



Common Genetic Variants Associated With Cognitive Performance Identified Using the Proxy-Phenotype Method

Citation

Rietveld, C. A., T. Esko, G. Davies, T. H. Pers, P. Turley, B. Benyamin, C. F. Chabris, et al. 2014. "Common Genetic Variants Associated with Cognitive Performance Identified Using the Proxy-Phenotype Method." *Proceedings of the National Academy of Sciences* 111 (38) (September 8): 13790–13794. doi:10.1073/pnas.1404623111.

Published Version

doi:10.1073/pnas.1404623111

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:23597722>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Title: Common Genetic Variants Associated with Cognitive Performance Identified Using Proxy-Phenotype Method

Authors: All authors and their affiliations appear at the end of the paper*

*Correspondence to: Daniel Benjamin (db468@cornell.edu)
or Philipp Koellinger (p.d.koellinger@uva.nl).

Abstract: We identify common genetic variants associated with cognitive performance using a two-stage approach, which we call the “proxy-phenotype method.” First, we conduct a genome-wide association study of educational attainment in a large sample ($N = 106,736$), which produces a set of 69 “education-associated single-nucleotide polymorphisms (SNPs).” Second, using independent samples ($N = 24,189$), we measure the association of these education-associated SNPs with cognitive performance. Three SNPs (rs1487441, rs7923609, rs2721173) are significantly associated with cognitive performance after correction for multiple hypothesis testing. In an independent sample of older Americans ($N = 8,652$), we also show that a polygenic score derived from the education-associated SNPs is associated with memory and absence of dementia. Convergent evidence from a set of bioinformatics analyses implicates four specific genes (*KNCMA1*, *NRXN1*, *POU2F3*, *SCRT*). All of these are associated with a particular neurotransmitter pathway involved in synaptic plasticity, the main cellular mechanism for learning and memory.

Significance Statement: We identify several common genetic variants associated with cognitive performance using a two-stage approach: we conduct a genome-wide association study of educational attainment to generate a set of candidates, then we measure the association of these variants with cognitive performance. In older Americans, we find that these variants are jointly associated with cognitive health. Bioinformatics analyses implicate a set of genes that are associated with a particular neurotransmitter pathway involved in synaptic plasticity, the main cellular mechanism for learning and memory. In addition to the substantive contribution, this paper also serves to demonstrate a “proxy-phenotype” approach to discovering common genetic variants that is likely to be useful for many phenotypes of interest to social scientists (such as personality traits).

Introduction: Twin and family studies have shown that at least a moderate share of variation in most facets of cognitive performance (i.e., performance by healthy individuals on cognitive tests) is associated with genetic factors (1, 2). However, despite considerable interest and effort, research to date has largely failed to identify common genetic variants associated with cognitive performance phenotypes (3–5), with the exception of *APOE* which predicts cognitive decline in older individuals (6–8). Existing studies have relied on one of two research strategies. The first is a candidate-gene design, in which researchers test a small number of genetic variants for association with the phenotype of interest, typically based on hypotheses derived from the known biological functions of the candidate genes. The candidate-gene associations that have been reported with cognitive performance (9), however, fail to replicate when larger samples are used (3). The second research strategy is a genome-wide association study (GWAS), in which researchers theoretically test hundreds of thousands of single-nucleotide polymorphisms (SNPs) for association with the phenotype and apply a threshold for “genome-wide” statistical significance—typically 5×10^{-8} —in order to account for multiple-hypothesis testing. For physical and medical phenotypes, GWASs have identified many novel associations that replicate (10). GWASs on cognitive performance, however, have not yet identified any genome-wide-significant associations (4, 5).

Here, we apply an alternative, two-stage research strategy, which we call the proxy-phenotype method. In the first stage, we conduct a GWAS on a “proxy phenotype” to identify a relatively small set of SNPs that are associated with the proxy phenotype. In the second stage, these SNPs serve as candidates that are tested in independent samples for association with the phenotype of interest, at a significance threshold corrected for the number of proxy-associated SNPs. In the study reported here, our phenotype of interest is cognitive performance, for which we use Spearman’s measure of general cognitive ability (usually abbreviated to *g*; it is the general factor measured by a battery of diverse cognitive tests (4)). Our proxy phenotype is educational attainment, as measured by self-reported years of schooling.

Rietveld et al. (11) had suggested the strategy of using SNPs associated with educational attainment as “empirically-based candidate genes” for association with cognitive performance; here we conduct that analysis and further develop the methodology for doing so. The SI Appendix contains our formal framework, building on that in (11), as well as power calculations under a range of assumptions. According to the framework, educational attainment is a good proxy phenotype for cognitive performance because cognitive performance is strongly genetically influenced and causally affects educational attainment. The genetic correlation between the two traits does not have straightforward implications for the statistical power to identify specific SNPs influencing cognitive performance; nonetheless, the high genetic correlation (estimated to be roughly 0.65 or higher (12–14)) may also provide a suggestive justification for the approach.

Results: In our first stage, we conducted a GWAS of educational attainment in a pooled “Education Sample” of 106,736 individuals. We used the same data, analysis protocol, and quantitative years-of-schooling measure as (11), except that we omit cohorts with high-quality measures of cognitive performance; we instead include these cohorts in the subsequent “Cognitive Performance Sample.” We chose our “inclusion threshold” of $p < 10^{-5}$ for selecting candidate SNPs based on ex ante power calculations whose goal was to maximize the number of

true positives among the candidates (see SI Appendix). Pruning for linkage disequilibrium the 927 SNPs that reach this threshold resulted in 69 approximately independent SNPs (see SI Appendix).

In our second stage, we tested these 69 “education-associated SNPs” for association with cognitive performance in the Cognitive Performance Sample, which comprises 24,189 genotyped subjects from 11 cohorts (see SI Appendix section 2). The specific cognitive tests differ across cohorts, but the cognitive performance measure in every cohort is calculated as Spearman’s g (see SI Appendix); previous research has found that g from different test batteries are highly correlated, especially if the batteries have many tests, or if the test is specifically constructed to measure g (15–17). We tested each SNP individually for association with cognitive performance using ordinary least squares, controlling for sex, age, and (depending on the cohort) at least four principal components of the genome-wide data (to reduce confounding from population stratification). At the cohort level, the analyses were conducted according to a prespecified plan that we preregistered on the Open Science Framework (see <https://osf.io/z7fe2/>). The cohorts’ results were then meta-analyzed using an inverse-variance weighting scheme. Two independent teams of analysts crosschecked and verified the results.

To confirm that the education-based first stage identifies reasonable candidate SNPs for cognitive performance, **Figure 1** plots the standardized regression coefficients from the regression of years-of-schooling on the education-associated SNPs in the Education Sample (with the reference allele chosen to ensure the coefficient is positive) against the standardized coefficients from the second-stage regression of cognitive performance on the SNPs in the Cognitive Performance Sample. The direction of the effect coincides in 53 out of 69 cases (two-sided binomial test, $p = 9.10 \times 10^{-6}$), indicating that this is a good context for applying the proxy-phenotype method. We were surprised that the correlation between the effect size on educational attainment and the effect size on cognitive performance is negative ($\rho = -0.25$; $p = 0.03$), although not significantly after dropping a possible outlier, the bottom-most point of the figure ($\rho = -0.14$ $p = 0.26$). Within our theoretical framework, a negative correlation suggests that SNPs that affect cognitive performance more strongly tend to affect other factors that matter for educational attainment (such as personality traits) less strongly, and vice-versa (see SI Appendix).

To provide a benchmark for evaluating our list of education-associated candidate SNPs, we generated (via a pre-specified algorithm) a list of “theory-based” candidate SNPs for cognitive performance drawn from published findings in the candidate-gene literature (see SI Appendix). (This list does not include the SNPs comprising the *APOE* haplotype because these SNPs were not available in the cohort GWAS results.) After applying the same pruning procedure as for the education-associated SNPs, our list of theory-based SNPs contains 24 independent SNPs, of which only one is in a genomic region close to an education-associated SNP. **Figure 2** overlays Q–Q plots for the theory-based and education-associated candidates. The education-associated candidates taken altogether are more strongly associated with cognitive performance than would be expected by chance ($z = 5.98$, $p = 1.12 \times 10^{-9}$). Whereas a visual inspection of the plot suggests that the theory-based candidates exhibit some association with cognitive performance, we cannot reject the null hypothesis for any SNP individually, nor for all of them taken together ($z = 1.19$, $p = 0.12$).

The top three education-associated SNPs—rs1487441, rs7923609, and rs2721173—show clear separation from the others in **Figure 2** and are significantly associated with cognitive performance after Bonferroni correction for multiple hypothesis testing (see **Table 1**). Consistent with the negative correlation in **Figure 1**, these SNPs are different from the three SNPs that reached genome-wide significance for association with educational attainment in the (11) analyses. After adjusting the SNPs’ estimated effect sizes (each $R^2 \approx 0.0006$) for the winner’s curse, we estimate each as $R^2 \approx 0.0002$ (see SI Appendix), or in terms of coefficient magnitude, each additional reference allele for each SNP is associated with ≈ 0.02 standard-deviation increase in cognitive performance (or 0.3 points on the typical “IQ” scale). This $R^2 \approx 0.0002$ is about the same as the R^2 for the known SNP associations with educational attainment (11) but far smaller than the largest effect sizes for complex physical traits such as height ($R^2 \approx 0.004$) and BMI ($R^2 \approx 0.003$) (18, 19).

Power calculations we report in the SI Appendix help shed light on why the proxy-phenotype method succeeded in identifying SNPs even though GWA studies to date on cognitive performance have not. A GWAS in our Cognitive Performance Sample of $N = 24,189$ —which is *larger* than the largest GWA studies ($N = 17,989$ in Benyamin et al. (2014) and $N = 3,511$ in Davies et al. (2011))—would have had power 0.06% to identify a SNP whose association has $R^2 = 0.0002$. In contrast, our proxy-phenotype approach had power 12%. Given this power and the rather stringent significance threshold ($0.05/69 \approx .00072$), Bayesian calculations using reasonable assumptions regarding priors suggest that the posterior probabilities that these three SNPs are associated with cognitive performance are high (see SI Appendix).

Turning from specific SNPs to the set of all 69 education-associated SNPs, we assess the explanatory power of a linear polygenic score that aggregates their coefficients (see SI Appendix). In pooled results from four family-based cohorts (4,463 individuals in total), we find that the score is significantly associated with cognitive performance ($p = 8.17 \times 10^{-4}$), with R^2 ranging approximately from 0.2% to 0.4% across samples. Using only within-family variation, the pooled coefficient has the same sign but is smaller and has a larger standard error ($p = 0.36$). Thus we cannot rule out that some of the score’s explanatory power is due to population stratification, although even without stratification, the non-significance of the within-family coefficient is not surprising given the low power of this test (see SI Appendix).

Next, we explore whether educational attainment might serve as a proxy phenotype for cognitive-health phenotypes (as opposed to cognitive performance in the normal range). Our sample comprises 8,652 European-descent individuals over the age of 50 from the Health and Retirement Study (HRS) (see SI Appendix). We confirm that, for the 60 out of 69 SNPs available in the HRS data, the direction of the effects on educational attainment generally coincides with the direction of the effects on the two cognitive-health phenotypes we study: “total word recall,” which is a test for memory problems (two-sided binomial test, $p = 0.0067$); and “total mental status,” which is a battery that screens for early signs of dementia ($p = 0.0775$). Next, we obtain the weights for a polygenic score by conducting a *de novo* meta-GWAS analysis of educational attainment just as in the first stage described above, but this time excluding the HRS from the Education Sample.

Figure 3 shows that the score is associated with both of the cognitive-health phenotypes. The strength of the protective effect is approximately constant across age categories from age 50 to 80, and becomes weaker for total word recall after age 80. These associations are essentially unaffected when we control for up to 20 principal components of the genome-wide data,

suggesting that the associations are not driven by population stratification (20). The R^2 of these associations range roughly 0.2%-0.4% (similar magnitudes as in the analysis of cognitive performance in the family-based cohorts). When we control for years of schooling, the estimated effect of the score falls roughly in half but remains statistically significant (see SI Appendix). The score is not associated with cognitive decline (i.e., the change in a cognitive phenotype across longitudinal survey waves), except for total word recall after age 80.

Finally, we used the 14 (out of 69) education-associated SNPs that are nominally significantly associated with cognitive performance ($p < .05$) to explore possible biological pathways in a set of bioinformatic analyses (see SI Appendix). Two of the 14 SNPs are in gene deserts, but the other 12 are in close vicinity to at least one gene predicted (based on its expression profile) to be involved in the nervous system (see SI Appendix). Among the most promising genes across these loci are *KNCMA1*, *NRXN1*, *POU2F3*, and *SCRT*, all of which are predicted to be involved in a glutamate neurotransmission pathway (labeled in REACTOME as “unblocking of NMDA receptor, glutamate binding, and activation”) that is involved in synaptic plasticity, a cellular mechanism for learning and memory. Using different methods (but some overlapping data), this same pathway has previously been implicated in human cognitive performance (21).

Discussion: This paper makes two contributions. First, we demonstrate that the “proxy-phenotype method” generates positive findings in a domain in which neither candidate-gene nor GWAS approaches have so far made substantial progress. Similar approaches have sometimes been used in prior work (e.g., to find rare structural variants associated with cognition; (22)), and there is existing work focused on the related idea of increasing statistical power in GWAS by analyzing correlated phenotypes jointly (23, 24).

We propose that the proxy-phenotype method, if systematically applied in social-science genetics, could be a useful complement to traditional gene discovery methods (such as GWAS) in cases where it affords greater statistical power. In the present case, it does so because (i) much larger genotyped samples are available for educational attainment than for cognitive performance, and (ii) some genetic variants are likely to be associated with educational attainment due to their more direct, stronger relationships with cognitive performance. For the same reasons, educational attainment might similarly serve as a proxy phenotype for personality traits such as persistence and self-control. In other contexts, the proxy-phenotype method may be better powered for different reasons. For example, for behavioral phenotypes with substantial measurement error—such as smoking, drinking, exercise, or eating habits—the proxy phenotype could be a medical outcome associated with the behavior (e.g., pulmonary disease for smoking, cirrhosis for alcohol consumption). We also note that, while our analysis plan specified that cohorts look up a relatively small set of education-associated SNPs in their existing GWAS results on cognitive performance, researchers with access to full GWAS results on the phenotype of interest could implement a more powerful version of the proxy-phenotype method. For example, first-stage results on the proxy phenotype could inform priors that are updated using GWAS results on the phenotype of interest.

We caution that the proxy-phenotype method (like theory-based candidate-SNP approaches) could generate an unacceptably high rate of false positives if it were applied when underpowered and if results were reported selectively. To avoid this, we propose a set of “best practices” that proxy-phenotype studies should follow: researchers should (a) conduct power calculations *ex ante* to justify the use of the method for a particular phenotype of interest, and report these

calculations in the SI; (b) circulate an analysis plan to all cohorts prior to conducting any analysis, and register the plan in a public repository; (c) commit to publishing all findings from the study, including null results; and (d) conduct Bayesian calculations of the credibility of any findings. We followed these procedures in this paper. While replication of findings in an independent cohort would be ideal, we anticipate that it will often be infeasible given the unavailability of genotyped samples that may motivate the proxy-phenotype approach in the first place.

The second contribution of this paper is to identify common genetic variants associated with cognitive phenotypes. Knowing the three significant SNPs is not useful for predicting any particular individual's cognitive performance because the effect sizes are far too small, but it does enable follow-up research—e.g., pinpointing the causal variants and then conducting knock-out experiments in animals—that may ultimately shed light on biological pathways underlying cognitive variation. The polygenic scores constructed from our results may prove useful for studying gene-environment interactions. In future work, the magnitude of explained variance will increase as researchers gain access to datasets with even larger first-stage samples. Our results suggest that such scores hold promise for eventually identifying individuals whose cognitive health at older ages is at greatest risk, which could allow for appropriate preparation and (if possible) preventative intervention.

Materials and Methods: See SI Appendix for all details on the samples and methods.

Acknowledgments: This research was carried out under the auspices of the Social Science Genetics Association Consortium (SSGAC), a cooperative enterprise among medical researchers and social scientists that coordinates genetic association studies for social science variables. Data for our analyses come from many studies and organizations, some of which are subject to an MTA (see SI Appendix). Results from the meta-analysis, the complete biological annotation, and a FAQ document describing the findings of this paper are available at the website of the consortium, <http://www.ssgac.org>.

The formation of the SSGAC was made possible by an EAGER grant from the NSF and a supplemental grant from the NIH/OBSSR (SES-1064089). This research was funded in part by the Ragnar Söderberg Foundation (E9/11), the Swedish Research Council (412-2013-1061), and by the NIA/NIH through grants P01-AG005842, P01-AG005842-20S2, P30-AG012810, and T32-AG000186-23. For a full list of acknowledgments, see SI Appendix.

References:

1. Bouchard TJ, McGue M (2003) Genetic and environmental influences on human psychological differences. *J Neurobiol* 54:4–45.
2. Plomin R, DeFries J, Knopik V, Neiderhiser J (2013) *Behavioral Genetics* (Worth Publishers).
3. Chabris CF et al. (2012) Most reported genetic associations with general intelligence are probably false positives. *Psychol Sci* 23:1314–1323.
4. Benyamin B et al. (2014) Childhood intelligence is heritable, highly polygenic and associated with FBNP1L. *Mol Psychiatry* 19:253–258.
5. Davies G et al. (2011) Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol Psychiatry* 16:996–1005.
6. Wisdom NM, Callahan JL, Hawkins KA (2011) The effects of apolipoprotein E on non-impaired cognitive functioning: A meta-analysis. *Neurobiol Aging* 32:63–74.
7. Lambert J-C et al. (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 45:1–9.
8. Davies G et al. (2014) A genome-wide association study implicates the APOE locus in nonpathological cognitive ageing. *Mol Psychiatry* 19:76–87.
9. Payton A (2009) The impact of genetic research on our understanding of normal cognitive ageing: 1995 to 2009. *Neuropsychol Rev* 19:451–77.
10. Visscher PMM, Brown MAA, McCarthy MII, Yang J (2012) Five years of GWAS discovery. *Am J Hum Genet* 90:7–24.
11. Rietveld CA et al. (2013) GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* (80-) 340:1467–1471.
12. Wainwright MA, Wright MJ, Geffen GM, Luciano M, Martin NG (2005) The genetic basis of academic achievement on the Queensland Core Skills Test and its shared genetic variance with IQ. *Behav Genet* 35:133–45.
13. Calvin C et al. (2012) Multivariate genetic analyses of cognition and academic achievement from two population samples of 174,000 and 166,000 school children. *Behav Genet* 42:699–710.
14. Marioni RE et al. (2014) Molecular genetic contributions to socioeconomic status and intelligence. *Intelligence* 44:26–32.

15. Johnson W, Bouchard TJ, Krueger RF, McGue M, Gottesman II (2004) Just one g: consistent results from three test batteries. *Intelligence* 32:95–107.
16. Ree MJ, Earles JA (1991) The stability of g across different methods of estimation. *Intelligence* 15:271–278.
17. Chabris CF (2007) in *Integrating the Mind: Domain General Versus Domain Specific Processes in Higher Cognition*, ed Roberts MJ (Psychology Press, Hove, UK), pp 449–491.
18. Lango Allen H et al. (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467:832–838.
19. Speliotes EK et al. (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 42:937–948.
20. Price AL et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909.
21. Hill WD et al. (2014) Human cognitive ability is influenced by genetic variation in components of postsynaptic signalling complexes assembled by NMDA receptors and MAGUK proteins. *Transl Psychiatry* 4:e341.
22. Stefansson H et al. (2014) CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* 505:361–366.
23. Ferreira MAR, Purcell SM (2009) A multivariate test of association. *Bioinformatics* 25:132–3.
24. Galesloot TE, van Steen K, Kiemeny LALM, Janss LL, Vermeulen SH (2014) A comparison of multivariate genome-wide association methods. *PLoS One* 9:e95923.

Table 1. The SNPs significantly associated with cognitive performance after Bonferroni correction (for full results see Table S4). The chromosome and basepair position are from the NCBI genome annotation (build 36), and the nearest gene from the SCAN database. “Allele frequency” refers to the Cognitive Performance Sample.

SNP	Chromosome	Basepair position	Nearest gene	Effective allele	Allele frequency	Years of Schooling (Education Sample)		Cognitive Performance (Cognitive Performance Sample)	
						Standardized coefficient	<i>P</i> -value	Standardized coefficient	<i>P</i> -value
rs1487441	6	98660615	<i>LOC100129158</i>	A	0.473	0.026	1.78×10^{-9}	0.036	1.24×10^{-4}
rs7923609	10	64803828	<i>JMJD1C</i>	A	0.521	-0.021	1.06×10^{-6}	-0.034	2.58×10^{-4}
rs2721173	8	145715237	<i>LRRC14</i>	T	0.473	-0.020	8.61×10^{-6}	-0.034	2.88×10^{-4}

Figure 1. The relationship between standardized coefficients from the first-stage regression of years of schooling on the education-associated SNPs in the Education Sample (x-axis) and standardized coefficients from the second-stage regression of cognitive performance on these SNPs in the Cognitive Performance Sample (y-axis). The reference allele is chosen such that the coefficient on years of schooling is positive. Each point represents one of the 69 education-associated SNPs. (The cloud of points is bounded away from zero effect on years of schooling because the criterion for including a SNP was its reaching $p < 10^{-5}$ in the GWAS on years of schooling in the Education Sample.) Since the standard deviation of years of schooling is approximately 3, a coefficient of 0.03—a typical size for a years-of-schooling standardized coefficient—means that each reference allele is associated with an increase of $0.03 \times 3 \approx 0.09$ years of educational attainment. In conventional “IQ” units that have a standard deviation of 15, a standardized regression coefficient on cognitive performance of 0.03 corresponds to ≈ 0.45 “IQ points.”

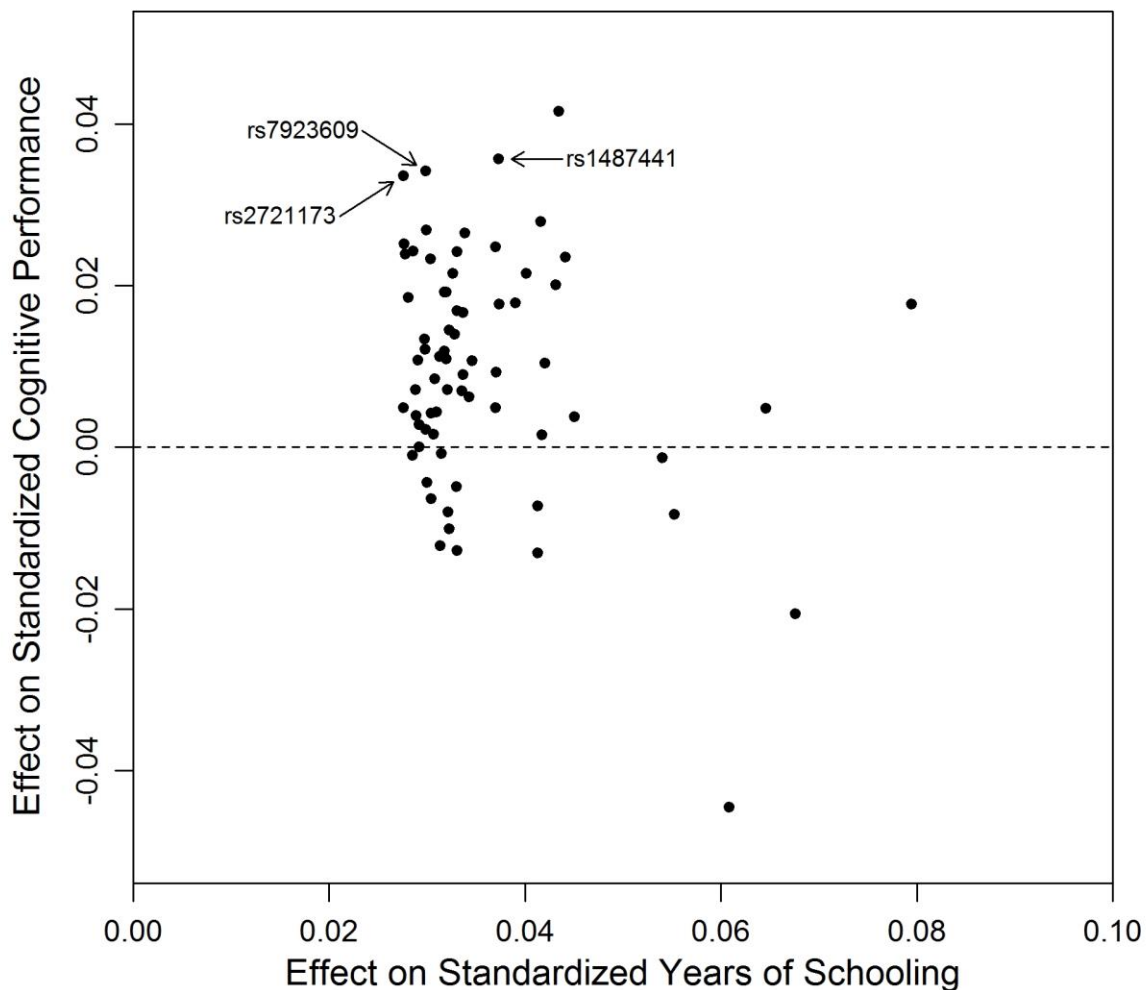


Figure 2. Q–Q plot for a regression of cognitive performance on the education-associated SNPs (the dark points) with 95% confidence interval around the null hypothesis (the darkly shaded region); and Q–Q plot for a regression of cognitive performance on the theory-based SNPs (the light points) with 95% confidence interval around the null hypothesis (the lightly shaded region). The table shows the nominal effect sizes and p -values for the three labeled SNPs, which are the SNPs are statistically significantly associated with cognitive performance after Bonferroni correction (for testing the 69 education-associated SNPs).

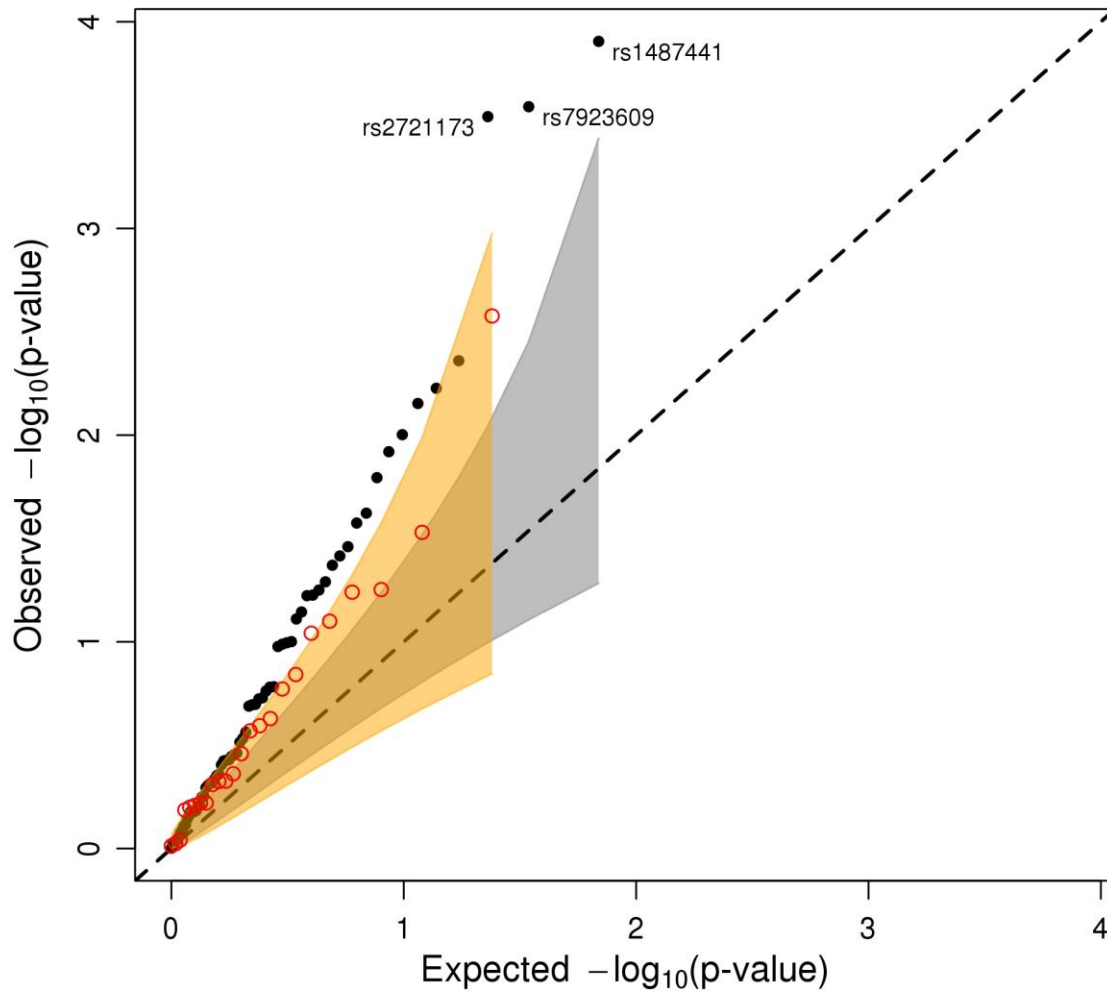
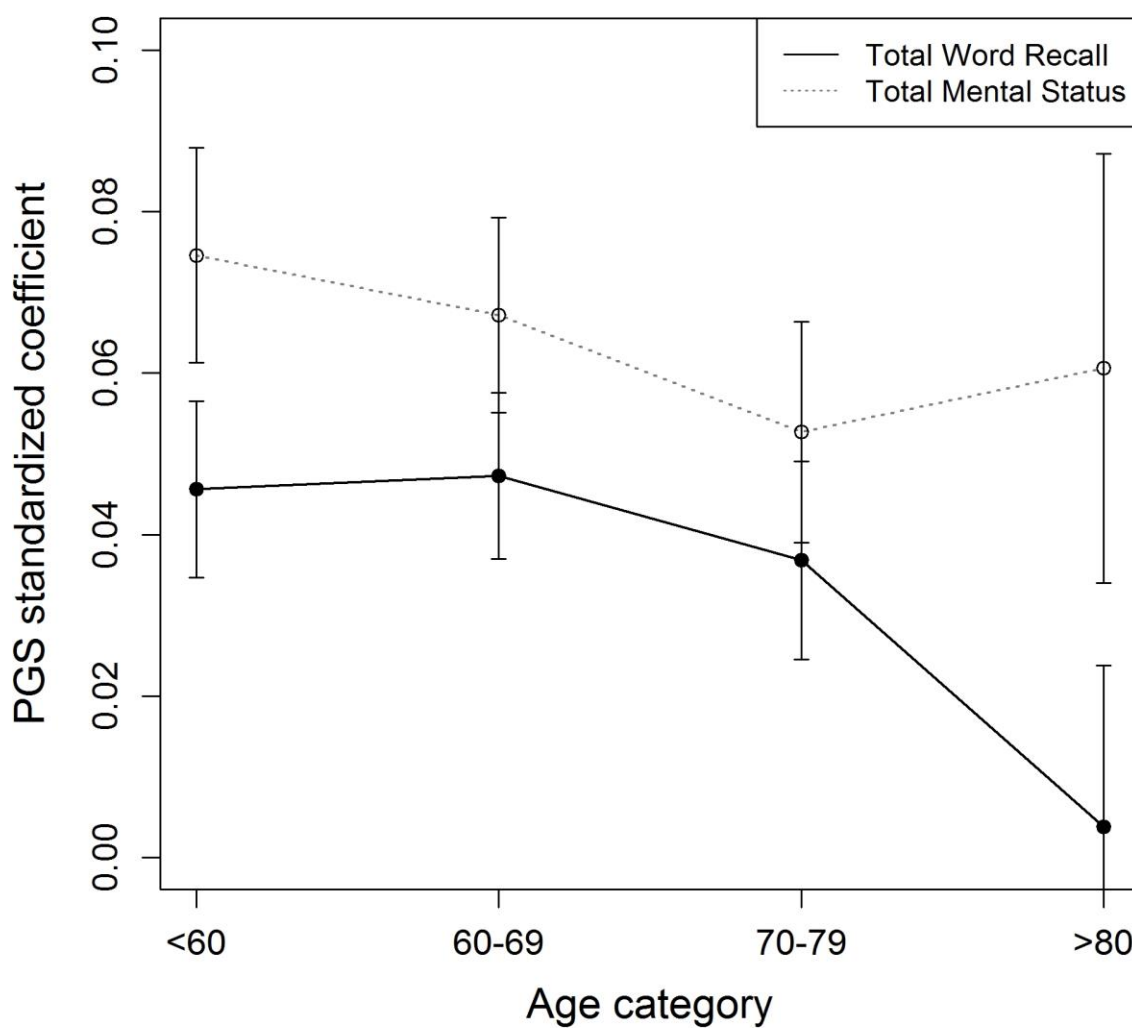


Figure 3. Coefficients from regression of standardized cognitive phenotype (Total Word Recall or Total Mental Status) on standardized polygenic score within age category, controlling for sex and clustering standard errors by individual (see SI Section 14 for details). Error bars show ± 1 standard error.



Authors:

Cornelius A. Rietveld,^{1,2} Tõnu Esko,^{3,4,5,6} Gail Davies,^{7,41} Tune H. Pers,^{3,4} Beben Benjamin,⁸ Christopher F. Chabris,⁹ Valur Emilsson,^{10,11} Andrew D. Johnson,¹² James J. Lee,^{13,14} Christiaan de Leeuw,^{15,16} Riccardo E. Marioni,^{7,8,17} Sarah E. Medland,¹⁸ Michael B. Miller,¹⁴ Olga Rostapshova,¹⁹ Patrick A. Turley,²⁰ Sven J. Van der Lee,²¹ Anna A.E. Vinkhuyzen,⁸ Najaf Amin,²¹ Dalton Conley,²² Jaime Derringer,²³ Cornelia M. van Duijn,^{21,24} Rudolf Fehrmann,²⁵ Lude Franke,²⁵ Generation Scotland,²⁶ Edward L. Glaeser,²⁰ Narelle K. Hansell,²⁷ Caroline Hayward,^{17,28} William G. Iacono,¹⁴ Carla A. Ibrahim-Verbaas,^{21,29} Vincent Jaddoe,^{2,30} Juha Karjalainen,²⁵ David Laibson,³¹ Paul Lichtenstein,³² David C. Liewald,⁷ Patrik K.E. Magnusson,³² Nicholas G. Martin,¹⁸ Matt McGue,¹⁴ George McMahon,³³ Nancy L. Pedersen,³² Steven Pinker,¹³ David J. Porteous,^{7,17} Danielle Posthuma,^{15,34,35} Fernando Rivadeneira,^{2,36} Blair H. Smith,³⁷ John M. Starr,^{7,38} Henning Tiemeier,^{2,34} Nicholas J. Timpson,³⁹ Maciej Trzaskowski,⁴⁰ André G. Uitterlinden,^{2,36} Frank C. Verhulst,³⁴ Mary E. Ward,³³ Margaret J. Wright,²⁷ George Davey-Smith,³⁹ Ian J. Deary,^{7,41} Magnus Johannesson,⁴² Robert Plomin,⁴⁰ Peter M. Visscher,^{8,43} Daniel J. Benjamin,^{44,*} David Cesarini,^{45,*} Philipp D. Koellinger,^{1,2,46,*} and the Social Science Genetic Association Consortium

Affiliations:

¹ Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands

² Department of Epidemiology, Erasmus Medical Center, Rotterdam 3000 CA, The Netherlands

³ Boston Children's Hospital, Boston, Massachusetts 02115, United States of America

⁴ Broad Institute of the Massachusetts Institute of Technology and Harvard, Massachusetts 02142, United States of America

⁵ Harvard Medical School, Boston, Massachusetts 02115, United States of America

⁶ Estonian Genome Center, University of Tartu, Tartu 51010, Estonia

⁷ Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh EH8 9JZ, Scotland, United Kingdom

⁸ Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia

⁹ Department of Psychology, Union College, Schenectady, New York 12308, United States of America

¹⁰ Icelandic Heart Association, Kopavogur 201, Iceland

¹¹ Faculty of Pharmaceutical Sciences, University of Iceland, 107 Reykjavík, Iceland

¹² Framingham Heart Study, National Heart, Lung, and Blood Institute, Framingham, Massachusetts 01702, United States of America

¹³ Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, United States of America

¹⁴ Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455-0344, United States of America

¹⁵ Department of Functional Genomics, VU University Amsterdam and VU Medical Center, 1081 HV Amsterdam, the Netherlands

- ¹⁶ Machine Learning Group, Intelligent Systems, Institute for Computing and Information Sciences, Faculty of Science, Radboud University Nijmegen, 6500 GL Nijmegen, The Netherlands
- ¹⁷ Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, United Kingdom
- ¹⁸ QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4029, Australia
- ¹⁹ Harvard Kennedy School, Harvard University, Cambridge, Massachusetts 02139, United States of America
- ²⁰ Department of Economics, Harvard University, Cambridge, Massachusetts 02138, United States of America
- ²¹ Genetic Epidemiology Unit, Department of Epidemiology and Biostatistics, Erasmus Medical Center, Rotterdam 3000 CA, the Netherlands
- ²² Department of Sociology, New York University, New York, New York 10012, United States of America
- ²³ Department of Psychology, University of Illinois, Urbana-Champaign, Illinois 61820, United States of America
- ²⁴ Centre for Medical Systems Biology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands
- ²⁵ Department of Genetics, University Medical Center Groningen, University of Groningen, 9713 GZ Groningen, The Netherlands
- ²⁶ Generation Scotland, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, United Kingdom
- ²⁷ Neuroimaging Genetics Group, QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4029, Australia
- ²⁸ MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, United Kingdom
- ²⁹ Department of Neurology, Erasmus Medical Center, Rotterdam 3000 CA, the Netherlands
- ³⁰ The Generation R Study Group, Erasmus Medical Center, 3000 CA Rotterdam, The Netherlands
- ³¹ Department of Economics, Harvard University, Cambridge, Massachusetts 02138, United States of America
- ³² Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden
- ³³ School of Social and Community Medicine, University of Bristol, Bristol BS8 2PR, United Kingdom
- ³⁴ Department of Child and Adolescent Psychiatry, Erasmus Medical Center, 3000 CB Rotterdam, The Netherlands
- ³⁵ Department of Clinical Genetics, VU University Medical Center, 1081 BT Amsterdam, The Netherlands
- ³⁶ Department of Internal Medicine, Erasmus Medical Center, Rotterdam 3000 CA, The Netherlands
- ³⁷ Medical Research Institute, University of Dundee, Dundee DD2 4RB, United Kingdom
- ³⁸ Alzheimer Scotland Dementia Research Centre, The University of Edinburgh, Edinburgh EH8 9JZ, Scotland, United Kingdom

- ³⁹ Medical Research Council Centre for Causal Analyses in Translational Epidemiology , School of Social and Community Medicine, University of Bristol, Bristol BS8 2PR, United Kingdom
- ⁴⁰ King's College London, Medical Research Council Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, London SE5 8AF, United Kingdom.
- ⁴¹ Department of Psychology, The University of Edinburgh, Edinburgh EH8 9JZ, Scotland, United Kingdom
- ⁴² Department of Economics, Stockholm School of Economics, Stockholm 113 83, Sweden
- ⁴³ University of Queensland Diamantina Institute, The University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland 4102, Australia
- ⁴⁴ Department of Economics, Cornell University, Ithaca, New York 14853, United States of America
- ⁴⁵ Center for Experimental Social Science, Department of Economics, New York University, New York, New York 10012, United States of America
- ⁴⁶ Faculty of Economics and Business, University of Amsterdam, Amsterdam 1018 TV, The Netherlands

* These authors contributed equally