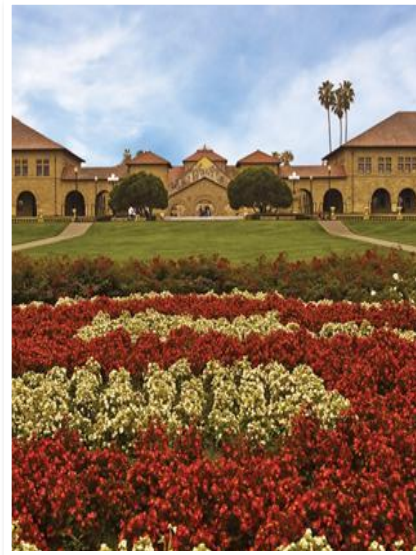


Schools & the Genetics of Educational Attainment: Evidence from Add Health

BEN DOMINGUE



What do we mean by “genetics”?

- Genome-wide data

	SNP 1	SNP 2	...	SNP 1,000,000
P1	0	1	...	2
P2	1	0	...	0
P3	1	2	...	1
⋮	⋮	⋮	⋮	⋮
P1000	2	1	...	2

1,000 × 1,000,000 matrix; each cell $\in \{0, 1, 2\}$.

What do we mean by “genetics”?

- Genome-wide data

	SNP 1	SNP 2	...	SNP 1,000,000
P1	0	1	...	2
P2	1	0	...	0
P3	1	2	...	1
⋮	⋮	⋮	⋮	⋮
P1000	2	1	...	2

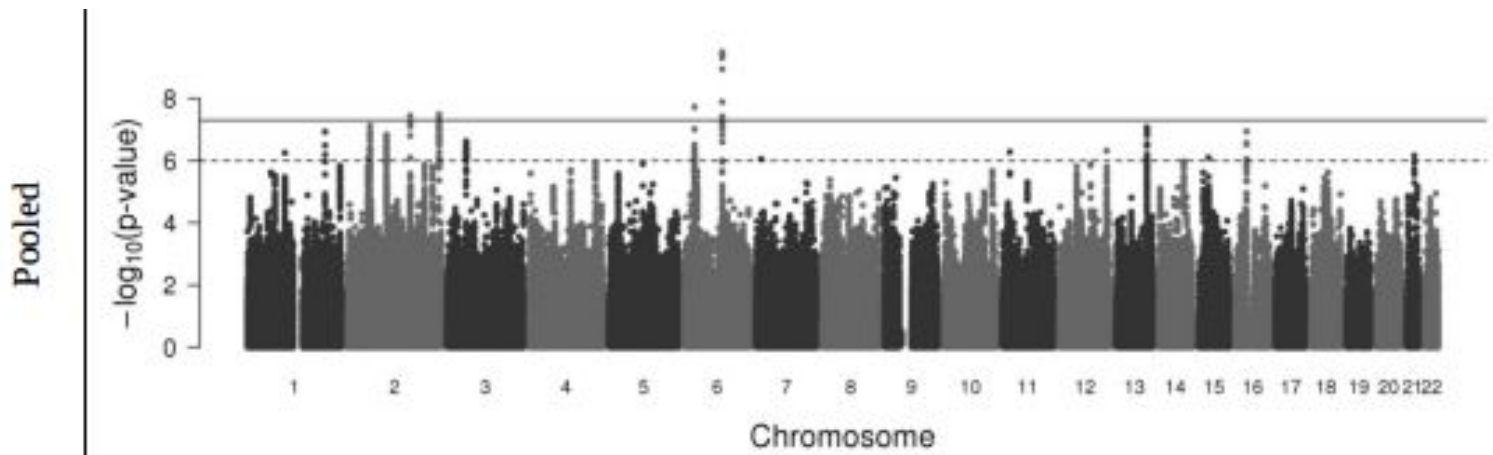
1,000 × 1,000,000 matrix; each cell $\in \{0, 1, 2\}$.

- How do we tie this large set of genetic variants to specific traits?

GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment

Social Science Genetic Association Consortium

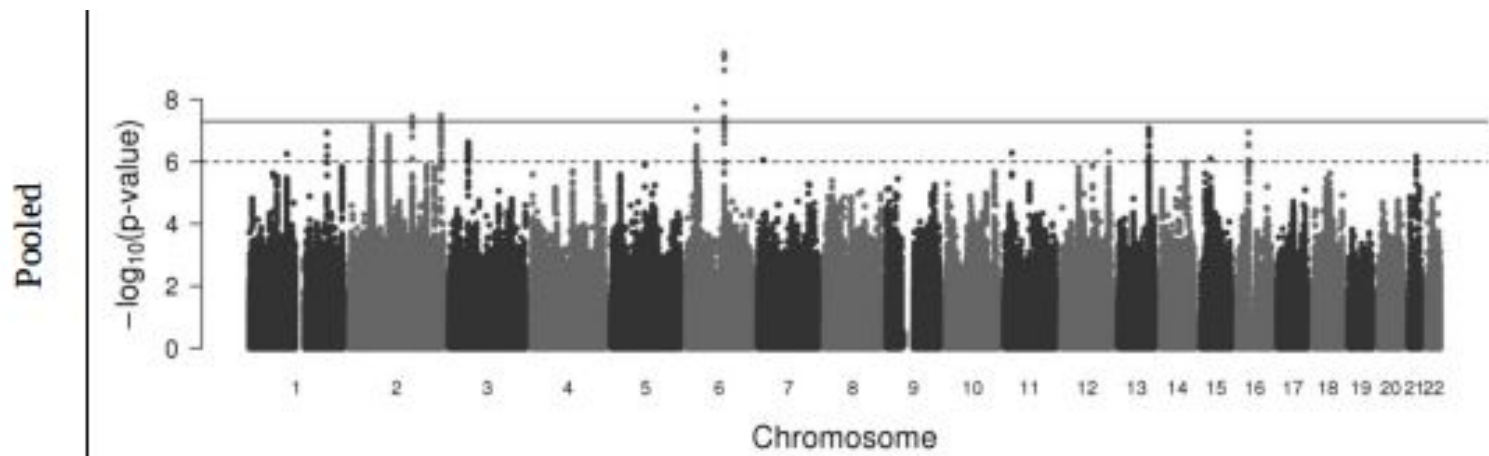
All authors with their affiliations appear at the end of this paper.



GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment

Social Science Genetic Association Consortium

All authors with their affiliations appear at the end of this paper.



- Second generation educational attainment GWAS now available.
- Genome-wide association studies produce information about small associations between millions of variants and outcomes of interest.
 - Can we simplify?

Genome-wide association study identifies 74 loci associated with educational attainment

Aysu Okbay, Jonathan P. Beauchamp, Mark Alan Fontana, James J. Lee, Tune H. Pers, Cornelius A. Rietveld, Patrick Turley, Guo-Bo Chen, Valur Emilsson, S. Fleur W. Meddens, Sven Oskarsson, Joseph K. Pickrell, Kevin Thom, Pascal Timshel, Ronald de Vlaming, Abdel Abdellaoui, Tarunveer S. Ahluwalia, Jonas Bacelis, Clemens Baumbach, Gyda Bjornsdottir, Johannes H. Brandsma, Maria Pina Concas, Jaime Derringer, Nicholas A. Furlotte, Tessel E. Galesloot *et al.*

[Affiliations](#) | [Contributions](#) | [Corresponding authors](#)

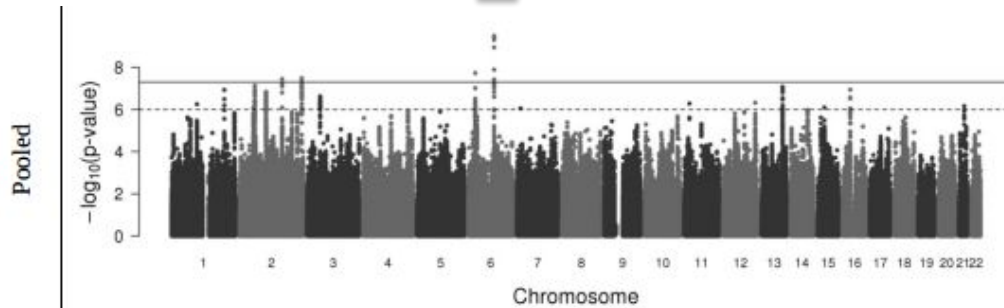
Nature 533, 539–542 (26 May 2016) | doi:10.1038/nature17671

Received 24 June 2015 | Accepted 16 March 2016 | Published online 11 May 2016

Polygenic Score (PGS)

	SNP 1	SNP 2	...	SNP 1,000,000
P1	0	1	...	2
P2	1	0	...	0
P3	1	2	...	1
⋮	⋮	⋮	⋮	⋮
P1000	2	1	...	2

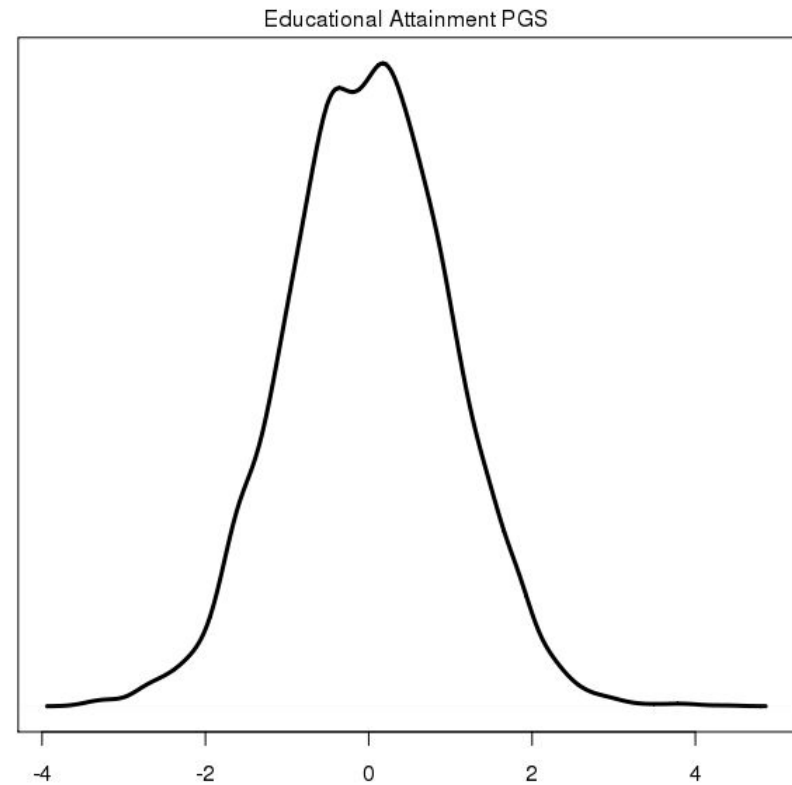
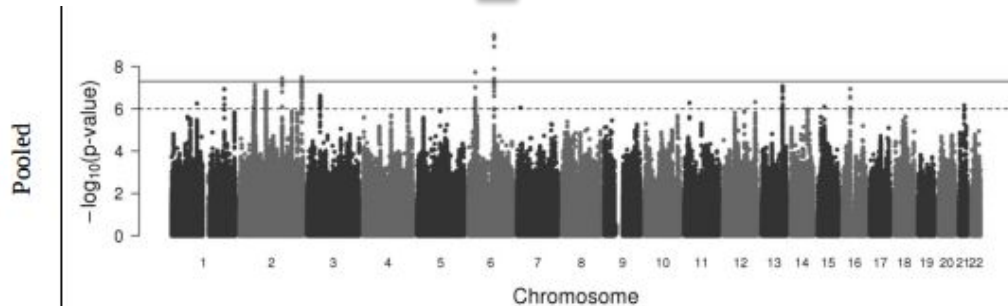
1,000 × 1,000,000 matrix; each cell $\in \{0, 1, 2\}$.



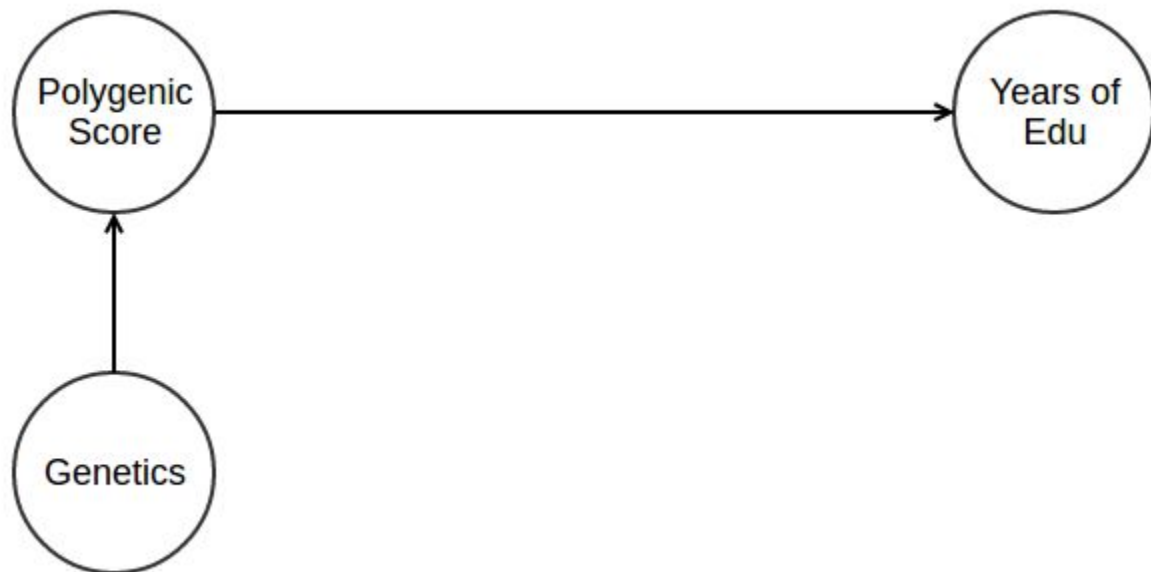
Polygenic Score (PGS)

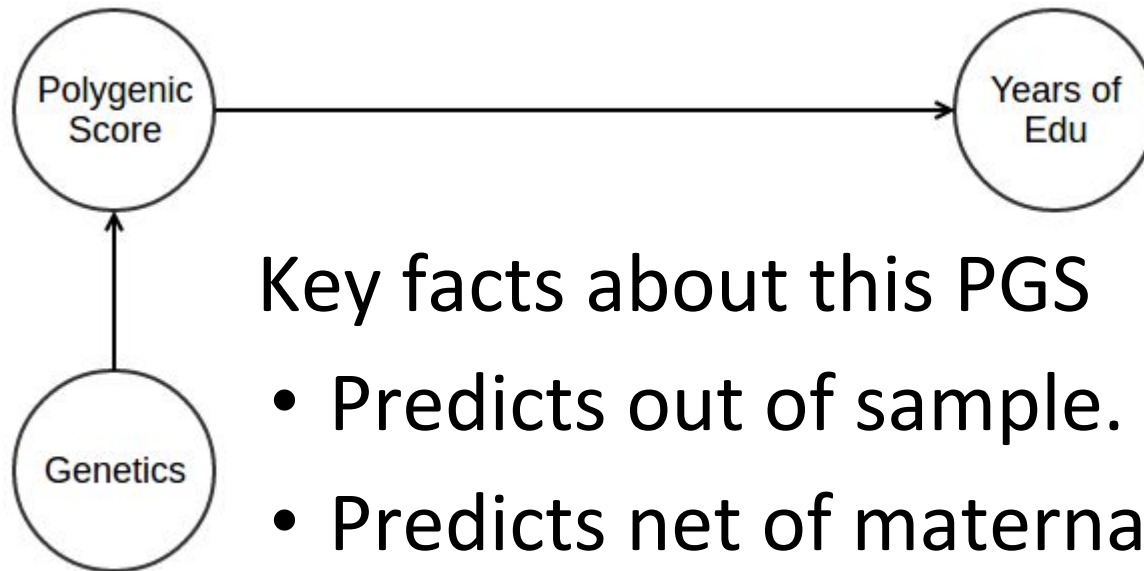
	SNP 1	SNP 2	...	SNP 1,000,000
P1	0	1	...	2
P2	1	0	...	0
P3	1	2	...	1
⋮	⋮	⋮	⋮	⋮
P1000	2	1	...	2

1,000 × 1,000,000 matrix; each cell $\in \{0, 1, 2\}$.



For technical reasons, we focus on non-Hispanic white respondents.





Key facts about this PGS

- Predicts out of sample.
- Predicts net of maternal education & maternal genetics.
- Predicts amongst siblings.

Replicability and Robustness of Genome-Wide-Association Studies for Behavioral Traits

Cornelius A. Rietveld^{1,2}, Dalton Conley³, Nicholas Eriksson⁴, Tõnu Esko⁵, Sarah E. Medland⁶, Anna A. E. Vinkhuyzen⁷, Jian Yang⁷, Jason D. Boardman^{8,9}, Christopher F. Chabris¹⁰, Christopher T. Dawes¹¹, Benjamin W. Domingue⁸, David A. Hinds⁴, Magnus Johannesson¹², Amy K. Kiefer⁴, David Laibson¹³, Patrik K. E. Magnusson¹⁴, Joanna L. Mountain⁴, Sven Oskarsson¹⁵, Olga Rostapshova¹³, Alexander Teumer¹⁶, Joyce Y. Tung⁴, Peter M. Visscher^{7,17}, Daniel J. Benjamin¹⁸, David Cesarini^{19,20}, Philipp D. Koellinger^{1,2,21}, and the Social Science Genetics Association Consortium

Is the Effect of Parental Education on Offspring Biased or Moderated by Genotype?

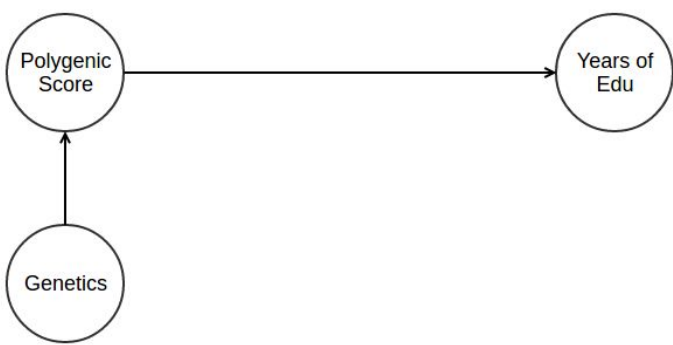
Dalton Conley, Benjamin W. Domingue, David Cesarini, Christopher Dawes, Cornelius A. Rietveld, Jason D. Boardman

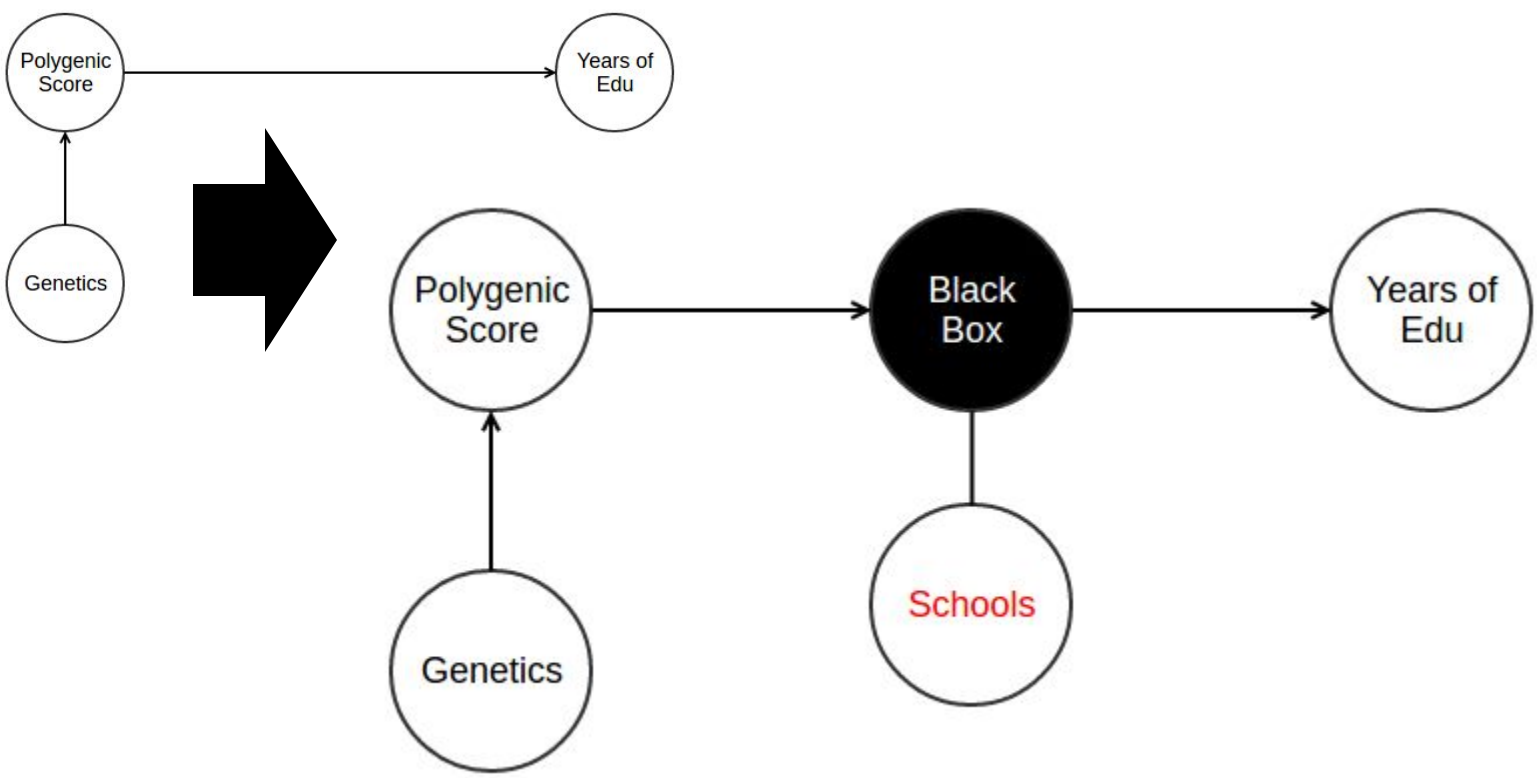
Sociological Science, February 25, 2015
DOI 10.15195/v2.a6

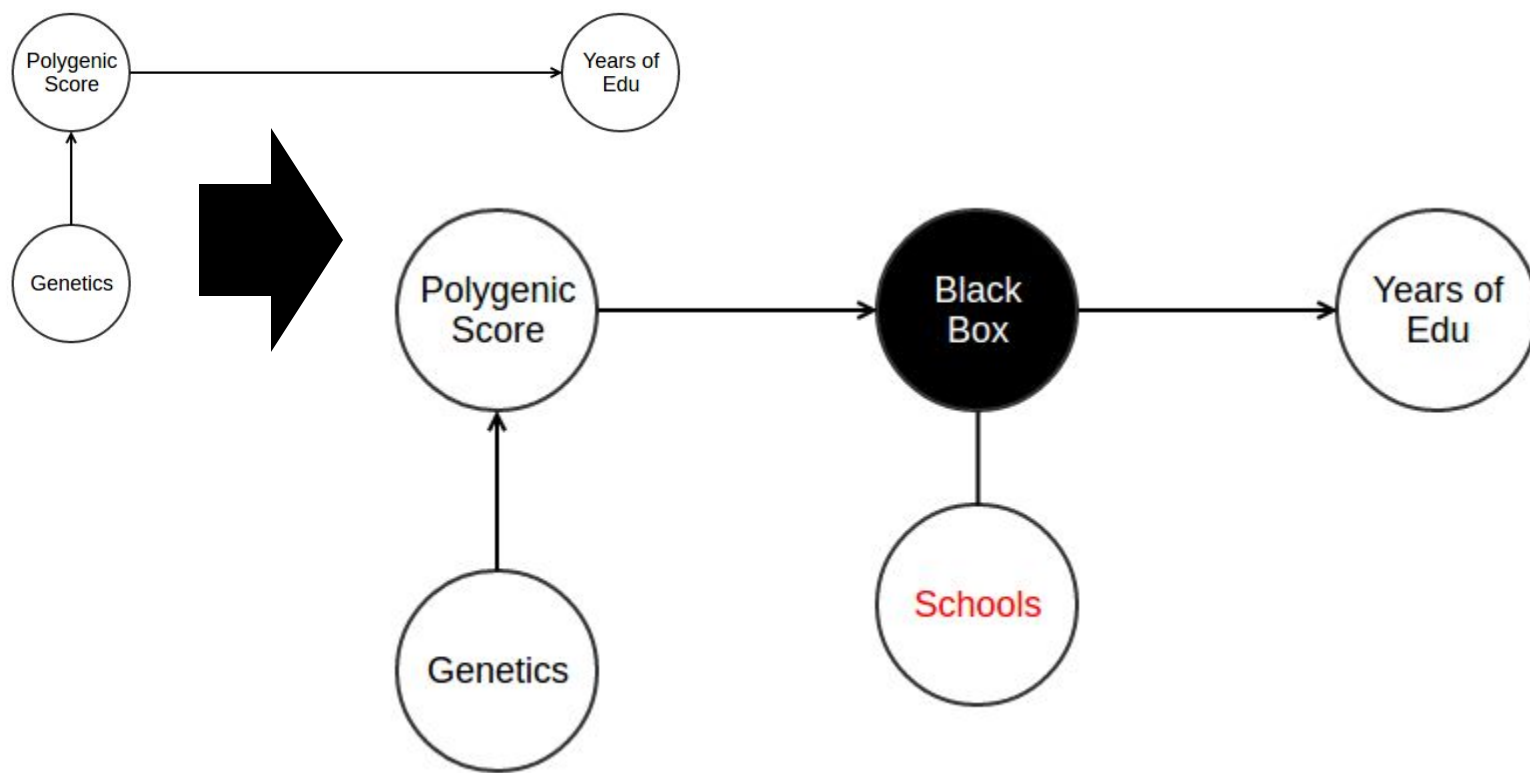
Polygenic Influence on Educational Attainment

New Evidence From the National Longitudinal Study of Adolescent to Adult Health

Benjamin W. Domingue, Daniel W. Belsky, Dalton Conley, Kathleen Mullan Harris, Jason D. Boardman
DOI: 10.1177/2332858415599972, Aug 2015







What is role of schools?

Questions

1. How are relevant genetics distributed across schools?
2. Any indication that schools are moderating genetic effects?



Add Health has been ranked as ICPSR's top downloaded study for two years in a row.



[Learn More about Add Health Public Data Use!](#)

1 2 3 4

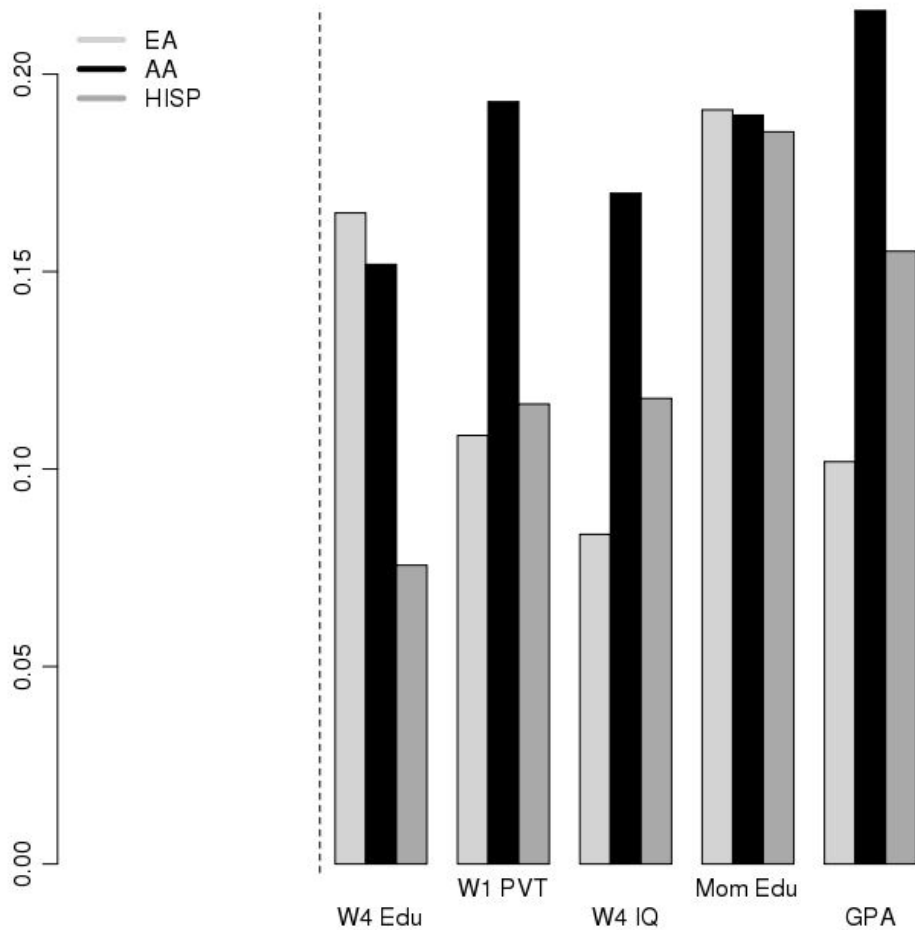
Add Health

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a longitudinal study of a nationally representative sample of adolescents in grades 7-12 in the United States during the 1994-95 school year. The Add Health cohort has been followed into young adulthood with four in-home interviews, the most recent in 2008, when the sample was aged 24-32*. Add Health combines longitudinal survey data on respondents' social, economic, psychological and physical well-being with contextual data on the family, neighborhood, community, school, friendships, peer groups, and romantic relationships, providing unique opportunities to study how social environments and behaviors in adolescence are linked to health and achievement outcomes in young adulthood. The fourth wave of interviews expanded the collection of biological data in Add Health to understand the social, behavioral, and biological linkages in health trajectories as the Add Health cohort ages through adulthood. Add Health plans to trace, locate, and re-interview cohort members in a Wave V follow-up during the period 2016-2018 to collect social, environmental, behavioral, and biological data with which to track the emergence of chronic disease as the cohort moves through their fourth decade of life.

*52 respondents were 33-34 years old at the time of the Wave IV interview.

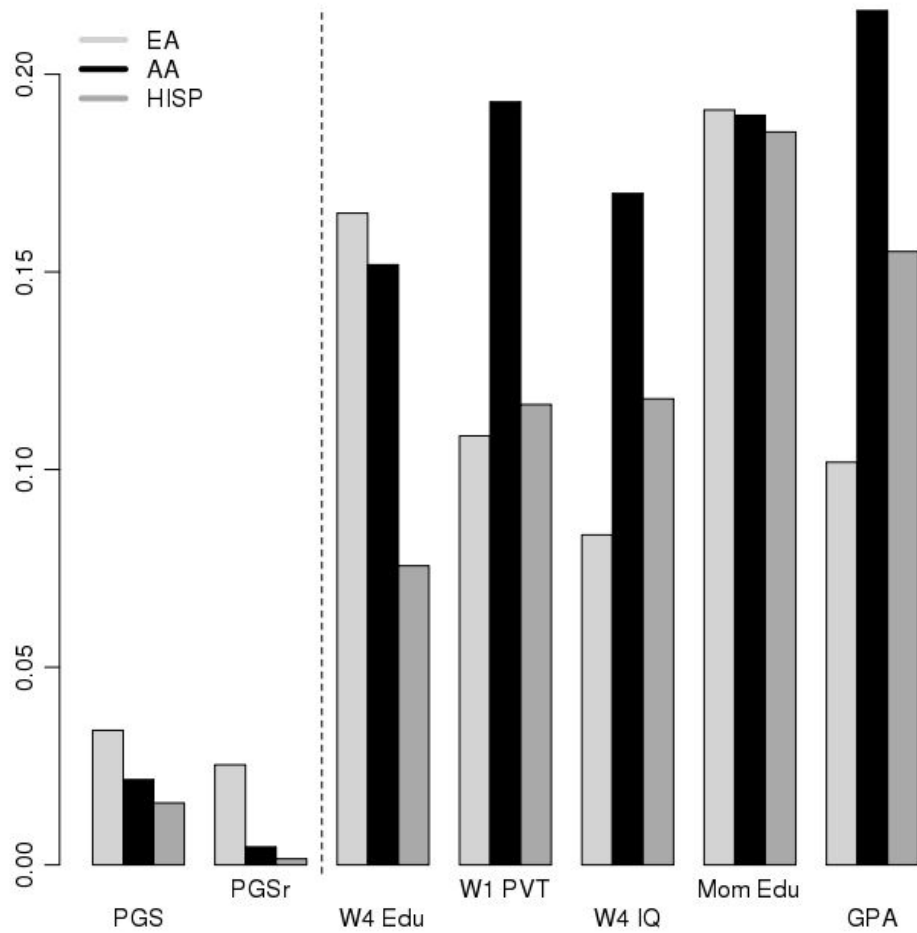
- Probability sample drawn from >100 schools.
- Subsample (siblings) first genotyped.
- Full cohort now genotyped.
 - 10,577 total
 - 6,524 non-Hispanic White

Clustering of PGS into schools

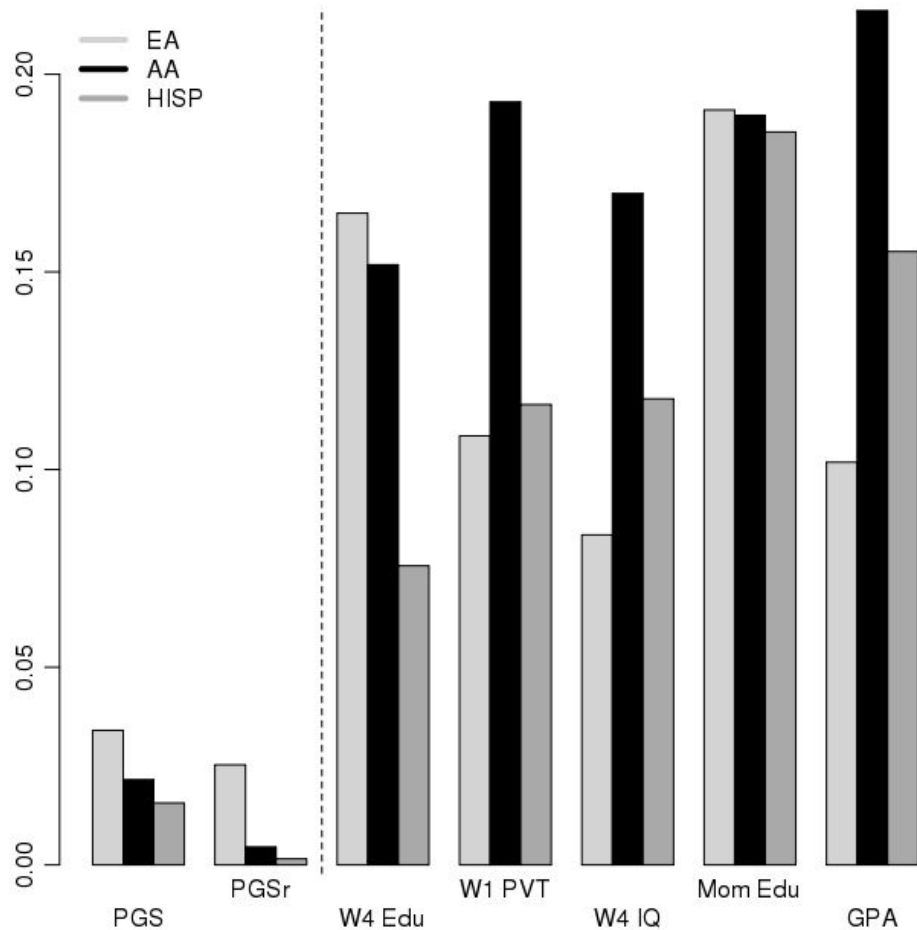


- ICC showing proportion of variation in outcomes that is between schools.
 - e.g., ~20% of the variation in maternal education is between schools.

Clustering of PGS into schools

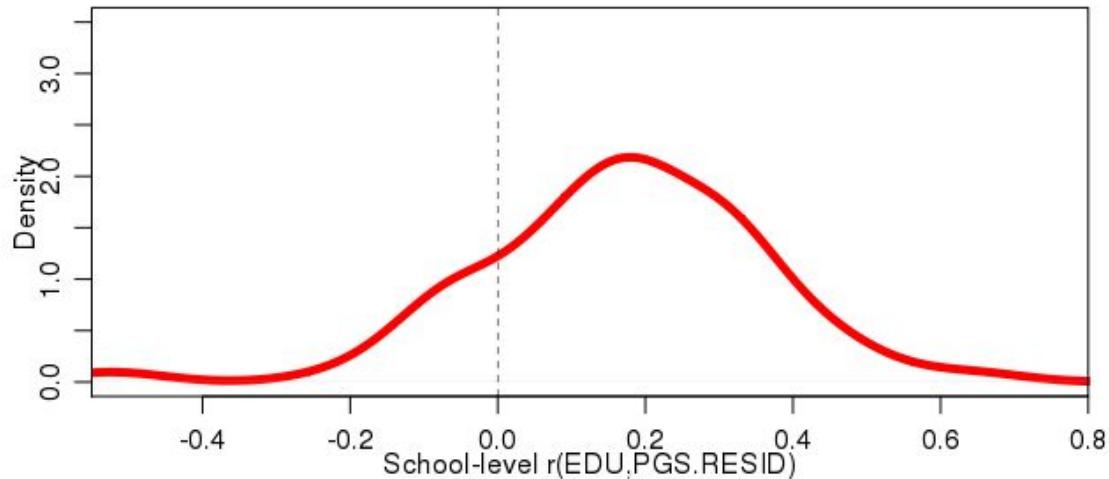


Clustering of PGS into schools



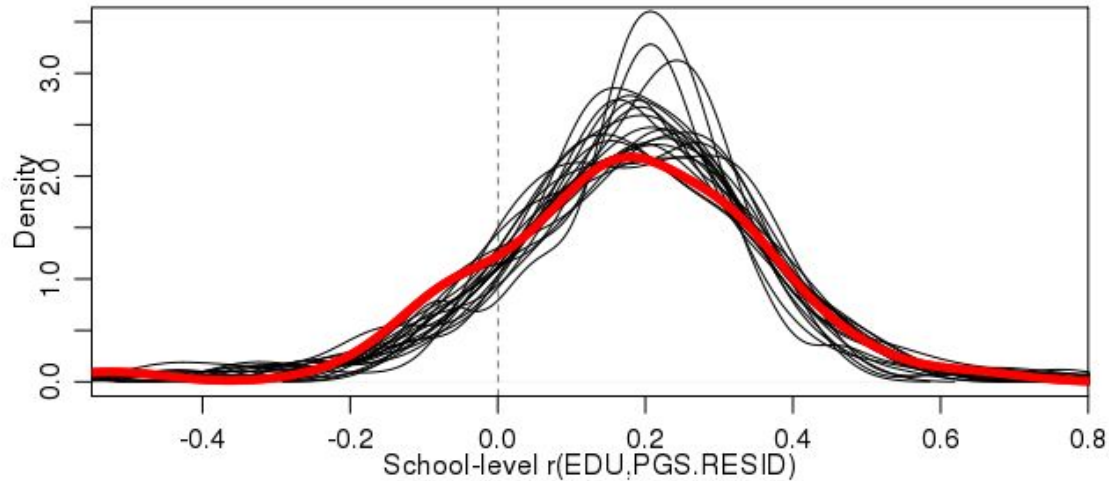
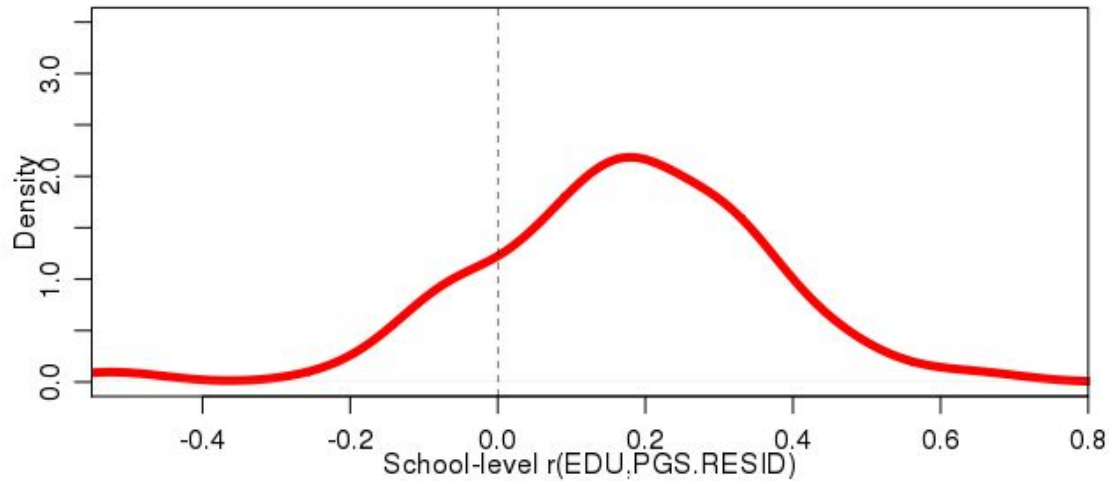
- What is effect of the clustering we do observe?
 - Genetics of EA student's peers are more predictive than their own.

How much variation is there in the PGS effect between schools?



- First computed correlations between educational attainment and PGS for each school.
- Then considered the distribution.

How much variation is there in the PGS effect between schools?



In closing

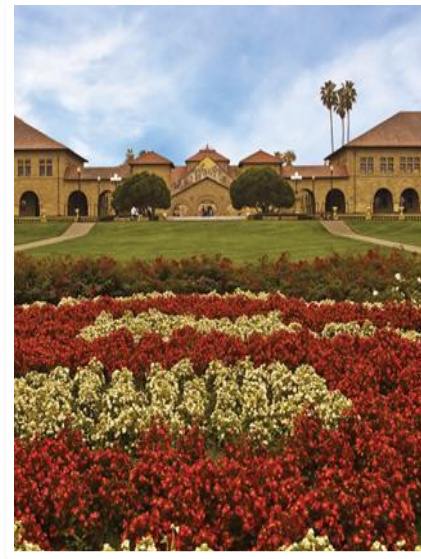
- Preliminary Findings:
 - It seems clear that there is interesting genetic clustering into schools.
 - Little evidence for school-level moderation of genetic effect.

In closing

- Preliminary Findings:
 - It seems clear that there is interesting genetic clustering into schools.
 - Little evidence for school-level moderation of genetic effect.
- Consistent with a model in which genetic risks are not evenly distributed but where there is little gene-environment interaction (GxE).
 - Key role for gene-environment correlation (rGE).
- Also of interest: upstream (developmental pathways) and downstream (later-life outcomes) associations with PGS.

Thanks!

bdomingue@stanford.edu



Collaborators: Dan Belsky, Kathleen Mullan Harris,
Jason Boardman