ARTICLE   OPEN

Check for updates

Schooling substantially improves intelligence, but neither lessens nor widens the impacts of socioeconomics and genetics

Nicholas Judd [1,2 ✉], Bruno Sauce [3] and Torkel Klingberg [1]

Schooling, socioeconomic status (SES), and genetics all impact intelligence. However, it is unclear to what extent their contributions are unique and if they interact. Here we used a multi-trait polygenic score for cognition (cogPGS) with a quasi-experimental regression discontinuity design to isolate how months of schooling relate to intelligence in 6567 children (aged 9–11). We found large, independent effects of schooling ($\beta \sim 0.15$), cogPGS ($\beta \sim 0.10$), and SES ($\beta \sim 0.20$) on working memory, crystallized (cIQ), and fluid intelligence (fIQ). Notably, two years of schooling had a larger effect on intelligence than the lifetime consequences, since birth, of SES or cogPGS-based inequalities. However, schooling showed no interaction with cogPGS or SES for the three intelligence domains tested. While schooling had strong main effects on intelligence, it did not lessen, nor widen the impact of these preexisting SES or genetic factors.

INTRODUCTION

Inter-individual differences in intelligence predict a variety of important life outcomes, such as life satisfaction, mortality, and educational achievement (EA)[1–3], showing correlations higher than 0.7 with achievement[4,5]. There has been considerable controversy on the source of these differences. Decades of research have shown intelligence to be highly associated with genetic differences (heritability estimates around 0.6) but also impacted by experiences and context (called here "environmental factors")[6–12].

Schooling is an important environmental factor, having a large impact on intelligence in children[13–15]. This has been shown in longitudinal studies controlling for prior intelligence[16] and in studies evaluating the cognitive effects of policy changes regarding compulsory schooling[17]. A third method to evaluate the impact of schooling (called regression discontinuity designs) exploits the fact children are put in a grade based on an arbitrary age cut-off, and allows us to separate the effect of chronological age from months of schooling[18]. Studies using this method also replicate the findings that schooling affects intelligence[19,20]. Thus, many lines of research provide converging evidence that schooling can change abilities often thought to be "fixed", such as fluid intelligence and working memory, with estimates of one year of additional schooling benefitting cognitive abilities somewhere between 1 to 5 IQ points, or 0.07 to 0.3 SD[13,19,21]. What is less clear is how the impact of schooling *interacts* with the environments that children experience and with their genetic predispositions—in other words, do school settings amplify preexisting differences (i.e., "rich-gets-richer"), or conversely, weaken existing differences between children (i.e., "catch up").

Much of a child's environment is captured by socioeconomic status (SES), a summary measure usually comprised of household income, parental education, and neighborhood quality[22]. While commonly implicated in initial differences in intelligence, SES has also been found to widen existing differences throughout development[23–25]. Recently, research on SES has been criticized

for neglecting the role of genetics—as parents not only hand down environments but also genes[26,27]. For example, a genetically informed twin study found a strong genetic influence on SES and its association with intelligence[28]. While there is a large body of research looking at how preexisting SES differences interplay with years of schooling[29], none of these previous studies have incorporated genetics.

Fortunately, genome-wide association studies (GWAS) with extremely large sample sizes and the viability of polygenic scores, have made it possible to incorporate genetically informative measures into a study. A polygenic score is an index that combines thousands/millions of DNA regions (each with only a tiny effect on the trait of interest) and gives a value to each individual representing their genetic propensity. Relevant to us here, a multi-trait cognitive polygenic score (cogPGS) was recently shown to predict 7–10% of the variance in cognitive performance[30]. This and similar polygenic scores correlate moderately (r ~0.3) with SES[31,32], but little is known about their unique contributions to different domains of intelligence. And even less known if, or how, they interplay with schooling.

Gene-by-environment (GE) interplay is often proposed as an explanation of how intelligence can show high heritability alongside malleability[33–36]. To date, there are only a few studies on GE interplay using polygenic scores. Of these studies, most have focused on SES as the environmental variable of interest and none have used the environmental variable of schooling, which is particularly relevant to intelligence. Unfortunately, GE-interplay results with SES have been inconsistent for educational achievement/attainment—some positive, negative, or null[37–41]. While there is a high correlation between educational achievement and intelligence, educational achievement is a broader concept thought to also include personality characteristics such as consciousness and openness to experience[42,43]. Furthermore, we are not aware of any research examining if schooling moderates' genetic effects on intelligence.

[1]Department of Neuroscience, Karolinska Institute, Stockholm, Sweden. [2]Cognitive Neuroscience Department, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands. [3]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. ✉email: nickkjudd@gmail.com

Schooling, SES, and genetics thus represent three substantial contributors to intelligence, but it is unclear to what extent their contributions are (1) unique and (2) interact with one another. In a sample of 3rd to 5th grade children, we estimated the unique contributions of a year of schooling, SES, and cogPGS on crystallized intelligence (cIQ), fluid intelligence (fIQ), and working memory (WM). We chose these domains of intelligence since they are particularly important for educational outcomes[4,16,44]. WM was included since WM has shown to be heavily impacted by schooling[21], malleable[45], and potentially relevant for GE-interplay[46]. Our main aim was to examine if GE-interplay is present for these domains of cognition. Specifically, we tested if the effect of (1) schooling or (2) SES is moderated by cogPGS, and (3) a three-way interaction between schooling, SES, and cogPGS. These interactions allowed us to test if schooling enhances or compensates for preexisting genetic and environmental inequity.

RESULTS

We included data from 6567 children (mean age = 9.88, range 8.92–11.00; Supplementary Table 1) recruited by the ABCD consortium to be representative of the United States in sex, race, ethnicity, SES, and urbanicity[47]. SES was defined as the first component of a probabilistic PCA, capturing 65% of the variance in total household income, highest parental education, and neighborhood quality. Due to modeling constraints, we had to exclude 1,086 children that were missing DNA data. This group had significantly lower SES (Cohen's $d = -0.23$, $p < 0.001$). SES and cogPGS were positively correlated with each other and each cognitive domain (Fig. 1 and Supplementary Fig. 4).

Schooling, SES, and cogPGS effect on cognition

To examine the effect of schooling on cIQ, fIQ, and WM development, we used a regression discontinuity design over grades 3 to 5, while controlling for age, sex, cogPGS, SES, and 20 ancestry-based principal components (Eq. 2, Supplementary Tables 3–5).

We found that schooling had a significant effect (all p's < 0.001) for each cognitive domain we examined (cIQ: $\beta = 0.13$, fIQ: $\beta = 0.10$, WM: $\beta = 0.09$). One year of schooling (i.e., 10 months) contributed 0.22 SD, 0.14 SD, and 0.14 SD to the development of cIQ, fIQ, and WM, respectively (Fig. 2). Notably, the ratio of 1 year of schooling to 1 year of chronological age differed depending on
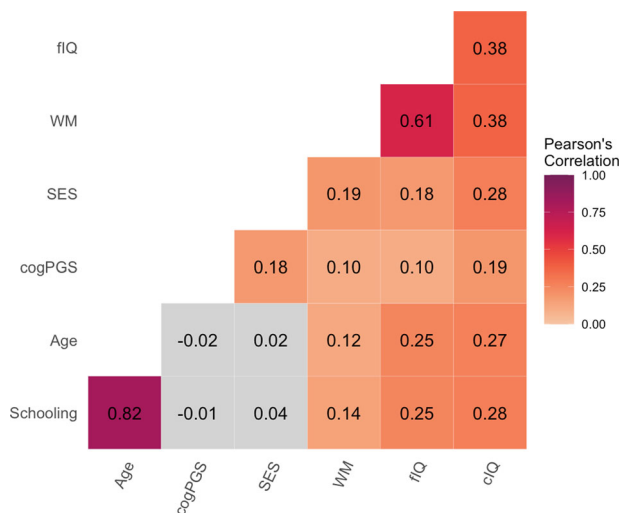
the domain. As expected, schooling affected cIQ (ratio = 1.1) more than fIQ (ratio = 0.54). Interestingly, WM, a subtask of fIQ, had the highest ratio of 2.2, indicating for this age range, 1 year of schooling had double the effect of 1 year of chronological age. There were no significant differences between the schooling coefficients for cIQ, fIQ, or WM ($|Z| < 1.96$).

Sex had a significant effect for each cognitive domain, with females performing better in fIQ ($\beta = 0.09$, $p < 0.001$) yet worse in cIQ ($\beta = -0.05$, $p = 0.023$) and WM ($\beta = -0.08$, $p < 0.001$). CogPGS and SES both showed independent significant effects (all p's < 0.001) for cIQ, fIQ, and WM. The effect sizes of cogPGS ($\beta = 0.16$) and SES ($\beta = 0.29$) were the largest for cIQ. CogPGS showed similarly sized effects for fIQ ($\beta = 0.09$) and WM ($\beta = 0.09$), yet SES had a slightly higher effect for WM ($\beta = 0.22$) than fIQ ($\beta = 0.18$).
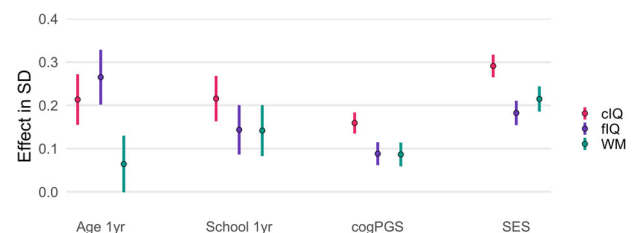
SES subcomponents

The analysis above showed that our combined measure of SES had effects on cIQ, fIQ, and WM. Since research has shown that different subcomponents of SES could relate differently to intelligence[48–50], we also performed an analysis of each sub-component. All components were significant ($p_{FDR} < 0.001$) for each cognitive domain. cIQ showed the largest effects for parental education ($\beta = 0.26$), family income ($\beta = 0.22$) and neighborhood quality ($\beta = 0.11$). In a similar fashion to the SES composite, WM had slightly larger effect sizes than fIQ for parental education (WM: $\beta = 0.19$; fIQ: $\beta = 0.17$), family income (WM: $\beta = 0.17$; fIQ: $\beta = 0.14$), and neighborhood quality (WM: $\beta = 0.08$; fIQ: $\beta = 0.06$).

Sibling analysis

Gene-environment correlations can inflate the estimation of cogPGS on phenotypes[33,51,52]. Therefore, we conducted a post hoc sibling analysis (families = 392, $n = 792$) to estimate within (β_w) and between (β_B) family effects of cogPGS on cIQ, fIQ, and WM. Within-family effects (β_w) are less confounded by the shared environment since the transmission of alleles is random, giving each sibling an equal probability of inheriting any given allele. Yet, there is no variance within families for SES and very little for schooling therefore, the sibling analysis is limited to genetics (cogPGS). We found cIQ's within-family effect ($\beta_w = 0.08$, $p_{FDR} = 0.009$) to be roughly half of the between-family effect ($\beta_B = 0.15$, $p_{FDR} = 0.009$). For fIQ we found a significant within-family effect ($\beta_w = 0.08$, $p_{FDR} = 0.013$), yet the between-family effect was not significant ($\beta_B = 0.10$, $p_{FDR} = 0.074$). For WM, neither the within-family ($\beta_w = 0.04$, $p_{FDR} = 0.235$) nor the between-family ($\beta_B = 0.10$, $p_{FDR} = 0.066$) effects were significant.

Gene-by-environment interplay

There were no significant interactions for any of the cognitive domains (Eqs. 3 and 4; Supplementary Tables 3–5). This included our two-way interaction terms of interest (i.e., schooling-cogPGS, schooling-SES, and cogPGS-SES) and the three-way interaction

Fig. 1 Corrected correlation plot. A correlation plot showing the relationships of our variables of interest, colored values indicate $p < 0.001$. These values are corrected for genetic PC's (see Supplementary Fig. 4 for more variables).

Fig. 2 Main effects for cognition. Independent effects of 1 year of age, 1 year of schooling, cogPGS, and SES from mixed-effects models predicting cIQ, fIQ, and WM (Eq. 2). All results are significant for predicting their respective cognitive domain (Supplementary Tables 3–5). Bars reflect 95% frequentist confidence intervals.
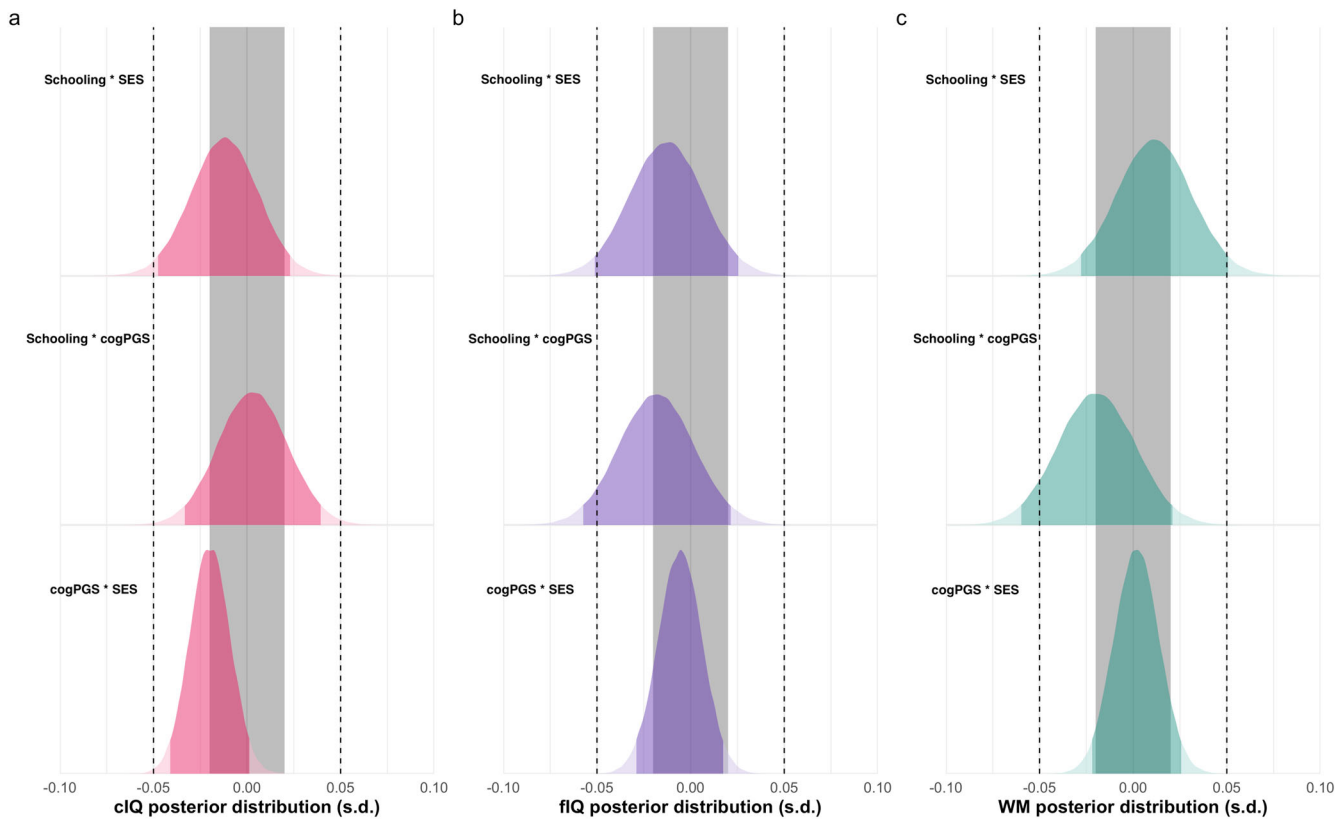
Fig. 3 Interaction effects for cognition. The estimated interaction effects of schooling-SES, schooling-cogPGS and cogPGS-SES on **a** cIQ, **b** fIQ, and **c** WM from a hierarchical Bayesian mixed-effects model. All values are in standard units. Dark blue shading indicates the 95% highest-density interval's (HDI) of the estimates posterior distribution. The gray band represents a region of practical equivalence (ROPE) of 0.02 SD's, while the dotted lines indicate a ROPE boundary of 0.05 SD's. Most of the interaction terms are fully within a 0.05 SD ROPE boundary providing evidence for the null hypothesis at this boundary, yet none fully lie within a ROPE boundary of 0.02 SD's (Supplementary Table 8).

term (i.e., schooling-cogPGS-SES). This null result held when we added principal component interactions with schooling, cogPGS, and SES for each model[53]. We, therefore, carried out post hoc Bayesian null hypothesis testing using region of practical equivalence (ROPE) boundaries with 95% highest-density intervals (HDI) for the two-way interaction terms of interest (Fig. 3, Eq. 3). We used weakly informative priors, centered around zero, for the two-way interaction terms (see Methods). Crucially, we used two ROPE boundaries in standard units of 0.05 and 0.02. This was done for two reasons, (1) a lack of consensus in the field on a minimal effect size of interest and (2) schooling representing a cumulative process[54–56].

Almost all interaction terms were entirely within a ROPE boundary of 0.05 (Fig. 3). This means that with a 95% probability, the interaction effect was less than 0.05 SD, adding evidence for the null hypothesis (Supplementary Table 8). The two exceptions were the schooling-cogPGS interaction (Fig. 3b, c), where 95% were slightly outside the 0.05 boundary for fIQ (3%) and WM (5%). However, most of the terms have a substantial part of their distribution overlapping with the ROPE boundaries of 0.02. This means we cannot confirm nor disprove the null hypothesis for a maximal effect size of interest of 0.02 SDs.

Post hoc *g* analysis
A single-factor confirmatory factor analysis fit the cognitive tasks well (RMSEA = 0.03 & CFI = 0.987; Supplementary Fig. 2). We then extracted factor scores (*g*) and used them as the dependent variable. The standardized effect of schooling on *g* was 0.136 SD ($p_{FDR} < 0.001$)—a similar magnitude as fIQ and WM's effects. Mirroring the other cognitive domains, no two- or three-way

interactions were significant following FDR correction (Supplementary Table 6).

Post hoc European ancestry analysis
CogPGS has less accuracy in non-European populations therefore, we conducted a post hoc reanalysis in the full model (Eq. 4) for each cognitive variable, including only those children whose genetic ancestry 4-means clustered on the first two principal components[57,58]. This subsample was less representative of the United States with substantially higher SES (Cohen's $d = 0.80$., $p < 0.001$). The results from that analysis stayed generally the same as the analysis of the full population (Supplementary Table 7), both in terms of effect sizes on the main results and the lack of significance on the interaction terms. In reference to the full sample, there was a slight increase in the size of the effect from CogPGS (β range = 0.005–0.020), yet this was followed by larger decreases from SES (β range = −0.046 to −0.080).

DISCUSSION
Schooling showed substantial and independent effects for each intelligence domain tested: cIQ, fIQ, and WM. In line with previous research, we found the raw effect of schooling on cIQ to be larger than for fIQ, though this difference was not significant[13]. This was also true for the relative influence of schooling compared to the effect of chronological age, as that ratio for cIQ was almost double that of fIQ (1.1 vs 0.54), showing almost near equal influences per year of schooling and age. Surprisingly, WM had the highest ratio (2.2), with the effect of schooling being more than double that of age. While this is in agreement with a previous study in younger

children[21], it should be interpreted with caution as the WM measure comprised of only one task[59].

As expected, SES and cogPGS were highly correlated with each other, highlighting the need to isolate the independent effects of each. Both had large, independent effects on cIQ, fIQ, and WM. In a follow-up analysis, we estimated the contribution of each SES component separately. Notably, all SES components were significant for each intelligence domain, with the effects from parental education and income being similar in size while neighborhood quality was roughly half.

A child's SES is not independent of their cogPGS[31], which makes it difficult to support causal inferences of these factors as well as interpret the interaction between them (i.e., the endogeneity problem)[60]. This gene-environment dependence can cause spurious gene-environment interactions[41,61]. Our sibling analysis sheds light on this issue. We found the within-family effect of cogPGS to be roughly half of the between-family effect for cIQ, in line with the literature[51,52]. This indicates the presence of passive genotype-environment correlations—whereby parents create family environments consistent with their genotypes, which in turn facilitate the development of their children's intelligence. Since we only had data from 392 families, a lack of statistical power is most likely the reason for our null findings for fIQ and WM. While our cogPGS estimate in the full sample should be interpreted with caution, previous research has shown SES to be the major source of these between-family effects[51,62].

Predominant theories of GE-interplay imply a positive sign—genetically endowed cognition influences one's proximal environment, and that environment, in turn, influences one's cognition in continuous, reciprocal interactions, such as the multiplier theory[34], the transactional model[35] and the bioecological model[36]. In line with this, a meta-analysis on twin research found the heritability of intelligence to increase with higher SES in the United States[63]. But this effect is far from consistent. The same meta-analysis did not replicate this finding on data collected outside the United States. Furthermore, a large twin study found negative and null results for heritability by SES interaction regarding mathematics and reading in Florida[64]. Two studies using a similar polygenic marker to ours did not find evidence of GE-interplay for EA[65,66], yet a study with 130,000 adults in the UK found a very small negative interaction of SES with neighborhood quality for fIQ and EA[40]. Crucially, the standardized effect size of this interaction was ($\beta < 0.02$ SD)—translating to less than a third IQ point throughout one's entire life —in turn having no practical effect for the individual.

We did not find any significant interaction between schooling and SES or cogPGS. One strength of our design is that schooling is, in principle, independent from cogPGS and SES. We are not aware of any other research looking at GE-interplay with schooling. A recent study did find low PGS children in high-SES schools to continue with mathematics much longer than genetically similar children in low-SES schools[39]. However, we see their result as more relevant to inform a cogPGS-SES effect rather than the gene-by-schooling interaction.

We expected our interaction terms to either compensate or accelerate preexisting differences. Schooling, for example, could increase (i.e., Mathew's effect) economic/genetic inequality or lessen these differences between children (i.e., catch-up effect). The Coleman Report, a seminal study with more than half a million students and over 3000 schools in 1966, controversially concluded schools did not contribute to widening achievement gaps between children[29]. Conversely, there is some evidence that schooling might lessen socioeconomic disparities between children (i.e., a catch-up effect) for cognitive skills[67].

Our study indicates schooling to not be a major driving force for either increasing or decreasing differences due to SES or cogPGS. Yet, we emphasize caution in interpreting these null effects as our range of schooling (3rd–5th grade) was limited, and Bayesian analysis showed that an effect of less than 0.02 SD could not be

ruled out (Fig. 3). Since any interaction effect with schooling could accumulate—that is, continue to increase each year—a very small (<0.05 SD) effect size could be of practical relevance[54,55]. For example, an interaction with schooling as small as 0.02 SD could accumulate over five years to 0.1 SD or roughly 30% of the largest SES main effect (i.e., cIQ = 0.29). This is, of course, a simplistic scenario assuming no counteracting mechanisms, yet it illustrates how very small effect sizes can become consequential[68,69]. In contrast, cogPGS–SES' interaction is a lifelong effect and does not have the potential to accumulate in the same way. However, our sample had a slightly lower SES (Cohen's $d = -0.23$) than the average for the United States, therefore, we cannot rule out an interaction at the lower tail of the SES distribution.

One limitation of this study is that the 1.1 million individuals used to estimate the cogPGS are heavily biased towards those of European descent and from higher SES areas[30,70]. This means our results regarding genetics should only be generalized to white populations. Furthermore, GWAS methods cannot detect certain types and sizes of GE interactions since they are intended to detect additive effects[71]. Another consideration is how to interpret findings with multi-trait GWAS's—in our case cognitive ability, mathematics, and educational attainment—since one of the supplementary phenotypes could be driving the results[72]. The extent of this issue depends on (1) the relative sample size differences between the GWAS included and (2) the genetic correlation of these traits. In our case, there are sample size differences between educational attainment and cognitive ability, yet the very high (r ~ 0.75) genetic correlation between these traits most likely mitigates this issue[73]. Lastly, a strength of our study is that we controlled for ancestry-based genetic PCs in the full model, rather than just correcting cogPGS. While this means SES's relationship with cognition is controlled for population stratification, it also brought the limitation that we had to exclude subjects without DNA, resulting in the average level of SES increasing.

We found that schooling causes relatively large increases in children's intelligence. The two years of schooling (3rd to 5th grade) caused a larger difference in intelligence than either SES or cogPGS. However, schooling did not change the rank order of individuals' intelligence. This was shown by the lack of significant two-way interactions between Schooling, SES, and cogPGS, although our power to detect potentially meaningful small effects for schooling was limited. Intriguingly, we did not find any interaction between SES and cogPGS, this means that children's genetic differences do not matter more, or less, for intelligence dependent upon their SES background.

METHODS

We used data from the Adolescent Brain Cognitive Development (ABCD) study, which was reviewed and approved by the central institutional review board at the University of California, San Diego. Written parental informed consent, along with child assent, was obtained for all participants. Part of the sample included siblings, along with non-representative amounts of monozygotic and dizygotic twins[47,74]. To correct this, we randomly selected one child per family. We also excluded children that reported ever repeating a grade, as this would bias our statistical analysis. This resulted in a total of 6567 children between grades three and five enrolled in regular private or public schooling.

Behavioral and demographic measures

Since recruitment was continuous throughout the year, this allowed us to measure schooling in months. We excluded summer vacation (15 June until 15 August) from our schooling variable by coding children recruited during summer as having the last month of schooling from the previous grade. This resulted in a variable

with 10 months of schooling per year. Chronological age was also measured in months to preserve anonymity.

Our endogenous variables were working memory (WM), fluid (fIQ), and crystallized intelligence (cIQ) from the NIH toolbox cognition battery[75]. These tasks have been shown to have good reliability and validity[59,76,77]. cIQ consisted of a picture vocabulary task and an oral reading recognition task. fIQ consisted of five tasks; pattern comparison processing speed test, a list-sorting WM test, a picture sequence memory test, a flanker task, and a dimensional change card sort task. One of the fluid tasks, list-sorting WM, was used as our WM measure. We used the sum scores provided by the NIH toolbox, rather than latent factors, to facilitate comparison with other studies. All outliers were brought to the fence (Tukey/Boxplot Method). Variables were then standardized to a mean of zero and a standard deviation of one.

SES was defined as the first principal component from a probabilistic PCA of total household income, highest parental education, and neighborhood quality. We used a probabilistic PCA because of the decent amounts of non-overlapping missing data in parental income ($n = 474$, ~7%) and neighborhood quality ($n = 263$, ~4%). Children missing more than one of the three measures were excluded ($n = 45$). Parental education was recorded to reflect middle school or less (1), some high school (2), high school graduate (3), some college/associates degree (4), bachelor's degree (5), a master's degree (6), or professional degree (7). Our measure of neighborhood quality was the area deprivation index calculated from the American Community Survey using the address of primary residency[78]. The SES composite and each subcomponent were standardized with a mean of zero and a standard deviation of one.

Genetic measures

Genotyping was done by the ABCD study and the data provided to us. Saliva samples were collected at the baseline visit the genotyping was performed using the Smokescreen array[79], consisting of 646,247 genetic variants.

Quality control (QC), imputation, and genetic PCA were performed by the National Bioinformatics Infrastructure Sweden (NBIS), as a service contracted by us. Before imputation, SNPs were excluded if they had high levels of missing data (SNP call rate <98%), departed from Hardy–Weinberg equilibrium as calculated in the lfa R package (sHWE) ($P < 1 \times 10^{-6}$), or had minor allele frequencies (MAF) <1%. Moreover, individuals with an absolute autosomal heterozygosity >0.2 or more than 2% missing genotypes were excluded. These filtering steps resulted in a cleaned dataset of 10,069 individuals and 430,622 variants. Subsequently, haplotypes were pre-phased with SHAPEIT2. Genetic markers were imputed using the IMPUTE4 software. As the reference population, we used the 1000 Genomes haplotypes—Phase 3 integrated variant set release in NCBI build 37 (hg19) coordinates. This is a mixed, multi-ethnic population dataset consisting of 2504 samples and 5008 haplotypes from populations of Europeans, Africans, East Asians, Southern Asians, and Americans (https://mathgen.stats.ox.ac.uk/impute/1000GP_Phase3.html). One of the main advantages of this imputation approach is that it provides better concordance in diverse human populations (for more information, see refs. [80,81]). After imputation, genotypes with an INFO score <0.3 or a MAF <0.001% were excluded. The final number of SNPs after imputation was 40,637,119 in a total of 10,069 individuals.

To check for outliers and to control for population structure, principal components of SNPs were obtained using the principal component analysis (PCA) module as implemented in RICOPILI. First, SNPs were pruned to ensure that there is little linkage disequilibrium between SNPs (R2 < 0.2, 200 SNPs window: plink – indep-pairwise 200 100 0.2). The LD pruning was repeated until 100 K SNPs were reached. Then, the resulting SNPs go into the PCA. We used the first 20 principal components (PCs) from that genetic PCA as covariates in our main analyses.

With the SNPs QCed and imputed, our group then created polygenic scores for cognitive performance (here called "cogPGS") for each participant using PRSice-2[82]. This was calculated by the sum of effect sizes of thousands of SNPs (weighted by how many of the effect alleles were present in each individual) that were discovered by a large genome-wide association study on educational attainment, mathematical ability, and general cognitive ability[30]. That study has available all effects sizes and p values of their SNPs on the website of the Social Science Genetics Association Consortium (https://www.thessgac.org/data). We used the data available by the consortium from a multi-trait analysis of GWAS[83], which, in our case, represents a joint polygenic score focused on a GWAS of cognitive performance and complemented by information from a GWAS on educational attainment, a GWAS on the highest-level math class completed, and a GWAS on self-reported math ability. This joint analysis is ideal because pairwise genetic correlations of these traits were high[30]. Furthermore, these GWAS had hundreds of thousands of individuals, and such a large sample size allows new studies to detect effects in samples of a few hundred individuals with 80% statistical power[30].

For the creation of cogPGS in our samples, we performed clumping and pruning to remove nearby SNPs that are correlated with one another. We used the SNPs below the original threshold of $P < 5e-5$ from the GWAS as we wanted to use the full GWAS sample (limited SNP release due to anonymity concerns). After adopting a clumping sliding window of 250 kb, with the linkage disequilibrium clumping set to $r^2 > 0.25$, this resulted in 5255 SNPs for our dataset. Finally, we normalized the polygenic scores (mean = 0, sd = 1).

Statistical analysis

We used random intercept mixed-effects models to predict the effect of schooling from our three cognitive variables of interest (i.e., fIQ, cIQ, and WM); Y denotes these outcome variables. Every model included a random intercept per collection site (j) to account for the clustering of individuals (i). Models were fit using maximum likelihood estimation with the R package lme4 (v. 1.1–27.1)[84,85]. All results are reported in standard units (mean = 0, SD = 1), unless specified otherwise. P values on all models were derived using Satterthwaite's degrees of freedom method with the lmerTest package (v. 3.1-3)[86], with an alpha level of 0.05.

We broke the equation into four steps to more clearly convey the identification assumptions behind our estimands[87]. A fuzzy regression discontinuity design (Eq. 1) was used to isolate the effect of one year of schooling from chronological age. This causal inference method is one of the best to isolate the effect of schooling, yet it relies on two critical assumptions: (1) age-based allocation rule and (2) linearity of the within-grade regression. Previous research has shown that, in practice, violation of these assumptions does not substantially bias the schooling or age coefficient[18,88]. Data collection was continuous throughout the school year therefore, we correct for months of schooling rather than grade, yet our instrument (age-based grade allocation) is on the level of grade. We included post hoc sensitivity checks using grades 3–5 and comparing the two most frequent, grades 4 and 5, for Eq. 1 (Supplementary Table 9a, b). Lastly, we also compared the schooling coefficients using a z-test on the difference between the coefficients. This method sums the standard error from the coefficients as error variance, assuming no correlation between the coefficients, thereby being conservative[89]. We used an absolute z-value of 1.96, corresponding to an alpha level of 0.05.

$$y_{ij} \sim \alpha_0 + \beta_1(\text{Age}_{ij}) + \beta_2(\text{Schooling}_{ij}) + \gamma_{0j} + \epsilon_{ij} \qquad (1)$$

We then add a fixed effect for cogPGS and SES, while controlling sex and 20 genetic principal components (Eq. 2). β_*

represents 20 fixed effects for each principal component. This model allowed us to estimate the causal effect of one year of schooling, along with estimating the independent effects of chronological age, sex, cogPGS, and SES. Crucially, schooling is the only exogenous variable. It is important to highlight that by adding 20 genetic principal components to the model, we are controlling all variables for population stratification, most notably SES's relationship with intelligence.

$$y_{ij} \sim \alpha_0 + \beta_1(\text{Age}_{ij}) + \beta_2(\text{Shooling}_{ij}) + \beta_3(\text{SES}_{ij}) + \beta_4(\text{cogPGS}_{ij}) \\ + \beta_5(\text{SexMale}_{ij}) + \beta_*(20\text{PC}_{ij}) + \gamma_{0j} + \epsilon_{ij}$$

(2)

In Eq. 3, we were interested in determining if schooling interacted with cogPGS (i.e., β_6) or SES (i.e., β_7). To properly isolate schooling's interaction effects and avoid spurious results, we included interactions of no interest with age (i.e., β_8 and β_9). We were also interested to see if SES interacted with cogPGS (i.e., β_{10}), yet we emphasize caution in interpreting this term as any interaction found could be spurious due to endogeneity between SES and cogPGS[41,90]. We refer to these interactions as GE-interplay, highlighting that they could be caused by gene-by-environment interaction (GxE) and/or gene-by-environment correlation (rGE). rGE refers to the fact that genotypes and environments are not randomly distributed, while GxE is an interaction in the classical sense whereby the environment (or genotype) interacts with the other[91]. Due to schooling's exogenous nature, a significant interaction would be most likely due to GxE, yet this is not the case for an interaction between SES and cogPGS which could be the result of either GxE or rGE.

$$y_{ij} \sim \alpha_0 + \beta_1(\text{Age}_{ij}) + \beta_2(\text{Schooling}_{ij})\beta_3(\text{SES}_{ij}) + \beta_4(\text{cogPGS}_{ij}) \\ + \beta_5(\text{SexMale}_{ij}) + \beta_6(\text{Schooling} \times \text{cogPGS}_{ij})\beta_7(\text{Schooling} \times \text{SES}_{ij}) \\ + \beta_8(\text{Age} \times \text{cogPGS}_{ij}) + \beta_9(\text{Age} \times \text{SES}_{ij}) \\ + \beta_{10}(\text{cogPGS} \times \text{SES}_{ij}) + \beta_*(20\text{PC}_{ij}) + \gamma_{0j} + \epsilon_{ij}$$

(3)

Lastly, we were interested in the presence of a three-way interaction between schooling, SES, and cogPGS (Eq. 4, β_{11}). To accomplish this, we added a three-way interaction term, which was then compared, using a likelihood-ratio test to Eq. 3.

$$y_{ij} \sim \alpha_0 + \beta_1(\text{Age}_{ij}) + \beta_2(\text{Schooling}_{ij}) + \beta_3(\text{SES}_{ij}) + \beta_4(\text{cogPGS}_{ij}) \\ + \beta_5(\text{SexMale}_{ij}) + \beta_6(\text{Schooling} \times \text{cogPGS}_{ij}) \\ + \beta_7(\text{Schooling} \times \text{SES}_{ij}) + \beta_8(\text{Age} \times \text{cogPGS}_{ij}) \\ + \beta_9(\text{Age} \times \text{SES}_{ij}) + \beta_{10}(\text{cogPGS} \times \text{SES}_{ij}) \\ + \beta_{11}(\text{Schooling} \times \text{cogPGS} \times \text{SES}_{ij}) + \beta_*(20\text{PC}_{ij}) + \gamma_{0j} + \epsilon_{ij}$$

(4)

SES subcomponents
In cases of significant SES composite findings, we also reported the results for each SES subcomponent separately. This was done for two reasons: (1) research showing specific SES measures to relate to intelligence more than others, and (2) call in the literature to treat SES measures independently[48–50]. We considered the entire subcomponent analysis a family, therefore, p values were corrected using a false discovery rate (FDR) for the number of terms tested.

Post hoc sibling analysis
To estimate the within and between-family effect cogPGS on cIQ, fIQ, and WM we conducted a sibling analysis using the procedure described by Selzam and colleagues 2019[51]. This allowed us to estimate the within-family direct genetic effects (β_w) and the between-family genetic effects (β_B). For families to be included in

the analysis they needed to have at least two non-identical siblings. We considered the entire sibling analysis a family; therefore, p values were corrected using false discovery rate (FDR) for the number of terms of interest tested. SES has been shown to largely explain between-family effects therefore cogPGS in our main analysis is, to some extent, controlled for these between-family effects[51,62].

Post hoc g analysis
We fit nine cognitive tasks into a single-factor solution using confirmatory factor analysis with the Lavaan package in R (Supplementary Fig. 2)[92]. This included all the tasks from the NIH toolbox cognition battery along with the overall score on a WM nback task and the total raw score from matrix reasoning on the Wechsler Intelligence Scale for Children. A single factor solution fit the data well (RMSEA = 0.03 and CFI = 0.987) when correlated covariances were added between the two crystallized tasks, the two WM tasks, and a cluster of three tasks sharing similar attributes (Flanker, Pattern comparison & Card sorting). We then extracted factor scores and fit them to Eqs. 1–4 while using FDR correction for multiple comparisons.

Post hoc European ancestry analysis
CogPGS has limited prediction accuracy in non-European samples[57]. We therefore, subsetted the data using 4-means clustering on the first two principal components ($n = 3751$), resulting in the exclusion of 43% of the sample[58]. The proportion is equal to the full ABCD sample, which is representative of the United States. Yet this subsample had significantly higher SES (Cohen's $d = 0.80$., $p < 0.001$). All variables were (re)standardized to this population with a mean of 0 and an SD of 1. We then fit Eq. 4 on this subset for cIQ, fIQ, WM, and g and used FDR correction on the p values (Supplementary Table 7).

Bayesian hierarchical modeling
To have a better understanding of our results, we conducted Bayesian mixed-effects models of Eq. 3 with the brms package (v. 2.15.0)[93] using Markov Chain Monte Carlo sampling with 10,000 iterations (1000 warmup) in 15 chains. We set weakly informative, normally distributed priors for our terms of interest, keeping the default non-informative priors unless otherwise specified. For cIQ age (β_1) and schooling (β_2) had normally distributed priors with a mean of .20 and a standard deviation (SD) of 0.15. CogPGS (β_4) had a prior with a mean of 0.15 and an SD of 0.1, while SES (β_3) had a larger mean of 0.25 and an SD of 0.2. All two-way interaction terms (β_{6-10}) had two-tailed priors with a mean of zero and an SD of 0.1. Gender and the 20 PCs had default flat priors. For fIQ and WM we set identical priors; these were a mean of 0.15 and an SD of 0.15 for age (β_1) and schooling (β_2). CogPGS (β_4) had a prior with a mean of 0.1 and an SD of 0.1, while SES (β_3) was double that with a mean of 0.2 and an SD of 0.2. All two-way interaction terms (β_{6-10}) had identical priors to cIQ, along with flat priors for gender and each principal component. We also fit Eq. 4 to test the null of the three-way interaction term where priors were identical to Eq. 3 while the three-way term (β_{11}) had a mean of 0 and an SD of 0.1. Notably, our results for the interactions in Eqs. 3 and 4 did not meaningfully change in comparison to flat priors.

We then computed the 95% highest-density intervals (HDI) of the posterior distributions for our three interaction terms of interest (β_6, β_7, β_{10}, β_{11}) for cIQ, fIQ, and WM. Crucially, we used two ROPE boundaries in standard units of 0.05 and 0.02. This was done for two reasons, (1) a lack of consensus in the field on a minimal effect size of interest, and (2) schooling representing a cumulative process. Since the schooling interaction terms (β_6, β_7, β_{11}) are cumulative, they could warrant a smaller minimal effect size of interest than the cogPGS-SES' interaction[54–56]. We then

calculated the percentage of the estimated 95% HDI lying within our ROPE boundaries, giving us the probability of the null hypothesis. If 95% of the estimates' posterior distribution lies within ROPE boundaries, the null hypothesis can be confirmed as any probable effect is considered too small to be meaningful.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

The data to replicate this analysis is on the NIHM data archive (https://nda.nih.gov) under release 3.0 "ABCDstudyNDA".

CODE AVAILABILITY

The code to replicate all analysis and figures will be made available upon publication at https://github.com/njudd/ABCDschooling.

REFERENCES

1. Calvin, C. M. et al. Childhood intelligence in relation to major causes of death in 68 year follow-up: prospective population study. *BMJ* **357**, j2708 (2017).
2. Batty, G. D. et al. IQ in early adulthood and mortality by middle age: cohort study of 1 million Swedish men. *Epidemiology* **20**, 100–109 (2009).
3. Deary, I. J., Cox, S. R. & Hill, W. D. Genetic variation, brain, and intelligence differences. *Mol. Psychiatry* https://doi.org/10.1038/s41380-021-01027-y (2021).
4. Deary, I. J., Strand, S., Smith, P. & Fernandes, C. Intelligence and educational achievement. *Intelligence* **35**, 13–21 (2007).
5. Frey, M. C. & Detterman, D. K. Scholastic assessment or g? The relationship between the Scholastic Assessment Test and general cognitive ability. *Psychol. Sci.* **15**, 373–378 (2004).
6. Flynn, J. R. Massive IQ gains in 14 nations: what IQ tests really measure. *Psychol. Bull.* **101**, 171–191 (1987).
7. Scarr, S. & McCartney, K. How people make their own environments: a theory of genotype–> environment effects. *Child Dev.* **54**, 424 (1983).
8. van Ijzendoorn, M. H., Juffer, F. & Poelhuis, C. W. K. Adoption and cognitive development: a meta-analytic comparison of adopted and nonadopted children's IQ and school performance. *Psychol. Bull.* **131**, 301–316 (2005).
9. Polderman, T. J. C. et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
10. Davies, G. et al. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol. Psychiatry* **16**, 996–1005 (2011).
11. Bouchard, T. J. Jr. Genetic influence on human intelligence (Spearman's g): how much? *Ann. Hum. Biol.* **36**, 527–544 (2009).
12. Haworth, C. M. A. et al. The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol. Psychiatry* **15**, 1112–1120 (2010).
13. Ritchie, S. J. & Tucker-Drob, E. M. How much does education improve intelligence? A meta-analysis. *Psychol. Sci.* **29**, 1358–1369 (2018).
14. Ceci, S. J. How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Dev. Psychol.* **27**, 703–722 (1991).
15. Lazar, I. et al. Lasting effects of early education: a report from the consortium for longitudinal studies. *Monogr. Soc. Res. Child Dev.* **47**, 1–151 (1982).
16. Clouston, S. A. P. et al. Benefits of educational attainment on adult fluid cognition: international evidence from three birth cohorts. *Int. J. Epidemiol.* **41**, 1729–1736 (2012).
17. Brinch, C. N. & Galloway, T. A. Schooling in adolescence raises IQ scores. *Proc. Natl Acad. Sci. USA* **109**, 425–430 (2012).
18. Luyten, H. An empirical assessment of the absolute effect of schooling: regression-discontinuity applied to TIMSS-95. *Oxf. Rev. Educ.* **32**, 397–429 (2006).
19. Cahan, S. & Cohen, N. Age versus schooling effects on intelligence development. *Child Dev.* **60**, 1239–1249 (1989).
20. Baltes, P. B. & Reinert, G. Cohort effects in cognitive development of children as revealed by cross-sectional sequences. *Dev. Psychol.* **1**, 169 (1969).
21. Roberts, G. et al. Schooling duration rather than chronological age predicts working memory between 6 and 7 years: Memory Maestros Study. *J. Dev. Behav. Pediatr.* **36**, 68–74 (2015).
22. Lubinski, D. Cognitive epidemiology: with emphasis on untangling cognitive ability and socioeconomic status. *Intelligence* **37**, 625–633 (2009).
23. Tucker-Drob, E. M. How many pathways underlie socioeconomic differences in the development of cognition and achievement? *Learn. Individ. Differ.* **25**, 12–20 (2013).
24. von Stumm, S. & Plomin, R. Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence* **48**, 30–36 (2015).
25. von Stumm, S. et al. Predicting educational achievement from genomic measures and socioeconomic status. *Dev. Sci.* **23**, e12925 (2020).
26. Hart, S. A., Little, C. & van Bergen, E. Nurture might be nature: cautionary tales and proposed solutions. *NPJ Sci. Learn* **6**, 2 (2021).
27. Smith-Woolley, E. et al. Differences in exam performance between pupils attending selective and non-selective schools mirror the genetic differences between them. *NPJ Sci. Learn* **3**, 3 (2018).
28. Trzaskowski, M. et al. Genetic influence on family socioeconomic status and children's intelligence. *Intelligence* **42**, 83–88 (2014).
29. Coleman, J. S. et al. The coleman report. *Equality of Educational Opportunity* (1966).
30. Lee, J. J. et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
31. Abdellaoui, A. et al. Genetic correlates of social stratification in Great Britain. *Nat. Hum. Behav.* **3**, 1332–1342 (2019).
32. Judd, N. et al. Cognitive and brain development is independently influenced by socioeconomic status and polygenic scores for educational attainment. *Proc. Natl Acad. Sci. USA* **117**, 12411–12418 (2020).
33. Sauce, B. & Matzel, L. D. The paradox of intelligence: heritability and malleability coexist in hidden gene-environment interplay. *Psychol. Bull.* **144**, 26–47 (2018).
34. Dickens, W. T. & Flynn, J. R. Heritability estimates versus large environmental effects: the IQ paradox resolved. *Psychol. Rev.* **108**, 346–369 (2001).
35. Tucker-Drob, E. M., Briley, D. A. & Harden, K. P. Genetic and environmental influences on cognition across development and context. *Curr. Dir. Psychol. Sci.* **22**, 349–355 (2013).
36. Bronfenbrenner, U. & Ceci, S. J. Nature-nuture reconceptualized in developmental perspective: a bioecological model. *Psychol. Rev.* **101**, 568–586 (1994).
37. Trejo, S. et al. Schools as moderators of genetic associations with life course attainments: evidence from the WLS and Add Heath. *Sociol. Sci.* **5**, 513–540 (2018).
38. Schmitz, L. L. & Conley, D. The effect of Vietnam-Era conscription and genetic potential for educational attainment on schooling outcomes. *Econ. Educ. Rev.* **61**, 85–97 (2017).
39. Harden, K. P., Domingue, B. W. & Belsky, D. W. Genetic associations with mathematics tracking and persistence in secondary school. *NPJ Sci. Learn.* **5**, 1 (2020).
40. Rask-Andersen, M., Karlsson, T., Ek, W. E. & Johansson, Å. Modification of heritability for educational attainment and fluid intelligence by socioeconomic deprivation in the UK Biobank. *Am. J. Psychiatry* **178**, 625–634 (2021).
41. Conley, D. et al. Is the effect of parental education on offspring biased or moderated by genotype? *Socio. Sci.* **2**, 82–105 (2015).
42. Demange, P. A. et al. Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat. Genet.* **53**, 35–44 (2021).
43. Borghans, L., Golsteyn, B. H. H., Heckman, J. J. & Humphries, J. E. What grades and achievement tests measure. *Proc. Natl Acad. Sci. USA* **113**, 13354–13359 (2016).
44. Bull, R., Espy, K. A. & Wiebe, S. A. Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev. Neuropsychol.* **33**, 205–228 (2008).
45. Klingberg, T. Training and plasticity of working memory. *Trends Cogn. Sci.* **14**, 317–324 (2010).
46. Sauce, B., Wiedenhoeft, J., Judd, N. & Klingberg, T. Change by challenge: a common genetic basis behind childhood cognitive development and cognitive training. *NPJ Sci. Learn* **6**, 1–16 (2021).
47. Garavan, H. et al. Recruiting the ABCD sample: design considerations and procedures. *Dev. Cogn. Neurosci.* **32**, 16–22 (2018).
48. Strenze, T. Intelligence and socioeconomic success: a meta-analytic review of longitudinal research. *Intelligence* **35**, 401–426 (2007).
49. Duncan, G. J. & Magnuson, K. A. *Socioeconomic Status, Parenting and Child Development* (Lawrence Erlbaum Associates, 2003).
50. Reardon, S. F. The widening income achievement gap. *Educ. Leadersh.* **70**, 10–16 (2013).
51. Selzam, S. et al. Comparing within- and between-family polygenic score prediction. *Am. J. Hum. Genet.* **105**, 351–363 (2019).
52. Trejo, S. & Domingue, B. W. Genetic nature or genetic nurture? Introducing social genetic parameters to quantify bias in polygenic score analyses. *Biodemography Soc. Biol.* **64**, 187–215 (2018).
53. Keller, M. C. Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol. Psychiatry* **75**, 18–24 (2014).

54. Kraft, M. A. Interpreting effect sizes of education interventions. *Educ. Res.* **49**, 241–253 (2020).

55. Abelson, R. P. A variance explanation paradox: when a little is a lot. *Psychol. Bull.* **97**, 129–133 (1985). vol.Preprint at.

56. Pogrow, S. How effect size (Practical Significance) misleads clinical practice: the case for switching to practical benefit to assess applied research findings. *Am. Stat.* **73**, 223–234 (2019).

57. Mostafavi, H. et al. Variable prediction accuracy of polygenic scores within an ancestry group. *Elife* **9**, e48376 (2020).

58. Pain, O. et al. Evaluation of polygenic prediction methodology within a reference-standardized framework. *PLoS Genet.* **17**, e1009021 (2021).

59. Anokhin, A. P. et al. Age-related changes and longitudinal stability of individual differences in ABCD neurocognition measures. *Dev. Cogn. Neurosci.* **54**, 101078 (2022).

60. Schmitz, L. & Conley, D. Modeling gene-environment interactions with quasi-natural experiments. *J. Pers.* **85**, 10–21 (2017).

61. Dudbridge, F. & Fletcher, O. Gene-environment dependence creates spurious gene-environment interaction. *Am. J. Hum. Genet.* **95**, 301–307 (2014).

62. Wang, B. et al. Robust genetic nurture effects on education: a systematic review and meta-analysis based on 38,654 families across 8 cohorts. *Am. J. Hum. Genet.* https://doi.org/10.1016/j.ajhg.2021.07.010 (2021).

63. Tucker-Drob, E. M. & Bates, T. C. Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychol. Sci.* **27**, 138–149 (2016).

64. Figlio, D. N., Freese, J., Karbownik, K. & Roth, J. Socioeconomic status and genetic influences on cognitive development. *Proc. Natl Acad. Sci. USA* **114**, 13441–13446 (2017).

65. de Zeeuw, E. L. et al. The moderating role of SES on genetic differences in educational achievement in the Netherlands. *NPJ Sci. Learn* **4**, 13 (2019).

66. Allegrini, A. G. et al. Multivariable G-E interplay in the prediction of educational achievement. *PLoS Genet.* **16**, e1009153 (2020).

67. Downey, D. B. & Condron, D. J. Fifty years since the Coleman report: rethinking the relationship between schools and inequality. *Sociol. Educ.* **89**, 207–220 (2016).

68. Anvari, F. et al. x Not All Effects Are Indispensable: Psychological Science Requires Verifiable Lines of Reasoning for Whether an Effect Matters. *Perspect Psychol Sci: J Assoc Psychol Sci*, https://doi.org/10.1177/17456916221091565. (2022).

69. Funder, D. C. & Ozer, D. J. Evaluating effect size in psychological research: sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* **2**, 156–168 (2019).

70. Fry, A. et al. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).

71. Domingue, B., Trejo, S., Armstrong-Carter, E. & Tucker-Drob, E. Interactions between polygenic scores and environments: methodological and conceptual challenges. *Sociol. Sci.* **7**, 365–386 (2020).

72. Becker, J. et al. Resource profile and user guide of the polygenic index repository. *Nat. Hum. Behav.* **5**, 1744–1758 (2021).

73. Abdellaoui, A. & Verweij, K. J. H. Dissecting polygenic signals from genome-wide association studies on human behaviour. *Nat. Hum. Behav.* **5**, 686–694 (2021).

74. Luciana, M. et al. Adolescent neurocognitive development and impacts of substance use: overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. *Dev. Cogn. Neurosci.* **32**, 67–79 (2018).

75. Heeringa, S. G. & Berglund, P. A. A Guide for population-based analysis of the adolescent brain cognitive development (ABCD) study baseline data. Preprint at https://doi.org/10.1101/2020.02.10.942011 (2020).

76. Heaton, R. K. et al. Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. *J. Int. Neuropsychol. Soc.* **20**, 588–598 (2014).

77. Gershon, R. C. et al. NIH Toolbox for assessment of neurological and behavioral function. *Neurology* **80**, S2–S6 (2013).

78. Kind, A. J. H. et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. *Ann. Intern. Med.* **161**, 765–774 (2014).

79. Baurley, J. W., Edlund, C. K., Pardamean, C. I., Conti, D. V. & Bergen, A. W. Smokescreen: a targeted genotyping array for addiction research. *BMC Genomics* **17**, 145 (2016).

80. Howie, B. N., Donnelly, P. & Marchini, J. 1,000 genomes haplotypes—Phase 3 integrated variant set release in NCBI build 37 (hg19) coordinates. Preprint at (2015).

81. 1000 Genomes Project Consortium. et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).

82. Choi, S. W. & O'Reilly, P. F. PRSice-2: polygenic risk score software for biobank-scale data. *Gigascience* **8**, giz082 (2019).

83. Turley, P. et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229–237 (2018).

84. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015). vol.Preprint at.

85. R Core Team. R: a language and environment for statistical computing. http://www.R-project.org/ (2014).

86. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw., Artic.* **82**, 1–26 (2017).

87. Lundberg, I., Johnson, R. & Stewart, B. M. What is your estimand? Defining the target quantity connects statistical evidence to theory. *Am. Sociol. Rev.* **86**, 532–565 (2021).

88. Cliffordson, C. Methodological issues in investigations of the relative effects of schooling and age on school performance: the between-grade regression discontinuity design applied to Swedish TIMSS 1995 data. *Educ. Res. Eval.* **16**, 39–52 (2010).

89. Loughnan, R. J. et al. Gene-experience correlation during cognitive development: evidence from the adolescent brain cognitive development (ABCD) StudySM. Preprint at *bioRxiv* https://doi.org/10.1101/637512 (2021).

90. Hanscombe, K. B. et al. Socioeconomic status (SES) and children's intelligence (IQ): in a UK-representative sample SES moderates the environmental, not genetic, effect on IQ. *PLoS ONE* **7**, e30320 (2012).

91. Plomin, R., DeFries, J. C. & Loehlin, J. C. Genotype-environment interaction and correlation in the analysis of human behavior. *Psychol. Bull.* **84**, 309–322 (1977).

92. Rosseel, Y. Lavaan: an R package for structural equation modeling and more. Version 0.5–12 (BETA). *J. Stat. Softw.* **48**, 1–36 (2012).

93. Bürkner, P.-C. brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017). vol. Preprint at.

AUTHOR CONTRIBUTIONS

N.J. made substantial contributions to the conception and design of the work; as well as the analysis and interpretation of data; drafted the work and substantively revised it. B.S. and T.K. have approved the submitted version and both made substantial contributions to the interpretation of the work, additional analyses as well as revising the manuscript.

FUNDING

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41539-022-00148-5.

Correspondence and requests for materials should be addressed to Nicholas Judd.

Reprints and permission information is available at http://www.nature.com/reprints

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