Gene discovery and polygenic prediction from a 1.1-million-person GWAS of educational attainment

AUTHORS:

James J. Lee[1,†], Robbee Wedow[2,3,4,†], Aysu Okbay[5,6,*†], Edward Kong[7], Omeed Maghzian[7], Meghan Zacher[8], Tuan Anh Nguyen-Viet[9], Peter Bowers[7], Julia Sidorenko[10,11], Richard Karlsson Linnér[5,6,12], Mark Alan Fontana[9,13], Tushar Kundu[9], Chanwook Lee[7], Hui Li[7], Ruoxi Li[9], Rebecca Royer[9], Pascal N. Timshel[14,15], Raymond K. Walters[16,17], Emily A. Willoughby[1], Loïc Yengo[10], *23andMe Research Team*[18], *COGENT (Cognitive Genomics Consortium)*[19], *Social Science Genetic Association Consortium*[18], Maris Alver[11], Yanchun Bao[20], David W. Clark[21], Felix R. Day[22], Nicholas A. Furlotte[23], Peter K. Joshi[21,24], Kathryn E. Kemper[10], Aaron Kleinman[23], Claudia Langenberg[22], Reedik Mägi[11], Joey W. Trampush[25,26], Shefali Setia Verma[27], Yang Wu[10], Max Lam[28,29], Jing Hua Zhao[22], Zhili Zheng[10,30], Jason D. Boardman[2,3,4], Harry Campbell[21], Jeremy Freese[31], Kathleen Mullan Harris[32,33], Caroline Hayward[34], Pamela Herd[35], Meena Kumari[20], Todd Lencz[36,37,38], Jian'an Luan[22], Anil K. Malhotra[36,37,38], Andres Metspalu[11], Lili Milani[11], Ken K. Ong[22], John R. B. Perry[22], David J. Porteous[39], Marylyn D. Ritchie[27], Melissa C. Smart[21], Blair H. Smith[40], Joyce Y. Tung[23], Nicholas J. Wareham[22], James F. Wilson[21,34], Jonathan P. Beauchamp[41], Dalton C. Conley[42], Tõnu Esko[11], Steven F. Lehrer[43,44,45], Patrik K. E. Magnusson[46], Sven Oskarsson[47], Tune H. Pers[14,15], Matthew R. Robinson[10,48], Kevin Thom[49], Chelsea Watson[9], Christopher F. Chabris[50], Michelle N. Meyer[51], David I. Laibson[7], Jian Yang[10,52], Magnus Johannesson[53], Philipp D. Koellinger[5,6,12], Patrick Turley[16,17,#], Peter M. Visscher[10,52,*#], Daniel J. Benjamin[9,45,54,*#], and David Cesarini[45,49,55,#]

[1] Department of Psychology, University of Minnesota Twin Cities, Minneapolis, Minnesota, USA
[2] Department of Sociology, University of Colorado Boulder, Boulder, Colorado, USA
[3] Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, Colorado, USA
[4] Institute of Behavioral Science, University of Colorado Boulder, Boulder, Colorado, USA
[5] Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[6] Department of Economics, School of Business and Economics, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[7] Department of Economics, Harvard University, Cambridge, Massachusetts 02138, USA
[8] Department of Sociology, Harvard University, Cambridge, Massachusetts, USA
[9] Center for Economic and Social Research, University of Southern California, Los Angeles, California, USA
[10] Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia
[11] Estonian Genome Center, University of Tartu, Tartu, Estonia
[12] Institute for Behavior and Biology, Erasmus University Rotterdam, Rotterdam, the Netherlands
[13] Center for the Advancement of Value in Musculoskeletal Care, Hospital for Special Surgery, New York, New York, USA

[14] The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, University of Copenhagen, Faculty of Health and Medical Sciences, Copenhagen 2100, Denmark
[15] Statens Serum Institut, Department of Epidemiology Research, Copenhagen 2300, Denmark
[16] Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA
[17] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA
[18] A full list of members and affiliations appears at the end of the paper
[19] A list of members and affiliations appears in the Supplementary Note
[20] Institute for Social and Economic Research, University of Essex, Colchester, UK
[21] Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, Scotland
[22] MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, Cambridge, UK
[23] 23andMe, Inc., Mountain View, California 94041, USA
[24] Institute of Social and Preventive Medicine, University Hospital of Lausanne, Lausanne, Switzerland
[25] BrainWorkup, LLC, Santa Monica, California, USA
[26] Department of Psychiatry and Behavioral Sciences, Keck School of Medicine, University of Southern California, Los Angeles, California, USA
[27] Department of Biomedical and Translational Informatics, Geisinger Health System, Danville, Pennsylvania, USA
[28] Institute of Mental Health, Singapore, Singapore
[29] Genome Institute Singapore, Singapore
[30] The Eye Hospital, School of Ophthalmology & Optometry, Wenzhou Medical University, Wenzhou, Zhejiang, China
[31] Department of Sociology, Stanford University, Stanford, California, USA
[32] Department of Sociology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[33] Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[34] MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland
[35] La Follette School of Public Affairs, University of Wisconsin-Madison, Madison, Wisconsin, USA
[36] Departments of Psychiatry and Molecular Medicine, Hofstra Northwell School of Medicine, Hempstead, New York, USA
[37] Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, Manhasset, New York, USA
[38] Psychiatry Research, The Zucker Hillside Hospital, Glen Oaks, California, USA
[39] Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland
[40] Division of Population Health Sciences, Ninewells Hospital and Medical School, University of Dundee, Dundee, Scotland

[41] Department of Economics, University of Toronto, Toronto, Ontario, Canada
[42] Department of Sociology, Princeton University, Princeton, New Jersey, USA
[43] School of Policy Studies, Queen's University, Kingston, Ontario, Canada
[44] Department of Economics, New York University Shanghai, Pudong, Shanghai, China
[45] National Bureau of Economic Research, Cambridge, MA, USA
[46] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[47] Department of Government, Uppsala University, Uppsala, Sweden
[48] Department of Computational Biology, University of Lausanne, Lausanne, Switzerland
[49] Department of Economics, New York University, New York, New York, USA
[50] Autism and Developmental Medicine Institute, Geisinger Health System, Lewisburg, Pennsylvania, USA
[51] Center for Translational Bioethics and Health Care Policy, Geisinger Health System, Danville, Pennsylvania, USA
[52] Queensland Brain Institute, University of Queensland, Brisbane, Australia
[53] Department of Economics, Stockholm School of Economics, Stockholm, Sweden
[54] Department of Economics, University of Southern California, Los Angeles, California, USA
[55] Center for Experimental Social Science, New York University, New York, New York, USA

† These authors contributed equally.
# These authors jointly directed the work.

* Correspondence to Daniel Benjamin, daniel.benjamin@gmail.com, Aysu Okbay, a.okbay@vu.nl, and Peter Visscher, peter.visscher@uq.edu.au.

ABSTRACT

We conduct a large-scale genetic association analysis of educational attainment in a sample of ~1.1 million individuals and identify 1,271 independent genome-wide-significant SNPs. For the SNPs taken together, we found evidence of heterogeneous effects across environments. The SNPs implicate genes involved in brain-development processes and neuron-to-neuron communication. In a separate analysis of the X chromosome, we identify 10 independent genome-wide-significant SNPs and estimate a SNP heritability of ~0.3% in both men and women, consistent with partial dosage compensation. A joint (multi-phenotype) analysis of educational attainment and three related cognitive phenotypes generates polygenic scores that explain 11-13% of the variance in educational attainment and 7-10% of the variance in cognitive performance. This prediction accuracy substantially increases the utility of polygenic scores as tools in research.

INTRODUCTION

Educational attainment (EA) is moderately heritable[1] and an important correlate of many social, economic, and health outcomes[2,3]. Because of its relationship with many health outcomes, measures of EA are available in most medical data sets. Partly for this reason, EA was the focus of the first large-scale genome-wide association study (GWAS) of a social-science phenotype[4] and has continued to serve as a "model phenotype" for behavioral traits (analogous to height for medical traits). Genetic associations with EA identified via GWAS have been used in follow-up work examining biologizcal[5] and behavioral mechanisms[6,7] and genetic overlap with health outcomes[8,9].

The largest (*N* = 293,723) GWAS of EA to date identified 74 approximately independent SNPs at genome-wide significance (hereafter, lead SNPs) and reported that a 10-million-SNP linear predictor (hereafter, polygenic score) had an out-of-sample predictive power of 3.2%[10]. Here, we expand the sample size to over a million individuals (*N* = 1,131,881). We identify 1,271 lead SNPs. In a subsample (*N* = 694,894), we also conduct genome-wide association analyses of variants on the X chromosome, identifying ten lead SNPs.

The dramatic increase in our GWAS sample size enables us to conduct a number of informative additional analyses. For example, we show that the lead SNPs have heterogeneous effects, and we perform within-family association analyses that probe the robustness of our results. Our biological annotation analyses, which focus on the results from the autosomal GWAS, reinforce the main findings from earlier GWAS in smaller samples, such as the role of many of the prioritized genes in brain development. However, the newly identified SNPs also lead to several new findings. For example, they strongly implicate genes involved in almost all aspects of neuron-to-neuron communication.

We found that a polygenic score derived from our results explains around 11% of EA variance. We also report additional GWAS of three phenotypes that are highly genetically correlated with EA: cognitive (test) performance (*N* = 257,841), self-reported math ability (*N* = 564,698), and hardest math class completed (*N* = 430,445). We identify 225, 618, and 365 lead SNPs, respectively. When we jointly analyze all four phenotypes using a recently developed method[11], we found that the explanatory power of polygenic scores based on the resulting summary statistics increases, to 12% for EA and 7-10% for cognitive performance.

RESULTS

Primary GWAS of *EduYears*

In our primary GWAS, we study EA, which is measured as number of years of schooling completed (*EduYears*). All association analyses were performed at the cohort level in samples restricted to European-descent individuals. We applied a uniform set of quality-control procedures to all cohort-level results. Our final sample-size-weighted meta-analysis produced association statistics for ~10 million SNPs from phase 3 of the 1000 Genomes Project[12].

The quantile-quantile plot of the meta-analysis (**Supplementary Figure 1**) exhibits substantial inflation ($\lambda_{GC} = 2.04$). According to our LD Score regression[13] estimates, only a small share (~5%) of this inflation is attributable to bias (**Supplementary Figure 2**, **Supplementary Table 1**). We used the estimated LD Score intercept (1.11) to generate inflation-adjusted test statistics.

Fig. 1 shows the Manhattan plot of the resulting P values. We identified 1,271 approximately independent (pairwise $r^2 < 0.1$) SNPs at genome-wide significance ($P < 5\times10^{-8}$), 995 of which remain if we adopt the stricter significance threshold ($P < 1\times10^{-8}$) proposed in a recent study (**Supplementary Table 2**, see Online Methods for a description of the clumping algorithm). The **Supplementary Note** and **Supplementary Table 3** reports the results from a conditional-joint analysis[14].

We used a Bayesian statistical framework to calculate winner's-curse-adjusted posterior distributions of the effect sizes of the lead SNPs (Online Methods). We found that the median effect size of the lead SNPs corresponds to 1.7 weeks of schooling per allele; at the 5th and 95th percentiles, 1.1 and 2.6 weeks, respectively. We also examined the replicability of the 162 single-SNP associations ($P < 5\times10^{-8}$) reported from the combined discovery and replication sample (N = 405,073) of the largest previous study[10]. In the subsample of our data (N = 726,808) that did not contribute to the earlier study's analyses, the SNPs replicate at a rate that closely matches theoretical projections derived from our Bayesian framework (**Supplementary Figure 3**).

Within-Family Association Analyses

We conducted within-family association analyses in four sibling cohorts (22,135 sibling pairs) and compared the resulting estimates to those from a meta-analysis that excluded the siblings (N = 1,070,751). The latter association statistics were adjusted for stratification bias using the LD Score intercept. **Fig. 2** shows the observed sign concordance for three sets of approximately independent SNPs, selected using P value cutoffs of 5×10^{-3}, 5×10^{-5}, and 5×10^{-8}. The concordance is substantially greater than expected by chance but weaker than predicted by our Bayesian framework, even after we extend the framework to account for inflation in GWAS coefficients due to assortative mating. In a second analysis based on all SNPs, we estimate that within-family effect sizes are roughly 40% smaller than GWAS effect sizes and that our assortative-mating adjustment explains at most one third of this deflation. (For comparison, when we apply the same method to height, we found that the assortative-mating adjustment fully explains the deflation of the within-family effects.)

Supplementary Note contains analyses and discussion of the possible causes of the remaining deflation we observe for *EduYears*. While the evidence is not conclusive, it suggests that the GWAS effect-size estimates may be biased upward by correlation between EA and a rearing environment conducive to EA. Consistent with this hypothesis, a recent paper[15] reports that a polygenic score for *EduYears* based entirely on parents' non-transmitted alleles is approximately 30% as predictive as a polygenic score based on transmitted alleles. (For height, the analogous estimate is only 6%.) The non-transmitted alleles affect parents' EA but can only influence the child's EA indirectly. If greater parental EA positively influences the rearing environment, then GWAS that control imperfectly for rearing environment will yield inflated estimates. The LD Score regression intercept does not capture this bias because the bias scales with the LD Score in the same way as a direct genetic effect.

Heterogeneous Effect Sizes

Because educational institutions vary across places and time, the effects of specific SNPs may vary across environments. Consistent with such heterogeneity, for the lead SNPs, we reject the joint null hypothesis of homogeneous cohort-level effects (P value = 9.7×10^{-12}; **Supplementary Figure 4**). Moreover, we found that the inverse-variance-weighted mean genetic correlation of *EduYears* across pairs of cohorts in our sample is 0.72 (SE = 0.14), which is statistically distinguishable from one (P value = 0.03).

Our finding of an imperfect genetic correlation replicates earlier results from smaller samples[16,17]. This imperfect genetic correlation is an important factor to consider in power calculations and study design. In the **Supplementary Note**, we report exploratory analyses that aim to identify specific sources of measurement heterogeneity or gene-environment interaction that may explain the imperfect genetic correlation. Unfortunately, the estimates are noisy, and the only strong finding was that SNP heritability was smaller in cohorts whose measure of *EduYears* is derived from questions with fewer response categories.

X-Chromosome GWAS Results

We supplemented our autosomal analyses with association analyses of SNPs on the X chromosome. We first conducted separate association analyses of males (N = 152,608) and females (N = 176,750) in the UK Biobank. We found a male-female genetic correlation close to unity. We also found nearly identical SNP heritability estimates for men and women, which is consistent with partial dosage compensation (i.e., on average the per-allele effect sizes are smaller in women) and implies that any contribution of common variants on the X chromosome to sex differences in the normal-range *variance* of cognitive phenotypes[18] is quantitatively negligible.

Next, we conducted a large (N = 694,894) meta-analysis of summary statistics from mixed-sex analyses (**Supplementary Figure 5**). We identified 10 lead SNPs and estimated a SNP heritability due to the X chromosome of ~0.3% (**Supplementary Table 4**). This heritability is lower than that expected for an autosome of similar length (**Supplementary Figure 6**, **Supplementary Table 5**). We cannot distinguish whether the lower heritability is due to smaller per-allele effect sizes for SNPs on the X chromosome or to the combination of haploidy in males and (partial) X-inactivation in females.

Biological Annotation

For biological annotation, we focus on the results from the autosomal meta-analysis of *EduYears*. Across an extensive set of analyses (see **Supplementary Figure 7** for a flowchart), all major conclusions from the largest previous GWAS of *EduYears*[10] continue to hold but are statistically stronger. For example, we applied the bioinformatics tool DEPICT[19] and found that, relative to other genes, genes near our lead SNPs are overwhelmingly enriched for expression in the central nervous system (**Fig. 3A**, **Supplementary Table 6**).

There are also many novel findings associated with the large number of genes newly implicated by our analyses: At the standard false discovery rate (FDR) threshold of 5%, the bioinformatics tool DEPICT[19] prioritizes 1,838 genes (**Supplementary Table 7**), a tenfold increase relative to the DEPICT results from an earlier GWAS of *EduYears*[10]. In what follows, we distinguish between the 1,703 "newly prioritized" genes and the 135 "previously prioritized" genes.

The **Supplementary Note** contains an extensive analysis of many of the newly prioritized genes and their brain-related functions. Here we highlight two especially noteworthy regularities. First, whereas previously prioritized genes exhibited especially high expression in the brain prenatally, newly prioritized genes show elevated levels of expression both pre- and postnatally (**Fig. 3B**). Many of the newly prioritized genes encode proteins that carry out online brain functions such as neurotransmitter secretion, the activation of ion channels and metabotropic pathways, and synaptic plasticity (**Supplementary Figure 8**).

Second, even though glial cells are at least as numerous as neurons in the human brain[20], gene sets related to glial cells (astrocytes, myelination, and positive regulation of gliogenesis) are absent from those identified as positively enriched (**Supplementary Table 8**). Furthermore, using stratified LD Score regression[21], we estimated relatively weak enrichment of genes highly expressed in glial cells (**Supplementary Table 9**): 1.08-fold for astrocytes (P = 0.07) and 1.09-fold for oligodendrocytes (P = 0.06) versus 1.33-fold for neurons ($P = 2.89\times10^{-11}$). Because myelination increases the speed with which signals are transmitted along axons[22], the absence of enrichment of genes related to glial cells may weigh against the hypothesis that differences across people in cognition are driven by differences in transmission speed.

The results also raise a number of possible targets for functional studies. Among SNPs within 50 kb of lead SNPs, 127 of them are identified by the fine-mapping tool CAVIARBF[23] as likely causal SNPs (posterior probability > 0.9) (**Supplementary Table 10**). Eight of these are non-synonymous, and one of these (rs61734410) is located in *CACNA1H* (**Supplementary Figure 9**), which encodes the pore-forming subunit of a voltage-gated calcium channel that has been implicated in the trafficking of NMDA-type glutamate receptors[24].

Polygenic Prediction

Polygenic predictors derived from earlier GWAS of *EduYears* have proven to be a valuable tool for researchers, especially in the social sciences[6,7]. We constructed polygenic scores for European-ancestry individuals in two prediction cohorts: the National Longitudinal Study of Adolescent to Adult Health (Add Health, N = 4,775), a representative sample of American adolescents; and the Health and Retirement Study (HRS, N = 8,609), a representative sample of Americans over age 50. We measure prediction accuracy by the "incremental R^2": the gain in coefficient of determination (R^2) when the score is added as a covariate to a regression of the phenotype on a set of baseline controls (sex, birth year, their interaction, and 10 principal components of the genetic relatedness matrix).

All scores are based on results from a meta-analysis that excluded the prediction cohorts. Our first four scores were constructed from sets of LD-pruned SNPs associated with *EduYears* at various P-value thresholds: 5×10^{-8}, 5×10^{-5}, 5×10^{-3}, and 1 (i.e., all SNPs). In both cohorts, the predictive power is greater for scores constructed with less stringent thresholds (**Supplementary Figure 10**). The sample-size-weighted mean incremental R^2 increases from 3.2% at $P < 5\times10^{-8}$ to 9.4% at $P \leq 1$. Our fifth score was generated from HapMap3 SNPs using the software LDpred[25]. Rather than dropping SNPs in LD with each other, LDpred is a Bayesian method which weights each SNP by (an approximation to) the posterior mean of its conditional effect, given other SNPs. This score was the most predictive in both cohorts, with an incremental R^2 of 12.7% in AddHealth and 10.6% in HRS (and a sample-size weighted average of 11.4%).

To put the predictive power of this score in perspective, **Fig. 4A** shows the mean college completion rate by polygenic-score quintile. The difference between the bottom and top quintiles in Add Health and HRS is, respectively, 45 and 36 percentage points (see **Supplementary**

Figure 11 for analogous analyses of high school completion and grade retention). **Fig. 4B** compares the incremental R^2 of the score to that of standard demographic variables. The score is a better predictor of *EduYears* than household income and a worse predictor than mother's or father's education. Controlling for all the demographic variables jointly, the score's incremental R^2 is 4.6% (**Supplementary Figure 12**).

We also found that the score has substantial predictive power for a variety of other cognitive phenotypes measured in the prediction cohorts (**Supplementary Figure 13**). For example, it explains 9.2% of the variance in overall grade point average in Add Health.

Because the discovery sample used to construct the score consisted of individuals of European ancestry, we would not expect the predictive power of our score to be as high in other ancestry groups[7,26,27]. Indeed, when our score is used to predict *EduYears* in a sample of African-Americans from the HRS (N = 1,519), the score only has an incremental R^2 of 1.6%, implying an attenuation of 85%. The **Supplementary Note** shows that this amount of attenuation is typical of what has been reported in previous studies.

Related Cognitive Phenotypes and MTAG

We performed genome-wide association analyses of three complementary phenotypes: cognitive performance (N = 257,841), self-reported math ability (*Math Ability*, N = 564,698), and highest math class taken (*Highest Math*, N = 430,445). For cognitive performance, we meta-analyzed published results from the COGENT Consortium[28] with results based on new analyses of the UKB, as did Davies et al.[29]. For the two math phenotypes, we studied new genome-wide analyses in samples of research participants from 23andMe. We identified 225, 618, and 365 genome-wide significant SNPs for *Cognitive Performance*, *Math Ability*, and *Highest Math*, respectively (**Supplementary Figures 14-16**, **Supplementary Tables 11-13**).

We conducted a multi-trait analysis of *EduYears* and our supplementary phenotypes to improve polygenic prediction accuracy. These phenotypes are well suited to joint analysis because their pairwise genetic correlations are high, in all cases exceeding 0.5 (**Supplementary Table 14**). We applied a recently developed method, Multi-Trait Analysis of GWAS, or MTAG[11], to summary statistics for the four phenotypes from meta-analyses that exclude the prediction cohorts. For all four phenotypes, MTAG increases the number of lead SNPs identified at genome-wide significance (**Supplementary Figures 17-20**, **Supplementary Table 15**). **Fig. 4C** shows the incremental R^2 for the polygenic scores based on GWAS and MTAG association statistics (but otherwise constructed using identical methods) when the target phenotype is either *EduYears* (left panel) or *Cognitive Performance* (right panel).

In Add Health, where our measure of cognitive performance is the respondent's score on a test of verbal cognition, the incremental R^2s of the GWAS and MTAG scores are 5.1% and 6.9%, respectively. To obtain a better measure prediction accuracy for cognitive performance, we used an additional validation cohort, the Wisconsin Longitudinal Study (WLS), which administered a cognitive test with excellent retest reliability and psychometric properties similar to those used in our discovery GWAS of cognitive performance. In the WLS, the MTAG score predicts 9.7% of the variance in *Cognitive Performance*, a substantial improvement over the 7.0% predicted by the GWAS score and approximately double the prediction accuracy reported in three recent GWASs of cognitive performance[29–31].

DISCUSSION

The results of this study illustrate what the advocates of GWAS anticipated: as sample sizes get large, thousands of lead SNPs will be identified, and polygenic predictors will attain non-trivial levels of predictive power. However, theoretical projections that failed to consider heterogeneity of effect sizes were optimistic[4]. Our and others' findings[16,17] suggest that imperfect genetic correlation across cohorts will be the norm for phenotypes that, like EA, are environmentally contingent.

For research at the intersection of genetics and neuroscience, the set of 1,271 lead SNPs we identify is a treasure trove for future analyses. For research in social science and epidemiology, the polygenic scores we construct—which explain 11-13% and 7-10% of the variance of EA and cognitive performance, respectively—will prove useful across at least three types of applications.

First, by examining associations between the scores and high-quality measures of endophenotypes, researchers may be able to disentangle the mechanisms by which genetic factors affect EA and cognitive phenotypes. Such studies are already being conducted with polygenic scores from earlier GWAS of EA[6,7], but they can now be well powered in samples as small as those from laboratory experiments. For example, if our polygenic score explains 10% of the variance in an endophenotype, then its effect can be detected at a 5% significance threshold with 80% power in a sample of only 75 individuals. Second, the polygenic scores can be used as control variables in randomized controlled trials (RCTs) of interventions that aim to improve academic and cognitive outcomes. Given the scores' current levels of predictive power, such use can now generate non-trivial gains in statistical power for the RCT. For example, if adding the polygenic score to the set of control variables in an RCT increases their joint explanatory power from 10% to 20%, then the gain in power from including the polygenic score is equivalent to increasing the RCT's sample size by 11% (for such calculations, see the SOM of Rietveld et al.[4]). Third, the polygenic scores can be used as a tool for exploring gene-environment interactions[32], which are known to be important for genetic effects on educational attainment and cognitive performance[1,33].

Our results also highlight two caveats to the use of the polygenic scores in research. First, our within-family analyses suggest that GWAS estimates may overstate the causal effect sizes: if

EA-increasing genotypes are associated with parental EA-increasing genotypes, which are in turn associated with rearing environments that promote EA, then failure to control for rearing environment will bias GWAS estimates. If this hypothesis is correct, some of the predictive power of the polygenic score reflects environmental amplification of the genetic effects. Without controls for this bias, it is therefore inappropriate to interpret the polygenic score for EA as a measure of genetic endowment.

Second, we found that our score for EA has much lower predictive power in an African-American sample than in a European-ancestry sample, and we anticipate that the score would also have reduced predictive power in other non-European-ancestry samples. Therefore, until polygenic scores are available that have as much predictive power in other ancestry groups, the score will be most useful in research that is focused on European-ancestry samples.

ACKNOWLEDGMENTS:

This research was carried out under the auspices of the Social Science Genetic Association Consortium (SSGAC). The research has also been conducted using the UK Biobank Resource under application numbers 11425 and 12512. We acknowledge the Swedish Twin Registry for access to data. The Swedish Twin Registry is managed by Karolinska Institutet and receives funding through the Swedish Research Council under the grant no 2017-00641. This study was supported by funding from the Ragnar Söderberg Foundation (E9/11, E24/15), the Swedish Research Council (421-2013-1061), The Jan Wallander and Tom Hedelius Foundation, an ERC Consolidator Grant (647648 EdGe), the Pershing Square Fund of the Foundations of Human Behavior, and the NIA/NIH through grants P01-AG005842, P01-AG005842-20S2, P30-AG012810, and T32-AG000186-23 to NBER, and R01-AG042568 to USC. A full list of acknowledgments is provided in the **Supplementary Note**.

CONTRIBUTOR LIST FOR THE 23andMe RESEARCH TEAM: Michelle Agee[23], Babak Alipanahi[23], Adam Auton[23], Robert K. Bell[23], Katarzyna Bryc[23], Sarah L. Elson[23], Pierre Fontanillas[23], Nicholas A. Furlotte[23], David A. Hinds[23], Bethann S. Hromatka[23], Karen E. Huber[23], Aaron Kleinman[23], Nadia K. Litterman[23], Matthew H. McIntyre[23], Joanna L. Mountain[23], Carrie A.M. Northover[23], J. Fah Sathirapongsasuti[23], Olga V. Sazonova[23], Janie F. Shelton[23], Suyash Shringarpure[23], Chao Tian[23], Joyce Y. Tung[23], Vladimir Vacic[23], Catherine H. Wilson[23], and Steven J. Pitts[23].

CONTRIBUTOR LIST FOR THE SOCIAL SCIENCE GENETIC ASSOCIATION CONSORTIUM RESEARCH TEAM: Aysu Okbay[5,6], Jonathan P. Beauchamp[41], Mark Alan Fontana[9,13], James J. Lee[1], Tune H. Pers[14,15], Cornelius A. Rietveld[12,56,57], Patrick Turley[16,17], Guo-Bo Chen[52], Valur Emilsson[58,59], S. Fleur W. Meddens[5,12,60], Sven Oskarsson[47], Joseph K.

Pickrell[61], Kevin Thom[49], Pascal Timshel[14,15], Ronald de Vlaming[12,56,57], Abdel Abdellaoui[62], Tarunveer S. Ahluwalia[14,63,64], Jonas Bacelis[65], Clemens Baumbach[66,67], Gyda Bjornsdottir[68], Johannes H. Brandsma[69], Maria Pina Concas[70], Jaime Derringer[71], Nicholas A. Furlotte[23], Tessel E. Galesloot[72], Giorgia Girotto[73], Richa Gupta[74], Leanne M. Hall[75,77], Sarah E. Harris[39,77], Edith Hofer[78,79], Momoko Horikoshi[80,81], Jennifer E. Huffman[34], Kadri Kaasik[82], Ioanna P. Kalafati[83], Robert Karlsson[46], Augustine Kong[68], Jari Lahti[82,84], Sven J. van der Lee[57], Christiaan de Leeuw[5,85], Penelope A. Lind[86], Karl-Oskar Lindgren[47], Tian Liu[87], Massimo Mangino[88,89], Jonathan Marten[34], Evelin Mihailov[11], Michael B. Miller[1], Peter J. van der Most[90], Christopher Oldmeadow[91,92], Antony Payton[93,94], Natalia Pervjakova[11,95], Wouter J. Peyrot[96], Yong Qian[97], Olli Raitakari[98], Rico Rueedi[99,100], Erika Salvi[101], Börge Schmidt[102], Katharina E. Schraut[21], Jianxin Shi[103], Albert V. Smith[58,104], Raymond A. Poot[69], Beate St Pourcain[105,106], Alexander Teumer[107], Gudmar Thorleifsson[68], Niek Verweij[108], Dragana Vuckovic[73], Juergen Wellmann[109], Harm-Jan Westra[110,111,112], Jingyun Yang[113,114], Wei Zhao[115], Zhihong Zhu[52], Behrooz Z. Alizadeh[90,116], Najaf Amin[57], Andrew Bakshi[52], Sebastian E. Baumeister[107,117], Ginevra Biino[118], Klaus Bønnelykke[63], Patricia A. Boyle[113,119], Harry Campbell[21], Francesco P. Cappuccio[120], Gail Davies[77,121], Jan-Emmanuel De Neve[122], Panos Deloukas[123,124], Ilja Demuth[125,126], Jun Ding[97], Peter Eibich[127,128], Lewin Eisele[102], Niina Eklund[95], David M. Evans[105,129], Jessica D. Faul[130], Mary F. Feitosa[131], Andreas J. Forstner[132,133], Ilaria Gandin[73], Bjarni Gunnarsson[68], Bjarni V. Halldórsson[68,134], Tamara B. Harris[135], Andrew C. Heath[136], Lynne J. Hocking[137], Elizabeth G. Holliday[91,92], Georg Homuth[138], Michael A. Horan[139], Jouke-Jan Hottenga[62], Philip L. de Jager[112,140,141], Peter K. Joshi[21,24], Astanand Jugessur[142], Marika A. Kaakinen[143], Mika Kähönen[144,145], Stavroula Kanoni[123], Liisa Keltigangas-Järvinen[82], Lambertus A. L. M. Kiemeney[72], Ivana Kolcic[146], Seppo Koskinen[95], Aldi T. Kraja[131], Martin Kroh[127], Zoltan Kutalik[99,100,147], Antti Latvala[74], Lenore J. Launer[148], Maël P. Lebreton[60,149], Douglas F. Levinson[150], Paul Lichtenstein[46], Peter Lichtner[151], David C. M. Liewald[77,121], LifeLines Cohort Study[152], Anu Loukola[74], Pamela A. Madden[136], Reedik Mägi[11], Tomi Mäki-Opas[95], Riccardo E. Marioni[41,77,153], Pedro Marques-Vidal[154], Gerardus A. Meddens[155], George McMahon[105], Christa Meisinger[67], Thomas Meitinger[151], Yusplitri Milaneschi[96], Lili Milani[11], Grant W. Montgomery[156], Ronny Myhre[142], Christopher P. Nelson[75,76], Dale R. Nyholt[156,157], William E. R. Ollier[93], Aarno Palotie[16,17,112,158,159,160], Lavinia Paternoster[105], Nancy L. Pedersen[46], Katja E. Petrovic[78], David J. Porteous[39], Katri Räikkönen[82,84], Susan M. Ring[105], Antonietta Robino[161], Olga Rostapshova[7,162], Igor Rudan[21], Aldo Rustichini[163], Veikko Salomaa[95], Alan R. Sanders[164,165], Antti-Pekka Sarin[159,166], Helena Schmidt[78,167], Rodney J. Scott[92,168], Blair H. Smith[169], Jennifer A. Smith[90], Jan A. Staessen[170,171], Elisabeth Steinhagen-Thiessen[125], Konstantin Strauch[172,173], Antonio Terracciano[174], Martin D. Tobin[175], Sheila Ulivi[161], Simona Vaccargiu[70], Lydia Quaye[88], Frank J. A. van Rooij[57,176], Cristina Venturini[88,89], Anna A. E. Vinkhuyzen[52], Uwe Völker[138], Henry Völzke[107], Judith M. Vonk[90], Diego Vozzi[161], Johannes Waage[63,64], Erin B. Ware[115,177], Gonneke Willemsen[62], John R. Attia[91,92], David A. Bennett[113,114], Klaus Berger[108], Lars Bertram[178,179], Hans Bisgaard[63], Dorret I. Boomsma[62], Ingrid B. Borecki[131], Ute Bültmann[180], Christopher F. Chabris[50], Francesco Cucca[181], Daniele Cusi[101,182], Ian J.

Deary[77,121], George V. Dedoussis[83], Cornelia M. van Duijn[57], Johan G. Eriksson[84,183], Barbara Franke[184], Lude Franke[185], Paolo Gasparini[73,161,186], Pablo V. Gejman[164,165], Christian Gieger[66], Hans-Jörgen Grabe[187,188], Jacob Gratten[52], Patrick J. F. Groenen[189], Vilmundur Gudnason[58,104], Pim van der Harst[108,185,190], Caroline Hayward[34], David A. Hinds[23], Wolfgang Hoffmann[107], Elina Hyppönen[191,192,193], William G. Iacono[1], Bo Jacobsson[65,142], Marjo-Riitta Järvelin[194,195,196,197], Karl-Heinz Jöckel[102], Jaakko Kaprio[74,95,159], Sharon L. R. Kardia[115], Terho Lehtimäki[198,199], Steven F. Lehrer[43,44,45], Patrik K. E. Magnusson[46], Nicholas G. Martin[200], Matt McGue[1], Andres Metspalu[11,201], Neil Pendleton[202,203], Brenda W. J. H. Penninx[96], Markus Perola[11,95], Nicola Pirastu[73], Mario Pirastu[70], Ozren Polasek[21,204], Danielle Posthuma[5,205], Christine Power[192], Michael A. Province[131], Nilesh J. Samani[75,76], David Schlessinger[97], Reinhold Schmidt[78], Thorkild I. A. Sørensen[14,105,206], Tim D. Spector[88], Kari Stefansson[68,104], Unnur Thorsteinsdottir[68,104], A. Roy Thurik[12,56,207,208], Nicholas J. Timpson[105], Henning Tiemeier[57,209,210], Joyce Y. Tung[23], André G. Uitterlinden[57,211], Veronique Vitart[34], Peter Vollenweider[154], David R. Weir[130], James F. Wilson[21,34], Alan F. Wright[34], Dalton C. Conley[42], Robert F. Krueger[1], George Davey Smith[105], Albert Hofman[57], David I. Laibson[7], Sarah E. Medland[86], Michelle N. Meyer[51], Jian Yang[10,52], Magnus Johannesson[53], Peter M. Visscher[10,52], Tõnu Esko[11], Philipp D. Koellinger[5,6,12], David Cesarini[45,49,55] & Daniel J. Benjamin[9,45,54].

[56] Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3062 PA, Rotterdam, The Netherlands

[57] Department of Epidemiology, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

[58] Icelandic Heart Association, Kopavogur, 201, Iceland

[59] Faculty of Pharmaceutical Sciences, University of Iceland, 107 Reykjavík, Iceland

[60] Amsterdam Business School, University of Amsterdam, Amsterdam, 1018 TV, The Netherlands

[61] New York Genome Center, New York, NY 10013, USA

[62] Department of Biological Psychology, VU University Amsterdam, Amsterdam, 1081 BT, The Netherlands

[63] COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, 2820, Denmark

[64] Steno Diabetes Center, Gentofte, 2820, Denmark

[65] Department of Obstetrics and Gynecology, Institute of Clinical Sciences, Sahlgrenska Academy, Gothenburg, SE 416 85, Sweden

[66] Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany

[67] Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany

[68] deCODE Genetics/Amgen Inc., Reykjavik, IS-101, Iceland

[69] Department of Cell Biology, Erasmus Medical Center Rotterdam, 3015 CN, The Netherlands

[70] Istituto di Ricerca Genetica e Biomedica U.O.S. di Sassari, National Research Council of Italy, Sassari, 07100, Italy

[71] Psychology, University of Illinois, IL 61820, Champaign, USA

[72] Institute for Computing and Information Sciences, Radboud University Nijmegen, Nijmegen, 6525 EC, The Netherlands
[73] Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, 34100, Italy
[74] Department of Public Health, University of Helsinki, Helsinki, FI-00014, Finland
[75] Department of Cardiovascular Sciences, University of Leicester, Leicester, LE3 9QP, UK
[76] NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, LE3 9QP, UK
[77] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
[78] Department of Neurology, General Hospital and Medical University Graz, Graz, 8036, Austria
[79] Institute for Medical Informatics, Statistics and Documentation, General Hospital and Medical University Graz, Graz, 8036, Austria
[80] Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Oxford, OX3 7LE, UK
[81] Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK
[82] Institute of Behavioural Sciences, University of Helsinki, Helsinki, FI-00014, Finland
[83] Nutrition and Dietetics, Health Science and Education, Harokopio University, Athens, 17671, Greece
[84] Folkhälsan Research Centre, Helsingfors, FI-00014, Finland
[85] Institute for Computing and Information Sciences, Radboud University Nijmegen, Nijmegen, 6525 EC, The Netherlands
[86] Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia
[87] Lifespan Psychology, Max Planck Institute for Human Development, Berlin, 14195, Germany
[88] Department of Twin Research and Genetic Epidemiology, King's College London, London, SE1 7EH, UK
[89] NIHR Biomedical Research Centre, Guy’s and St. Thomas’ Foundation Trust, London, SE1 7EH, UK
[90] Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, 9700 RB, The Netherlands
[91] Public Health Stream, Hunter Medical Research Institute, New Lambton, NSW 2305, Australia
[92] Faculty of Health and Medicine, University of Newcastle, Newcastle, NSW 2300, Australia
[93] Centre for Integrated Genomic Medical Research, Institute of Population Health, The University of Manchester, Manchester, M13 9PT, UK
[94] School of Psychological Sciences, The University of Manchester, Manchester, M13 9PL, UK
[95] Department of Health, THL-National Institute for Health and Welfare, Helsinki, FI-00271, Finland
[96] Psychiatry, VU University Medical Center & GGZ inGeest, Amsterdam, 1081 HL, The Netherlands
[97] Laboratory of Genetics, National Institute on Aging, Baltimore, MD 21224, USA
[98] Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, 20521, Finland
[99] Department of Medical Genetics, University of Lausanne, Lausanne, 1005, Switzerland
[100] Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland

[101] Department Of Health Sciences, University of Milan, Milano, 20142, Italy
[102] Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, Essen, 45147, Germany
[103] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892-9780, USA
[104] Faculty of Medicine, University of Iceland, Reykjavik, 101, Iceland
[105] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK
[106] School of Oral and Dental Sciences, University of Bristol, Bristol, BS1 2LY, UK
[107] Institute for Community Medicine, University Medicine Greifswald, Greifswald, 17475, Germany
[108] Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, 9700 RB, The Netherlands 107
[109] Institute of Epidemiology and Social Medicine, University of Muenster, Muenster, 48149, Germany
[110] Divisions of Genetics and Rheumatology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, MA 02115, Boston, USA
[111] Partners Center for Personalized Genetic Medicine, Boston, MA 02115, USA
[112] Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[113] Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL 60612, USA
[114] Department of Neurological Sciences, Rush University Medical Center, Chicago, IL 60612, USA
[115] Department of Epidemiology, University of Michigan, Ann Arbor, MI 48109, USA
[116] Department of Gastroenterology and Hepatology, University of Groningen, University Medical Center Groningen, Groningen, 9713 GZ, The Netherlands
[117] Institute of Epidemiology and Preventive Medicine, University of Regensburg, Regensburg, D-93053, Germany
[118] Institute of Molecular Genetics, National Research Council of Italy, Pavia, 27100, Italy
[119] Department of Behavioral Sciences, Rush University Medical Center, Chicago, IL 60612, USA
[120] Warwick Medical School, University of Warwick, Coventry, CV4 7AL, UK
[121] Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
[122] Saïd Business School, University of Oxford, Oxford, OX1 1HP, UK
[123] William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK
[124] Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, 21589, Saudi Arabia
[125] The Berlin Aging Study II; Research Group on Geriatrics, Charité – Universitätsmedizin Berlin, Germany, Berlin, 13347, Germany
[126] Institute of Medical and Human Genetics, Charité-Universitätsmedizin, Berlin, Berlin, 13353, Germany
[127] German Socio- Economic Panel Study, DIW Berlin, Berlin, 10117, Germany
[128] Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, OX3 7LF, UK
[129] The University of Queensland Diamantina Institute, The Translational Research Institute, Brisbane, QLD 4102, Australia

[130] Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48109, USA
[131] Department of Genetics, Division of Statistical Genomics, Washington University School of Medicine, St. Louis, MO 63018, USA
[132] Institute of Human Genetics, University of Bonn, Bonn, 53127, Germany
[133] Department of Genomics, Life and Brain Center, University of Bonn, Bonn, 53127, Germany
[134] Institute of Biomedical and Neural Engineering, School of Science and Engineering, Reykjavik University, Reykjavik 101, Iceland
[135] Laboratory of Epidemiology, Demography, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892-9205, United States
[136] Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA
[137] Division of Applied Health Sciences, University of Aberdeen, Aberdeen, AB25 2ZD, UK
[138] Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, 17475, Germany
[139] Manchester Medical School, The University of Manchester, Manchester, 9PT, UK
[140] Program in Translational NeuroPsychiatric Genomics, Departments of Neurology & Psychiatry, Brigham and Women's Hospital, Boston, MA 02115, USA
[141] Harvard Medical School, Boston, MA 02115, USA
[142] Department of Genes and Environment, Norwegian Institute of Public Health, Oslo, N-0403, Norway
[143] Department of Genomics of Common Disease, Imperial College London, London, W12 0NN, UK
[144] Department of Clinical Physiology, Tampere University Hospital, Tampere, 33521, Finland
[145] Department of Clinical Physiology, University of Tampere, School of Medicine, Tampere, 33014, Finland
[146] Public Health, Medical School, University of Split, 21000 Split, Croatia
[147] Institute of Social and Preventive Medicine, Lausanne University Hospital (CHUV), Lausanne, 1010, Switzerland
[148] Neuroepidemiology Section, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892-9205, USA
[149] Amsterdam Brain and Cognition Center, University of Amsterdam, 1018 XA, Amsterdam, The Netherlands
[150] Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA 94305-5797, USA
[151] Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany 155
[152] LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, Groningen, 9713 BZ, The Netherlands
[153] Medical Genetics Section, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK
[154] Department of Internal Medicine, Internal Medicine, Lausanne University Hospital (CHUV), Lausanne, 1011, Switzerland
[155] Tema BV, 2131 HE Hoofddorp, The Netherlands
[156] Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia

[157] Institute of Health and Biomedical Innovation, Queensland Institute of Technology, Brisbane, QLD 4059, Australia
[158] Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114, USA
[159] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, 00014, Finland
[160] Department of Neurology, Massachusetts General Hospital, Boston, MA 02114, USA
[161] Medical Genetics, Institute for Maternal and Child Health IRCCS "Burlo Garofolo", Trieste, 34100, Italy
[162] Social Impact, Arlington, VA 22201, USA
[163] Department of Economics, University of Minnesota Twin Cities, Minneapolis, MN 55455, USA
[164] Department of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem, Evanston, IL 60201-3137, USA
[165] Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL 60637, USA
[166] Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki 00300, Finland
[167] Research Unit for Genetic Epidemiology, Institute of Molecular Biology and Biochemistry, Center of Molecular Medicine, General Hospital and Medical University, Graz, Graz, 8010, Austria
[168] Information Based Medicine Stream, Hunter Medical Research Institute, New Lambton, NSW 2305, Australia
[169] Medical Research Institute, University of Dundee, Dundee, DD1 9SY, UK
[170] Research Unit Hypertension and Cardiovascular Epidemiology, Department of Cardiovascular Science, University of Leuven, Leuven, 3000, Belgium
[171] R&D VitaK Group, Maastricht University, Maastricht, 6229 EV, The Netherlands
[172] Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany
[173] Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig Maximilians-Universität, Munich, 81377, Germany
[174] Department of Geriatrics, Florida State University College of Medicine, Tallahassee, FL 32306, USA
[175] Department of Health Sciences and Genetics, University of Leicester, Leicester, LE1 7RH, UK
[176] Department of Internal Medicine, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
[177] Research Center for Group Dynamics, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104, USA
[178] Platform for Genome Analytics, Institutes of Neurogenetics & Integrative and Experimental Genomics, University of Lübeck, Lübeck, 23562, Germany
[179] Neuroepidemiology and Ageing Research Unit, School of Public Health, Faculty of Medicine, The Imperial College of Science, Technology and Medicine, London SW7 2AZ, UK
[180] Department of Health Sciences, Community & Occupational Medicine, University of Groningen, University Medical Center Groningen, Groningen, 9713 AV, The Netherlands

[181] Istituto di Ricerca Genetica e Biomedica (IRGB), Consiglio Nazionale delle Ricerche, c/o Cittadella Universitaria di Monserrato, Monserrato, Cagliari, 9042, Italy
[182] Institute of Biomedical Technologies, Italian National Research Council, Segrate (Milano), 20090, Italy
[183] Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, 00014, Finland
[184] Departments of Human Genetics and Psychiatry, Donders Centre for Neuroscience, Nijmegen, 6500 HB, The Netherlands
[185] Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen. 9700 RB, The Netherlands
[186] Sidra, Experimental Genetics Division, Sidra, Doha 26999, Qatar
[187] Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, 17475, Germany
[188] Department of Psychiatry and Psychotherapy, HELIOS-Hospital Stralsund, Stralsund, 18437, Germany
[189] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, 3062 PA, The Netherlands
[190] Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, 1105 AZ, The Netherlands
[191] Centre for Population Health Research, School of Health Sciences and Sansom Institute, University of South Australia, SA5000, Adelaide, Australia
[192] South Australian Health and Medical Research Institute, Adelaide, SA5000, Australia
[193] Population, Policy and Practice, UCL Institute of Child Health, London, WC1N 1EH, UK
[194] Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment & Health, School of Public Health, Imperial College London, London, W2 1PG, UK
[195] Center for Life Course Epidemiology, Faculty of Medicine, University of Oulu, Oulu, FI-90014 , Finland
[196] Unit of Primary Care, Oulu University Hospital, Oulu, 90029 OYS, Finland
[197] Biocenter Oulu, University of Oulu, FI-90014 Oulu, Finland
[198] Fimlab Laboratories, Tampere, 33520, Finland
[199] Department of Clinical Chemistry, University of Tampere, School of Medicine, Tampere, 33014, Finland
[200] Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia
[201] Institute of Molecular and Cell Biology, University of Tartu, Tartu, 51010, Estonia
[202] Centre for Clinical and Cognitive Neuroscience, Institute Brain Behaviour and Mental Health, Salford Royal Hospital, Manchester, M6 8HD, UK
[203] Manchester Institute Collaborative Research in Ageing, University of Manchester, Manchester, M13 9PL, UK
[204] Faculty of Medicine, University of Split, Croatia, Split 21000, Croatia
[205] Department of Clinical Genetics, VU Medical Centre, Amsterdam, 1081 HV, The Netherlands
[206] Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospitals, The Capital Region, Frederiksberg, 2000, Denmark
[207] Montpellier Business School, Montpellier, 34080, France
[208] Panteia, Zoetermeer, 2715 CA, The Netherlands

[209] Department of Psychiatry, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
[210] Department of Child and Adolescent Psychiatry, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
[211] Department of Internal Medicine, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

AUTHOR CONTRIBUTIONS: D.J.B., D.C., P.T., and P.M.V. designed and oversaw the study. A.O. was the study's lead analyst, responsible for quality control and meta-analyses. Analysts who assisted A.O. in major ways include: E.K. (quality control), O.M. (COJO, MTAG, quality-control), T.A.N-V. (figure preparation), H.L. (quality control), C.L. (quality control), J.S. (UKB association analyses), and R.K.L. (UKB association analyses). P.B. and E.K. conducted the within-family association analyses. The cross-cohort heritability and genetic-correlation analyses were conducted by R.W. and M.Z. The analyses of the X chromosome in UK Biobank were conducted by J.S.; A.O. ran the meta-analysis. J.J.L. organized and oversaw the bioinformatics analyses, with assistance from T.E., E.K., K.T., T.H.P., and P.N.T. Polygenic-prediction analyses were designed and conducted by A.O., K.T., and R.W. Besides the contributions explicitly listed above, T.K., R.L., and R.R. conducted additional analyses for several subsections. C.W. helped with coordinating among the participating cohorts. J.P.B., D.C.C., T.E., M.J., J.J.L., P.D.K., D.I.L., S.F.L., S.O., M.R.R., K.T., and J.Y. provided helpful advice and feedback on various aspects of the study design. All authors contributed to and critically reviewed the manuscript. E.K., J.J.L., and R.W. made especially major contributions to the writing and editing.

COMPETING FINANCIAL INTERESTS: Anil Malhotra is a consultant to Genomind Inc., Informed DNA, Concert Pharmaceuticals, and Biogen. Nicholas A. Furlotte, Aaron Kleinman, and Joyce Tung are employees of 23andMe, Inc.

REFERENCES

1. Branigan, A. R. *et al.* Variation in the Heritability of Educational Attainment: An International Meta-Analysis. *Soc. Forces* **92,** 109–140 (2013).
2. Conti, G., Heckman, J. & Urzua, S. The Education-Health Gradient. *Am. Econ. Rev.* **100,** 234–238 (2010).
3. Cutler, D. M. & Lleras-Muney, A. Education and Health: Evaluating Theories and Evidence. in *Making Americans Healthier: Social and Economic Policy as Health Policy* (eds. House, J., Schoeni, R., Kaplan, G. & Pollack, H.) (Russell Sage Foundation, 2008).
4. Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science (80-.).* **340,** 1467–1471 (2013).
5. Pickrell, J. K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48,** 709–717 (2016).
6. Belsky, D. W. *et al.* The Genetics of Success. *Psychol. Sci.* **27,** 957–972 (2016).

7. Domingue, B. W., Belsky, D. W., Conley, D., Harris, K. M. & Boardman, J. D. Polygenic Influence on Educational Attainment: New evidence from The National Longitudinal Study of Adolescent to Adult Health. *AERA Open* **1,** 1–13 (2015).

8. Marioni, R. E. *et al.* Genetic variants linked to education predict longevity. *Proc. Natl. Acad. Sci.* **113,** 13366–13371 (2016).

9. Anttila, A. V *et al.* Analysis of shared heritability in common disorders of the brain. *bioRxiv* 48991 (2016). doi:http://dx.doi.org/10.1101/048991

10. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533,** 539–542 (2016).

11. Turley, P. *et al.* MTAG: Multi-Trait Analysis of GWAS. *Nat. Genet.* **in press,** (2017).

12. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526,** 68–74 (2015).

13. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).

14. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44,** 369–375 (2012).

15. Kong, A. *et al.* The nature of nurture: effects of parental genotypes. *bioRxiv* 219261 (2017). doi:10.1101/219261

16. de Vlaming, R. *et al.* Meta-GWAS Accuracy and Power (MetaGAP) Calculator Shows that Hiding Heritability Is Partially Due to Imperfect Genetic Correlations across Studies. *PLOS Genet.* **13,** e1006495 (2017).

17. Tropf, F. C. *et al.* Hidden heritability due to heterogeneity across seven populations. *Nat. Hum. Behav.* (2017). doi:10.1038/s41562-017-0195-1

18. Johnson, W., Carothers, A. & Deary, I. J. Sex Differences in Variability in General Intelligence: A New Look at the Old Question. *Perspect. Psychol. Sci.* **3,** 518–531 (2008).

19. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6,** 5890 (2015).

20. Azevedo, F. A. C. *et al.* Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* **513,** 532–541 (2009).

21. Finucane, H. K. *et al.* Partitioning heritability by functional category using GWAS summary statistics. *Nat. Genet.* **47,** 1228–1235 (2015).

22. Reed, T. E. & Jensen, A. R. Arm nerve conduction velocity (NCV), brain NCV, reaction time, and intelligence. *Intelligence* **15,** 33–47 (1991).

23. Chen, W., McDonnell, S. K., Thibodeau, S. N., Tillmans, L. S. & Schaid, D. J. Incorporating functional annotations for fine-mapping causal variants in a Bayesian framework using summary statistics. *Genetics* **204,** 933–958 (2016).

24. Wang, G. *et al.* CaV3.2 calcium channels control NMDA receptor-mediated transmission: a new mechanism for absence epilepsy. *Genes Dev.* **29,** 1535–51 (2015).

25. Vilhjálmsson, B. J. *et al.* Modeling linkage disequilibrium increases accuracy of polygenicrisk scores. *Am. J. Hum. Genet.* **97,** 576–592 (2015).

26. Martin, A. R. *et al.* Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **100,** 635–649 (2017).

27. Scutari, M., Mackay, I. & Balding, D. Using Genetic Distance to Infer the Accuracy of Genomic Prediction. *PLoS Genet.* **12,** e1006288 (2016).

28. Trampush, J. W. *et al.* GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive

function: a report from the COGENT consortium. *Mol. Psychiatry* **22,** 336–345 (2017).

29. Davies, G. *et al.* Ninety-nine independent genetic loci influencing general cognitive function include genes associated with brain health and structure (N = 280,360). *bioRxiv* (2017).

30. Sniekers, S. *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat Genet* **49,** 1107–1112 (2017).

31. Savage, J. E. *et al.* GWAS meta-analysis (N=279,930) identifies new genes and functional links to intelligence. *bioRxiv* (2017).

32. Schmitz, L. L. & Conley, D. The Effect of Vietnam-Era Conscription and Genetic Potential for Educational Attainment on Schooling Outcomes. *Econ. Educ. Rev.* (2017). doi:10.1016/j.econedurev.2017.10.001

33. Heath, A. C. *et al.* Education policy and the heritability of educational attainment. *Nature* **314,** 734–736 (1985).

34. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478,** 483–489 (2011).

Fig. 1. Manhattan Plot for GWAS of *EduYears* (*N* = 1,131,881). P values and the mean χ^2 shown in figure are based on inflation-adjusted test statistics. The x-axis is chromosomal position, and the y-axis is the significance on a $-\log_{10}$ scale. The dashed line marks the threshold for genome-wide significance ($P = 5\times10^{-8}$).

Fig 2. Sign Concordance in Within-Family Association Analyses. The set of LD-pruned SNPs is limited to SNPs with **(a)** $P < 5\times10^{-3}$, **(b)** $P < 5\times10^{-5}$, or **(c)** $P < 5\times10^{-8}$. Each panel compares the observed sign concordance between within-family and GWAS estimates to the distributions expected (i) by chance alone (pink); (ii) according to a Bayesian framework that adjusts the GWAS estimates for bias due to winner's curse (green); and (iii) according to the same framework with an additional adjustment for bias due to assortative mating (blue). These results are based on a GWAS sample size of 1,070,751 individuals and a within-family sample of 22,135 sibling pairs (44,270 individuals).

Fig. 3. Tissue-specific expression of genes in DEPICT-defined loci. (a) We took microarray measurements from the Gene Expression Omnibus[19] and determined whether the genes overlapping *EduYears*-associated loci (as defined by DEPICT) are significantly overexpressed (relative to genes in random sets of loci) in each of 180 tissues/cell types. These types are grouped in the figure by Medical Subject Headings (MeSH) first-level term. The y-axis is the one-sided P value from DEPICT on a $-\log_{10}$ scale. The 28 dark bars correspond to tissues/cell types in which the genes are significantly overexpressed (FDR < 0.01), including all 22 classified as part of the central nervous system (see **Supplementary Table 6** for identifiers of all tissues/cell types). **(b)** Whereas genes prioritized by DEPICT in a previous analysis based on a smaller sample[10] tend to be more strongly expressed in the brain prenatally (red curve), the 1,703 newly prioritized genes show a flat trajectory of expression across development (blue curve). Both groups of DEPICT-prioritized genes show elevated levels of expression relative to protein-coding genes that are not prioritized (gray curve). Analyses were based on RNA-seq data from the BrainSpan Developmental Transcriptome[34]. These results are based on the full GWAS sample of 1,131,881 individuals. Error bars represents 95% confidence intervals.

Fig. 4. Prediction Accuracy. (a) Mean prevalence of college completion by *EduYears* PGS quintile. Error bars show the 95% confidence interval for the mean. **(b)** Incremental R^2 of the *EduYears* PGS compared to that of other variables. (**c**) Incremental R^2 of the PGS for *EduYears* and *Cognitive Performance* constructed from the respective GWAS or MTAG summary statistics. Error bars for the R^2 values show bootstrapped 95% confidence intervals with 1000 iterations each. Sample sizes are N = 4,775 for Add Health and N = 8,609 for HRS.

ONLINE METHODS

This article is accompanied by a **Supplementary Note** with further details.

Genome-wide association study meta-analyses. Our primary analysis extends the (combined discovery and replication) sample of a previous genome-wide association study (GWAS) of educational attainment[10] from N = 405,072 to N = 1,131,881 individuals. We performed a sample-size-weighted meta-analysis of 71 quality-controlled cohort-level results files using the METAL software[35]. The meta-analysis combines 59 cohort-level results files from the previous study with 12 new results files: 8 from cohorts that were not included in the previous study[10] and 4 from cohorts that updated their results in larger samples.

All cohort-level analyses were restricted to European-ancestry individuals that passed the cohort's quality control and whose *EduYears* was measured at an age of at least 30. The *EduYears* phenotype was constructed by mapping each major educational qualification that can be identified from the cohort's survey measure to an International Standard Classification of Education (ISCED) category and imputing a years-of-education equivalent for each ISCED category. Details on cohort-level phenotype measures, genotyping, imputation, association analyses, and quality-control filters are described in **Supplementary Tables 16-19**.

We used the estimated intercept from LD Score regression[13] to inflation-adjust the test statistics. We then used the clumping algorithm described below to determine the number of approximately independent SNPs identified at any given P value threshold.

Clumping algorithm. Our clumping algorithm is iterative and has been used previously[10]. We describe it here for the case of identifying lead SNPs among the set of SNPs reaching $P < 5\times10^{-8}$; the algorithm is the same when determining sets of approximately independent SNPs for other P value thresholds.

First, the SNP with the smallest P value in the pooled meta-analysis results is identified as the lead SNP of the first clump. Next, all SNPs in LD with the lead SNP are also assigned to this clump. SNPs are defined to be in LD with each other if they are on the same chromosome and the squared correlation of their genotypes is $r^2 > 0.1$. To determine the second lead SNP and second clump, the first clump is removed, and the same steps are applied to the remaining SNPs. The process is repeated until no SNPs with P value below 5×10^{-8} remain. Each locus is defined by a lead SNP and the SNPs assigned to its clump. Hence, each lead SNP maps to exactly one locus, and each locus maps to exactly one lead SNP.

We perform the clumping in Plink[36]. Note that we measure the LD between every pair of SNPs on each chromosome without regard to the physical distance between them. Therefore, if two SNPs on the same chromosome have pairwise r^2 above 0.1, then they cannot both be lead SNPs. On the other hand, it is possible for two SNPs in close physical proximity both to be lead SNPs, provided their pairwise r^2 is below 0.1. The **Supplementary Note** reports analyses of the sensitivity of the number of lead SNPs and loci to alternative definitions and to the choice of the reference file used to estimate LD.

Conditional and joint multiple-SNP analysis (COJO). Given a P value threshold specified by the user, COJO[14] is a method that identifies a set of SNPs such that, in a multivariate regression of the phenotype on all the SNPs in the set, every SNP has a P value below threshold. COJO uses the meta-analysis summary statistics together with LD estimates from a reference simple. Our COJO analysis was conducted using a reference sample of approximately unrelated individuals of European ancestry from UK Biobank. We specified the P value threshold 5×10^{-8}. The analyses were restricted to SNPs satisfying recommended quality-control filters. The **Supplementary Note** contains additional details.

Bayesian framework for calculating winner's-curse-adjusted posterior effect-size distributions. We assume that the marginal effect size of each SNP is drawn from the following mixture distribution:

$$\beta_j \sim \begin{cases} N(0, \tau^2) & \text{with probability } \pi \\ 0 & \text{otherwise,} \end{cases}$$

where τ^2 is the effect-size variance for non-null SNPs and π is the fraction of non-null SNPs in our data. We estimate the parameters τ^2 and π by maximum likelihood. Given their values, the posterior distribution of SNP j can be calculated from Bayes' Rule. Relative to the GWAS effect estimate, the mean of the posterior distribution is shrunken toward zero (because zero is the mean of the prior distribution) and is not biased by the winner's curse. Further details and a derivation of the likelihood function used in the maximum-likelihood estimation are provided on p. 59 in the Supplementary Note of a previous SSGAC study[37].

To calculate the 5th, 50th, and 95th percentile of the effect-size distribution of our lead SNPs, we simulated effect sizes from each lead SNP's posterior distribution and identified the 5th, 50th, and 95th percentiles of the complete set of simulated effect sizes.

As described below, we also use this Bayesian framework in our GWAS and MTAG replication analyses and in our within-family analyses.

Replication of lead SNPs from Okbay et al.'s combined-stage analysis. We conducted a replication analysis of the 162 lead SNPs identified at genome-wide significance in Okbay et al.'s[10] pooled (discovery and replication) meta-analysis (N = 405,073). Of the 162 SNPs, 158 pass quality-control filters in our updated meta-analysis. To examine their out-of-sample replicability, we calculated Z-statistics from the subsample of our data (N = 726,808) that was not included in Okbay et al. Let the Z-statistics of association from, respectively, Okbay et al., the new data, and our final EA3 meta-analysis, be denoted by Z_1, Z_2 and Z. Since our meta-analysis used sample-size weighting[35], Z_2 is implicitly defined by:

$$Z = \sqrt{\frac{N_1}{N}} Z_1 + \sqrt{\frac{N_2}{N}} Z_2,$$

where SNP subscripts have been dropped and N's are sample sizes. Because this formula holds when Z_1 and Z_2 are independent, the implicitly-defined Z_2 is interpreted as the additional information contained in the new data.

Of the 158 SNPs, we found that 154 have matching signs in the new data (for the remaining four SNPs, the estimated effect is never statistically distinguishable from zero at $P < 0.10$). Of the 154 SNPs with matching signs, 143 are significant at $P < 0.01$, 119 are significant at $P < 10^{-5}$, and 97 are significant at $P < 5{\times}10^{-8}$. The replication results are shown graphically in **Supplementary Figure 3.** To help interpret these results, we used the Bayesian framework described above to calculate the expected replication record under the hypothesis that all 158 SNPs are true associations. The posterior distributions of the SNPs' effect sizes are calculated using parameters estimated from Okbay et al.'s summary statistics: $(\hat{\tau}^2, \hat{\pi}) = (5.02 \times 10^{-6}, 0.33)$.

Within-family analyses. We conducted within-family association analyses on a sample of 22,135 sibling pairs from *STR-Twingene, STR-SALTY*, *UKB*, and *WLS*. For each cohort, we standardized *EduYears* within the cohort and then residualized this variable using the same controls as in the GWAS. We then regressed the sibling difference in the residuals on the sibling difference in genotype. We restricted analyses to SNPs with minor allele frequency above 5% in each of the sibling cohorts and meta-analyzed the cohort-level results using inverse-variance weighting.

We followed Okbay et al.[37] to compare the signs of the within-family estimates to the signs of the estimates from a GWAS meta-analysis that we re-ran after removing the sibling samples (N = 1,070,751). We benchmarked our observed fraction of concordant signs against the three theoretical benchmarks shown in **Fig. 2**. The theoretical benchmarks are calculated using posterior distributions for the GWAS effect sizes obtained from our Bayesian statistical framework. Treating each benchmark as a null hypothesis, we conducted one-sided binomial tests where the alternative hypothesis is that the observed sign concordance falls short of the benchmark. We conducted this test for sets of approximately independent SNPs selected at the P value thresholds 5×10^{-8}, 5×10^{-5}, and 5×10^{-3} (**Supplementary Table 20** and **Fig. 2**).

We also performed regression-based comparisons of the within-family estimates and the GWAS estimates (**Supplementary Table 21** and **Supplementary Figure 21**). Further details, including a derivation of our assortative-mating adjustment, can be found in the **Supplementary Note.**

Joint F-test of heterogeneity. When the SNPs are considered individually, for all but one of the 1,271 lead SNPs, we fail to reject a null hypothesis of homogenous effects across cohorts at the Bonferroni-adjusted P value threshold of 0.05/1,271. We generated an omnibus test statistic for heterogeneity by summing the Cochran Q-statistics for heterogeneity across all 1,271 lead SNPs[38]. Because the software used for meta-analysis does not report Q-statistics, we inferred these values based on the reported heterogeneity P values. To do so, we treated each lead SNP as if it were available for each of the 71 cohorts in the meta-analysis, which implies that the Q-statistic for each lead SNP has a χ^2 distribution with 70 degrees of freedom. The sum of these Q-statistics is therefore (approximately) χ^2-distributed with $70 \times 1{,}271 = 88{,}970$ degrees of freedom. This gave us an omnibus Q-statistic of 91,830, with corresponding P value equal to 9.68×10^{-12}.

Cross-cohort genetic correlation. We estimated the genetic correlation of *EduYears* across all pairs of cohorts with non-negative heritability estimates (**Supplementary Table 22**). We used bivariate LD Score regression[39] implemented by the LDSC software with a European reference population, filtered to HapMap3 SNPs. The estimated genetic correlations of *EduYears* between each of our 933 pairs of cohorts is shown in **Supplementary Table 23**.

We calculated the inverse-variance-weighted mean of the genetic-correlation estimates. The genetic correlation across pairs of cohorts will be correlated across all observations that share one of their cohorts in common. Therefore, to obtain correct standard errors, we used the node-jackknife variance estimator described by Cameron and Miller[40]. As detailed in **Supplementary Note**, we also estimated the variance of SNP heritability of *EduYears* across cohorts, and we conducted analyses to assess the extent to which we can predict variation in SNP heritability and genetic correlation of *EduYears* based on several observable cohort characteristics (**Supplementary Tables 24** and **25**).

X chromosome. We performed association analyses of SNPs on the X chromosome in our two largest cohorts, *UKB* (N = 329,358) and *23andMe* (N = 365,536). The *UKB* analyses were conducted in a sample of conventionally unrelated European-ancestry individuals, yielding a smaller sample size than the autosomal *UKB* analyses (**Supplementary Table 26**). Imputed genotypes for the X chromosome were not included in the data officially released by *UKB*. We therefore imputed the data ourselves using the 1000 Genomes Project[41] as our reference panel.

In both cohorts, the association analyses were performed on a pooled male-female sample with male genotypes coded 0/2. Except for this allele coding in males, all major aspects of the *23andMe* analysis were identical to those described for the autosomal analyses; see **Supplementary Tables 17-19** for details.

Both sets of association results underwent the same set of quality-control filters as the autosomal analyses prior to meta-analysis. Additionally, we dropped a small number of SNPs with male-female allele frequency differences above 0.005 in *UKB*. The meta-analysis was conducted in METAL[35], using sample-size weighting. Only SNPs that were present in both cohorts' results files were used. To adjust the test statistics for bias, we inflated the standard errors using the LD Score regression intercept estimated from our main autosomal analysis ($\sqrt{1.113}$).

Heritability of the X chromosome and dosage compensation. To estimate SNP heritability for males and females, we use the equation

$$\mathrm{E}[\chi_i^2] = 1 + \frac{N_i h_i^2}{M_{\mathrm{eff}}},$$

where $i \in \{m, f\}$ indicates males or females, $\mathrm{E}[\chi_i^2]$ is the expected χ^2 statistic, h_i^2 is the SNP heritability for the X chromosome, N_i is the GWAS sample size, and M_{eff} is the effective number of SNPs (which is assumed to be the same in males and females). We replaced $\mathrm{E}[\chi_i^2]$ with its sample analog and M_{eff} with its estimated value, and then we solved for h_i^2.

Let $\gamma = h_m^2/h_f^2$ denote the dosage compensation ratio. The ratio takes on a value between 0.5 (zero dosage compensation) and 2 (full dosage compensation). Based on the above equation, we estimated it as

$$\hat{\gamma} = \frac{(\hat{\chi}_m^2 - 1)N_f}{(\hat{\chi}_f^2 - 1)N_m},$$

where $\hat{\chi}_i^2$ is the mean χ^2 statistic. (Equivalently, our γ estimate is equal to the ratio of our SNP heritability estimates.)

Biological annotation. We used DEPICT[19] (downloaded February 2016 from https://github.com/perslab/depict) to identify the tissues/cell types where the causal genes are strongly expressed, detect enrichment of gene sets, and prioritize likely causal genes. We ran DEPICT as described previously[10] with the following exceptions: we used 37,427 human Affymetrix HGU133a2.0 platform microarrays[19], discarded gene sets that were not well reconstituted[42], and relaxed the significance threshold for defining a matching SNP in the simulated null GWAS from 5×10^{-4} to 5×10^{-3}. "Previously prioritized" genes were prioritized by DEPICT (in the sense of achieving FDR < 0.05) both in Okbay et al.[10] and in the current work; "newly prioritized genes," on the other hand, were not prioritized in Okbay et al.[10]. We used expression data from the BrainSpan Developmental Transcriptome[34] and calculated the average expression in the brain of all DEPICT-prioritized *EduYears* genes (**Supplementary Table 7**) as a function of developmental stage (**Supplementary Table 8**, **Supplementary Figure 22**).

In addition to the analyses presented in the main text, we determined which functional systems are *least* implicated by DEPICT (**Supplementary Table 27**) and how enrichment of gene sets differs across phenotypes (**Supplementary Table 28**).

We tested the robustness of our DEPICT results using the bioinformatics tools MAGMA[43] and PANTHER[44,45]. For MAGMA, we used the "multi=snp-wise" option, mapping a SNP to a gene if it resides within the gene boundaries or 5kb of either endpoint. We estimated LD using a reference panel of Europeans in 1000 Genomes phase 3, and we defined a gene as significant if its joint *P* value falls below the threshold corresponding to FDR < 0.05 (**Supplementary Table 29**). For PANTHER, we used the binomial overrepresentation test with the DEPICT-prioritized genes as input (**Supplementary Table 30**).

We also used stratified LD Score regression[21] to partition the heritability of the trait between SNPs of different types. In addition to the baseline SNP-level annotations (**Supplementary Table 31**), we tested a number of novel annotation types, described more fully in the **Supplementary Note.** We tested the heritability enrichment of neural cell types (**Supplementary Table 9**), various SNP-level annotations assembled by Pickrell[46] (**Supplementary Figure 23**, **Supplementary Table 32**), developmental stages (**Supplementary Table 33**), and genes that are broadly expressed or specifically expressed in a particular tissue (**Supplementary Figure 24**, **Supplementary Table 34**). We also applied LD Score regression to DEPICT-reconstituted gene sets (**Supplementary Table 35**) and binary gene sets (**Supplementary Table 36** and **Supplementary Figure 25**).

We used the tool CAVIARBF[23,47] in a fine-mapping exercise to identify candidate causal SNPs. We used the 74 baseline annotations employed by stratified LD Score regression as well as 451 annotations from from Pickrell[46]. We applied a MAF filter of 0.01 and a sample-size filter of 400,000 and only considered SNPs within a 50-kb radius of a lead SNP. We computed exact Bayes factors by averaging over prior variances of 0.01, 0.1, and 0.5; we set the sample size to the mean sample size of our considered SNPs; and we added 0.2 to the main diagonal of the LD matrix because we used a reference panel for LD estimation. To incorporate annotations, we used the elastic net setting with parameters selected via 5-fold cross-validation. The resulting annotation effect sizes and list of candidate causal SNPs are given in **Supplementary Tables 37** and **10**. Regional association plots of four noteworthy candidates are shown in **Supplementary Figure 9**.

Polygenic prediction. Prediction analyses were performed using the National Longitudinal Study of Adolescent to Adult Health (Add Health), the Health and Retirement Study (HRS), and the Wisconsin Longitudinal Study (WLS). Polygenic scores were constructed using HapMap3 SNPs that meet the following conditions: (i) the variant has a call rate greater than 98% in the prediction cohort; (ii) the variant has a minor allele frequency (MAF) greater than 1% in the prediction cohort; and (iii) the allele frequency discrepancy between the meta-analysis and the prediction cohort does not exceed 0.15. To calculate the SNP weights we use the software package LDpred[25], assuming a fraction of causal variants equal to 1, and then we construct the scores in PLINK.

All prediction exercises were performed with an OLS or probit regression of a phenotype on our score and a set of controls consisting of a full set of dummy variables for year of birth, an indicator variable for sex, a full set of interactions between sex and year of birth, and the first 10 principal components of the variance-covariance matrix of the genetic relatedness matrix.

Our measure of prediction accuracy is the incremental R^2. To calculate this value, we first regress a phenotype on our set of controls without the polygenic score. Next, we re-run the same regression but with the score included as a regressor. For quantitative phenotypes, our measure of predictive power is the change in R^2. For binary outcomes, we calculated the incremental pseudo-R^2 from a Probit regression. To obtain 95% confidence intervals, we bootstrapped the incremental R^2's with 1000 repetitions (**Supplementary Table 38** and **Supplementary Figures 13, 26, 27 and 28.**

Prediction of other phenotypes. In addition to *EduYears*, we also used our polygenic score to predict a number of other phenotypes. In the HRS and Add Health, we analyzed three binary variables related to educational attainment: (i) High School Completion, (ii) College Completion, and (iii) Grade Retention (i.e., retaking a grade).

In additional analyses in Add Health, we predicted an augmented version of the Peabody Picture Vocabulary test, measured when participants were 12–20 years old. Peabody scores were age-standardized. We also predicted a number of Grade Point Average variables (range: 0.0 to 4.0) from the third wave of Add Health, when transcripts were collected from respondents' high schools. We analyzed Overall GPA, Math GPA, Science GPA, and Verbal GPA, controlling for high school fixed effects.

In additional analyses in the HRS, we predicted several cognitive phenotypes. Total Cognition is the sum of four cognitive measures measured in waves 3 through 10: an immediate word recall task, a delayed word recall task, a naming task, and a counting task. Verbal Cognition measures the subject's ability to define five words. To evaluate changes over time, we also studied wave-to-wave changes in Total Cognition and Verbal Cognition. Our next cognitive outcome, Alzheimer's, is an indicator variable equal to 1 for subjects who report having been diagnosed with Alzheimer's disease, and 0 otherwise. Since the HRS data are longitudinal, the unit of analysis for our 4 cognitive outcomes is a person-year. For these analyses, because an individual took the cognitive tests at different ages, in our set of controls we replaced our person-specific age variable with age at assessment (which differs for an individual across the cognitive outcomes); we also clustered all standard errors at the person level.

In the WLS, we measured cognitive performance using a respondent's raw score on a Henmon-Nelson test of mental ability[48].

For all of these additional prediction exercises, results are shown in **Supplementary Table 38** and depicted in **Figure 4A** and **Supplementary Figures 13** and **11**.

Benchmarking the Predictive Power of the *EduYears* Polygenic Score. To benchmark our score's predictive power, we compared its predictive power to the predictive power of other common variables: mother's education, father's education, both mother's and father's education, verbal cognition, household income, and a binary indicator for marital status. For each variable, we calculated the variable's incremental R^2 using the same procedures as those described above, with the same set of control variables. (For "mother's and father's education," we calculated the incremental R^2 from adding both variables as regressors.) The results of this analysis are shown in **Supplementary Table 39A** and depicted in **Figure 4B** and **Supplementary Figure 12**.

We also evaluated the attenuation in the incremental R^2 of the polygenic score in predicting *EduYears* when we control for available demographic variables one at a time: marital status, household income, mother's education, and father's education. We next controlled for both mother's and father's education, and finally, we controlled for the full set of demographic controls. The results of this analysis are shown in **Supplementary Table 39B** and **Supplementary Figure 12.**

GWAS of *Cognitive Performance*, *Math Ability* and *Highest Math*. The GWAS of *Math Ability* (N = 564,698) and *Highest Math* (N = 430,445) phenotypes were conducted exclusively among research participants of the personal genomics company *23andMe* who answered survey questions about their mathematical background. In our analyses of *Cognitive Performance*, we combined a published study of general cognitive ability (N = 35,298) conducted by the COGENT consortium[28] with new genome-wide association analyses of cognitive performance in the UK Biobank (N = 222,543). The phenotype measures are described in detail in **Supplementary Table 40**. Our new genome-wide analyses of *Cognitive Performance* in *UKB*, and *Math Ability* and *Highest Math* in *23andMe*, were conducted using methods identical to those for *EduYears* in *UKB* and *23andMe*, respectively (**Supplementary Table 19**).

For *Cognitive Performance*, we conducted a sample-size-weighted meta-analysis (N = 257,841), imposing a minimum-sample-size filter of 100,000. We similarly applied minimum-sample-size filters to the *Math Ability* (N > 500,000) and *Highest Math* (N > 350,000) results. We adjusted the test statistics using the estimated intercepts from LD Score regressions (1.073 for *Math Ability*, 1.105 for *Highest Math*, and 1.046 for *Cognitive Performance*). The summary statistics underwent quality control using the same procedures applied to the *EduYears* results files.

The lists of lead SNPs were obtained by applying the same clumping algorithm used in the *EduYears* analyses (**Supplementary Tables 11-13**). Manhattan plots from the analyses are shown in **Supplementary Figures 14-16.**

MTAG of *Cognitive Performance*, *Math Ability* and *Highest Math*. We performed a joint analysis of our GWAS results on *EduYears*, *Cognitive Performance*, *Math Ability*, and *High Math* using MTAG[11]. **Supplementary Table 14** shows moderately high pairwise genetic correlations, ranging from 0.51 to 0.85, which motivate the multivariate analysis. The MTAG analyses were restricted to SNPs that passed MTAG-recommended filters in all files with summary statistics. We dropped (i) SNPs with minor allele frequency below 1% or (ii) SNPs with sample sizes below a cutoff (66.6% of the 90th percentile), leaving approximately 7.1 million SNPs found in all four results files. **Supplementary Table 41** reports the increases in effective sample size from using MTAG for each set of GWAS results.

Supplementary Table 15 lists all the lead SNPs in the MTAG analysis. **Supplementary Figures 17-20** show inverted Manhattan plots that compare the MTAG and GWAS results, restricted to the set of SNPs that pass MTAG filters.

Polygenic scores were constructed from MTAG results using the same procedures as for the GWAS results. **Supplementary Figure 29** and **Supplementary Tables 42** and **43** compare the predictive power of scores constructed from MTAG results in the Add Health and WLS cohorts (see **Supplementary Note** for details).

To examine the credibility of the MTAG-identified lead SNPs of our lowest-powered GWAS, *Cognitive Performance*, we conducted a replication analysis. We re-ran MTAG with GWAS results that exclude COGENT cohorts, and we used the COGENT meta-analysis as our replication sample. In addition to applying the MTAG filters above, we limited the analysis to SNPs for which the COGENT results file contains summary statistics based on analyses of at least 25,000 individuals. The MTAG-identified lead SNPs for *Cognitive Performance* from our restricted sampled are reported in **Supplementary Table 44**. We used our Bayesian framework to calculate the expected replication record of the MTAG results under the hypothesis that the MTAG-identified lead SNPs are true positives, given sampling variation and adjusted for winner's curse and differences in SNP heritability across the samples.

DATA AVAILABILITY AND ACCESSION CODES

Summary statistics can be downloaded from www.thessgac.org/data. We provide association results for all SNPs that passed quality-control filters in a GWAS meta-analysis of *EduYears* that excludes the research participants from 23andMe. SNP-level summary statistics from analyses based entirely or in part on 23andMe data can only be reported for up to 10,000 SNPs. We provide summary statistics for all lead SNPs identified in our GWAS analyses of *Cognitive Performance*, *Math Ability*, and *Highest Math* and the MTAG analyses of our four phenotypes. For the complete *EduYears* GWAS, which includes 23andMe, clumped results for the 3,575 SNPs with $P < 10^{-5}$ are provided; this *P*-value threshold was chosen such that the total number of SNPs across the analyses that include data from 23andMe does not exceed 10,000. Contact information for each of the cohorts included in this paper can be found in the **Supplementary Note**.

CODE AVAILABILITY:

All software used to perform these analyses are available online.

URLs:

Social Science Genetic Association Consortium (SSGAC) website: http://www.thessgac.org/#!data/kuzq8.

Minimac2: https://genome.sph.umich.edu/wiki/Minimac2

BEAGLE v2.1.2: http://faculty.washington.edu/browning/beagle/b3.html

IMPUTE2 v2.3.1: http://mathgen.stats.ox.ac.uk/impute/impute_v2.html

PBWT: https://github.com/richarddurbin/pbwt

IMPUTE4: https://jmarchini.org/impute-4/

ShapeIT v2.r790: http://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html

BOLT-LMM: https://data.broadinstitute.org/alkesgroup/BOLT-LMM/

SNPTEST v2.4.1: https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html

REGSCAN v0.2.0: https://www.geenivaramu.ee/en/tools/regscan

METAL, release 2011-03-25: http://csg.sph.umich.edu/abecasis/metal/

EasyQC v9.0: http://www.uni-regensburg.de/medizin/epidemiologie-praeventivmedizin/genetische-epidemiologie/software/

ldsc v1.0.0: https://github.com/bulik/ldsc

Plink, 1.90b3p: http://zzz.bwh.harvard.edu/plink/plink2.shtml

LDpred v0.9.09: https://bitbucket.org/bjarni_vilhjalmsson/ldpred

Stata v14.2: https://www.stata.com/install-guide/windows/download/

DEPICT (downloaded Feb 2015): https://data.broadinstitute.org/mpg/depict/

MAGMA v1.06b: https://ctg.cncr.nl/software/magma

PANTHER release 20170403: http://www.geneontology.org

CAVIARBF v0.2.1: https://bitbucket.org/Wenan/caviarbf

MTAG software v1.0.1: https://github.com/omeed-maghzian/mtag

METHODS-ONLY REFERENCES

35. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26,** 2190–2191 (2010).
36. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4,** 1–16 (2015).
37. Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48,** 624–633 (2016).
38. Cochran, W. G. The Combination of Estimates from Different Experiments. *Biometrics* **10,** 101 (1954).
39. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47,** 1236–1241 (2015).
40. Cameron, A. C. & Miller, D. Robust inference with dyadic data. *mimeo* (2014). doi:10.1201/b10440
41. The 1000 Genomes Project Consortium *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491,** 56–65 (2012).
42. Fehrmann, R. S. N. *et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat. Genet.* **47,** 115–125 (2015).
43. de Leeuw, C. A. *et al.* MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* **11,** e1004219 (2015).
44. Liu, J. Z. *et al.* A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* **87,** 139–145 (2010).
45. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8,** 1551–1566 (2013).
46. Pickrell, J. K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am. J. Hum. Genet.* **94,** 559–573 (2014).
47. Chen, W. *et al.* Fine mapping causal variants with an approximate Bayesian method using marginal test statistics. *Genetics* **200,** 719–736 (2015).
48. Henmon, V. A. C. Henmon-Nelson Tests of Mental Ability, High School Examination-Grades 7 to 12-Forms A, B, and C. Teacher's Manual. (1946).