.54 [0.46, 0.61]	1.23 [1.16, 1.28]	0.07 [0.00, 0.15]	0.00 [−0.07, 0.07]	0.18 [0.03, 0.30]	0.09 [0.04, 0.14]	0.03 [0.00, 0.09]	0.40 [0.39, 0.40]
Performance IQ	0.50 [0.37, 0.63]	1.03 [1.02, 1.04]	0.08 [0.01, 0.14]	0.11 [0.05, 0.17]	0.08 [−0.06, 0.23]	0.06 [0.01, 0.11]	0.02 [0.01, 0.07]	0.08 [0.07, 0.08]
Information	0.50 [0.40, 0.58]	1.12 [1.08, 1.17]	0.12 [0.05, 0.18]	−0.01 [−0.07, 0.06]	0.13 [0.02, 0.24]	0.05 [0.01, 0.08]	0.02 [0.00, 0.05]	0.34 [0.33, 0.34]
Block Design	0.56 [0.43, 0.69]	1.01 [1.01, 1.02]	0.03 [0.00, 0.10]	0.13 [0.09, 0.18]	0.09 [−0.05, 0.20]	0.08 [0.03, 0.11]	0.03 [0.01, 0.07]	0.03 [0.02, 0.03]
Picture Completion	0.28 [0.00, 0.46]	1.01 [0.99, 1.03]	0.03 [0.00, 0.10]	0.02 [−0.03, 0.07]	0.13 [−0.05, 0.23]	0.03 [−0.01, 0.05]	0.02 [0.00, 0.05]	0.13 [0.13, 0.13]
Vocabulary (IN)	0.55 [0.48, 0.62]	1.23 [1.17, 1.29]	0.02 [0.00, 0.09]	0.01 [−0.05, 0.09]	0.20 [0.03, 0.31]	0.11 [0.05, 0.15]	0.04 [0.01, 0.09]	0.37 [0.37, 0.38]
Vocabulary (FU3)	0.33 [0.00, 0.46]	1.08 [1.00, 1.16]	0.06 [0.00, 0.15]	0.09 [0.01, 0.19]	0.27 [0.00, 0.42]	0.12 [0.04, 0.22]	0.10 [0.03, 0.20]	0.37 [0.37, 0.38]
ICAR-16	0.62 [0.42, 0.74]	1.09 [1.03, 1.15]	0.04 [0.00, 0.15]	−0.03 [−0.11, 0.06]	0.08 [−0.01, 0.17]	0.02 [−0.04, 0.06]	0.01 [0.00, 0.03]	0.19 [0.08, 0.29]

Note: a = the effect of the genetic score on the phenotype; q = the variance of the genetic score, i.e., the additive genetic variance; s^2 = variance of environmental factors shared by siblings reared together; m = direct effect of maternal phenotype on offspring phenotype; p = direct effect of paternal phenotype on offspring phenotype; w = covariance between additive genetic factors and family environment; x = variance of the shared environment induced by parental phenotypes; μ = correlation between spouses; IN = intake; FU3 = follow-up 3.

Table S13: Predictive validity and N valid scores for PGS_{EA} (Lee et al., 2018) on various IQ and demographic variables in full combined sample of white parents and offspring.

Phenotype	N	R^2	p-value	95% CI
Total IQ	983	.114	4.28×10^{-23}	.077, .151
Verbal IQ	985	.154	3.76×10^{-33}	.113, .195
Performance IQ	984	.031	1.07×10^{-8}	.003, .034
Information	987	.102	3.23×10^{-31}	.066, .138
Block Design	984	.031	2.55×10^{-8}	.010, .052
Picture Completion	985	.016	6.87×10^{-3}	~ 0, .031
Vocabulary	985	.124	2.33×10^{-16}	.086, .162
ICAR-16	751	.061	2.70×10^{-11}	.028, .094
Highest degree	770	.081	4.57×10^{-15}	.044, .118
Years of education	768	.061	1.19×10^{-10}	.028, .094
WRAT	1401	.082	9.65×10^{-28}	.055, .109
Family SES	1400	.060	4.44×10^{-11}	.036, .084

Note: Reported values of R^2 were computed with generation (1 = parents, 2 = offspring) as indicator variable. To account for non-independence of offspring from the same family, bootstrap resampling (100 iterations) over families was used for statistical inference. 95% CIs are estimated with the *CI.Rsq* function of R's "psychometric" package. *WRAT* = Wide Range Achievement Test.

Table S14: Predictive validity and N valid scores for PGS_{EA} (Lee et al., 2018) on various IQ and demographic variables in adopted, biological, white and Asian offspring

	Adoptive			Biological			Asian			White		
	N	R^2	p-value	N	R^2	p-value	N	R^2	p-value	N	R^2	p-value
Intake measures												
Family SES	553	.001	.480	422	.075	$< .001$	365	.006	.125	524	.039	$< .001$
Family income	464	.001	.451	302	.034	.001	310	.002	.436	390	.011	.035
Total IQ	556	.082	$< .001$	430	.080	$< .001$	365	.043	$< .001$	536	.100	$< .001$
Verbal IQ	556	.079	$< .001$	430	.101	$< .001$	365	.049	$< .001$	536	.126	$< .001$
Performance IQ	557	.046	$< .001$	430	.019	.004	365	.020	.007	536	.028	$< .001$
Information	557	.057	$< .001$	431	.084	$< .001$	365	.038	$< .001$	537	.111	$< .001$
Block Design	557	.050	$< .001$	430	.031	$< .001$	365	.022	.004	536	.037	$< .001$
Picture Completion	557	.011	.013	430	~ 0	.715	365	.005	.169	536	.002	.297
Vocabulary	556	.067	$< .001$	430	.075	$< .001$	365	.042	$< .001$	536	.091	$< .001$
Follow-up 3 measures												
Highest degree	329	.074	$< .001$	272	.049	$< .001$	217	.077	$< .001$	337	.071	$< .001$
Years of education	330	.092	$< .001$	273	.027	.007	218	.106	$< .001$	338	.049	$< .001$
Vocabulary	362	.059	$< .001$	296	.060	$< .001$	237	.013	.085	368	.099	$< .001$
ICAR-16	319	.053	$< .001$	265	.053	$< .001$	214	.040	.003	325	.060	$< .001$

Note: Analyses are conducted using full sample of offspring with valid scores. To account for non-independence of offspring from the same family, bootstrap resampling (200 iterations) over families was used for statistical inference.

12. References

van Beijsterveldt, C. E. M., Overbeek, L. I. H., Rozendaal, L., McMaster, M. T. B., Glasner, T. J., Bartels, M., Vink, J. M., Martin, N. G., Dolan, C. V., & Boomsma, D. I. (2016). Chorionicity and heritability estimates from twin studies: The prenatal environment of twins and their resemblance across a large number of traits. *Behavior Genetics*, *46*, 304–314. doi:10.1007/s10519-015-9745-3.

Bergen, S. E., Gardner, C. O., & Kendler, K. S. (2007). Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: A meta-analysis. *Twin Research and Human Genetics*, *10*, 423–433. doi:10.1375/twin.10.3.423.

Bouchard Jr., T. J. (2009). Genetic influence on human intelligence (spearman's *g*): How much? *Annals of Human Biology*, *36*, 527–544. doi:10.1080/03014460903103939.

Bouchard Jr., T. J. (2013). The Wilson Effect: The increase in heritability of IQ with age. *Twin Research and Human Genetics*, *16*, 923–930. doi:10.1017/thg.2013.54.

Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science*, *24*, 1704–1713. doi:10.1177/0956797613478618.

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52–64. doi:10.1016/j.intell.2014.01.004.

Devlin, B., Daniels, M., & Roeder, K. (1997). The heritability of IQ. *Nature*, *388*, 468–471. doi:10.1038/41319.

Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J. C., van Beijsterveldt, C. E. M., Bartels, M., Posthuma, D., Boomsma, D. I., Davis, O. S. P., Kovas, Y., Corley, R., DeFries, J. C., Hewitt, J. K., Olson, R. K., Rhea, S.-A., Wadsworth, S. J., Iacono, W. G., McGue, M., Thompson, L. A., Hart, S. A., Petrill, S. A., Lubinski, D., & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry*, *15*, 1112–1120. doi:10.1038/mp.2009.55.

Hill, W. G., Goddard, M. E., & Visscher, P. M. (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLOS Genetics*, *4*, e1000008. doi:10.1371/journal.pgen.1000008.

Hivert, V., Sidorenko, J., Rohart, F., Goddard, M. E., Yang, J., Wray, N. R., Yengo, L., & Visscher, P. M. (2021). Estimation of non-additive genetic variance in human complex traits from a large sample of unrelated individuals. *American Journal of Human Genetics*, *108*, 786–798. URL: https://doi.org/10.1016/j.ajhg.2021.02.014. doi:10.1016/j.ajhg.2021.02.014.

Jensen, A. R. (1983). Effects of inbreeding on mental-ability factors. *Personality and Individual Differences*, *4*, 71–87. doi:10.1016/0191-8869(83)90054-5.

Joshi, P. K., Esko, T., Mattsson, H., Eklund, N., Gandin, I., Nutile, T., Jackson, A. U., Schurmann, C., Smith, A. V., Zhang, W., Okada, Y., Stančáková, A., Faul, J. D., Zhao, W., Bartz, T. M., Concas, M. P., Franceschini, N., Enroth, S., Vitart, V., Trompet, S., Guo, X., Chasman, D. I., O'Connel, J. R., Corre, T., Nongmaithem, S. S., Chen, Y., Mangino, M., Ruggiero, D., Traglia, M., Farmaki, A.-E., Kacprowski, T., Bjonnes, A., van der Spek, A., Wu, Y., Giri, A. K., Yanek, L. R., Wang, L., Hofer, E., Rietveld, C. A., McLeod, O., Cornelis, M. C., Pattaro, C., Verweij, N., Baumbach, C., Abdellaoui, A., Warren, H. R., Vuckovic, D., Mei, H., Bouchard, C., Perry, J. R. B., Cappellani, S., Mirza, S. S., Benton, M. C., Broeckel, U., Medland, S. E., Lind, P. A., Malerba, G., Drong, A., Yengo, L., Bielak, L. F., Zhi, D., van der Most, P. J., Shriner, D., Mägi, R., Hemani, G., Karaderi, T., Wang, Z., Liu, T., Demuth, I., Zhao, J. H., Meng, W., Lataniotis, L., van der Laan, S. W., Bradfield, J. P., Wood, A. R., Bonnefond, A., Ahluwalia, T. S., Hall, L. M., Salvi, E., Yazar, S., Carstensen, L., de Haan, H. G., Abney, M., Afzal, U., Allison, M. A., Amin, N., Asselbergs, F. W., Bakker, S. J. L., Barr, R. G., Baumeister, S. E., Benjamin, D. J., Bergmann, S., Boerwinkle, E., Bottinger, E. P., Campbell, A., Chakravarti, A., Chan, Y., Chanock, S. J., Chen, C. et al. (2015). Directional dominance on stature and cognition in diverse human populations. *Nature*, *523*, 459–462. doi:10.1038/nature14618.

Keller, M. C., Medland, S. E., Duncan, L. E., Hatemi, P. K., Neale, M. C., Maes, H. H., & Eaves, L. J. (2009). Modeling extended twin family data I: Description of the cascade model. *Twin Research and Human Genetics*, *12*, 8–18. doi:`10.1375/twin.12.1.8`.

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., Yengo, L., Alver, M., Bao, Y., Clark, D. W., Day, F. R., Furlotte, N. A., Joshi, P. K., Kemper, K. E., Kleinman, A., Langenberg, C., Mägi, R., Trampush, J. W., Verma, S. S., Wu, Y., Lam, M., Zhao, J. H., Zheng, Z., Boardman, J. D., Campbell, H., Freese, J., Harris, K. M., Hayward, C., Herd, P., Kumari, M., Lencz, T., Luan, J., Malhotra, A. K., Metspalu, A., Milani, L., Ong, K. K., Perry, J. R. B., Porteous, D. J., Ritchie, M. D., Smart, M. C., Smith, B. H., Tung, J. Y., Wareham, N. J., Wilson, J. F., Beauchamp, J. P., Conley, D. C., Esko, T., Lehrer, S. F., Magnusson, P. K. E., Oskarsson, S., Pers, T. H., Robinson, M. R., Thom, K., Watson, C., Chabris, C. F., Meyer, M. N., Laibson, D. I., Yang, J., Johannesson, M., Koellinger, P. D., Turley, P., Visscher, P. M., Benjamin, D. J., & Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*, 1112–1121. doi:`10.1038/s41588-018-0147-3`.

Lynch, M., & Walsh, B. (1998). *Genetics and the Analysis of Quantitative Traits*. Sunderland, MA: Sinauer.

Maki-Tanila, A., & Hill, W. G. (2014). Influence of gene interaction on complex trait variation with multilocus models. *Genetics*, *198*, 355–367. doi:`10.1534/genetics.114.165282`.

Marceau, K., McMaster, M. T. B., Smith, T. F., Daams, J. G., van Beijsterveldt, C. E. M., Boomsma, D. I., & Knopik, V. S. (2016). The prenatal environment in twin studies: A review on chorionicity. *Behavior Genetics*, *46*, 286–303. doi:`10.1007/s10519-016-9782-6`.

Pazokitoroudi, A., Chiu, A. M., Burch, K. S., Pasaniuc, B., & Sankararaman, S. (2021). Quantifying the contribution of dominance deviation effects to complex trait variation in biobank-scale data. *American Journal of Human Genetics*, *108*, 799–808. URL: `https://doi.org/10.1016/j.ajhg.2021.03.018`. doi:`10.1016/j.ajhg.2021.03.018`.

Schull, W. J., & Neel, J. V. (1965). *The Effects of Inbreeding on Japanese Children*. New York, NY: Harper and Row.

Vilhjálmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., Genovese, G., Loh, P. R., Bhatia, G., Do, R., Hayeck, T., Won, H. H., Neale, B. M., Corvin, A., Walters, J. T., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., Belliveau, R. A., Bene, J., Bergen, S. E., Bevilacqua, E., Bigdeli, T. B., Black, D. W., Bruggeman, R., Buccola, N. G., Buckner, R. L., Byerley, W., Cahn, W., Cai, G., Campion, D., Cantor, R. M., Carr, V. J., Carrera, N., Catts, S. V., Chambert, K. D., Chan, R. C., Chen, R. Y., Chen, E. Y., Cheng, W., Cheung, E. F., Chong, S. A., Cloninger, C. R., Cohen, D., Cohen, N., Cormican, P., Craddock, N., Crowley, J. J., Curtis, D., Davidson, M., Davis, K. L., Degenhardt, F., Del Favero, J., Delisi, L. E., Demontis, D., Dikeos, D., Dinan, T., Djurovic, S., Donohoe, G., Drapeau, E., Duan, J., Dudbridge, F., Durmishi, N., Eichhammer, P., Eriksson, J., Escott-Price, V., Essioux, L., Fanous, A. H., Farrell, M. S., Frank, J., Franke, L., Freedman, R., Freimer, N. B., Friedl, M., Friedman, J. I., Fromer, M., Georgieva, L., Gershon, E. S., Giegling, I., Giusti-Rodrguez, P., Godard, S., Goldstein, J. I., Golimbet, V., Gopal, S., Gratten, J., Grove, J., De Haan, L., Hammer, C. et al. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *American Journal of Human Genetics*, *97*, 576–592. doi:`10.1016/j.ajhg.2015.09.001`.