Sociogenomics: Some recent history and a novel approach for examining GxE

PDHP Workshop
University of Michigan

BEN DOMINGUE
bdomingue@stanford.edu

Stanford cepa | Center for Education Policy Analysis

Stanford MEDICINE | Center for Population Health Sciences

Genetics: A quick disclaimer

- Genetics: A good candidate mechanism for explaining how people who start in similar situations end up different
  - Terrible for explaining why people who start in radically different situations don’t end up the same.

Genetics: A quick disclaimer

- Genetics: A good candidate mechanism for explaining how people who start in similar situations end up different
  - Terrible for explaining why people who start in radically different situations don't end up the same.
- Lots of social science is focused on the structural forces that lead to unequal distribution of opportunity (i.e., radically different starting points). Rightly so.

Genetics: A quick disclaimer

- Genetics: A good candidate mechanism for explaining how people who start in similar situations end up different
  - Terrible for explaining why people who start in radically different situations don’t end up the same.
- Lots of social science is focused on the structural forces that lead to unequal distribution of opportunity (i.e., radically different starting points). Rightly so.
- But, there is more to the story of human behavior.
  - If environments mattered so thoroughly, siblings would be way more similar than they happen to be.

Genetics: A quick disclaimer

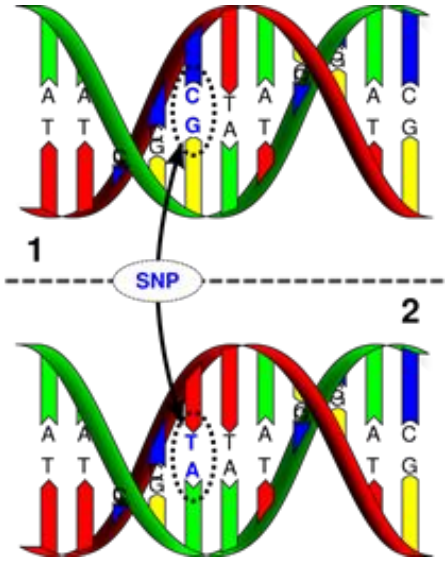
- Genetics: A good candidate mechanism for explaining how people who start in similar situations end up different
  - Terrible for explaining why people who start in radically different situations don’t end up the same.
- Lots of social science is focused on the structural forces that lead to unequal distribution of opportunity (i.e., radically different starting points). Rightly so.
- But, there is more to the story of human behavior.
  - If environments mattered so thoroughly, siblings would be way more similar than they happen to be.
- Understanding why siblings (and otherwise similar people) differ can tell us about developmental cascades and perhaps offer insight into
  - Opportunities for more effective intervention.
  - More refined notions of environmental impacts.

Genetics: A quick disclaimer

- Genetics: A good candidate mechanism for explaining how people who start in similar situations end up different
  - Terrible for explaining why people who start in radically different situations don't end up the same.
- Lots of social science is focused on the structural forces that lead to unequal distribution of opportunity (i.e., radically different starting points). Rightly so.
- But, there is more to the story of human behavior.
  - If environments mattered so thoroughly, siblings would be way more similar than they happen to be.
- Understanding why siblings (and otherwise similar people) differ can tell us about developmental cascades and perhaps offer insight into
  - Opportunities for more effective intervention.
  - More refined notions of environmental impacts.

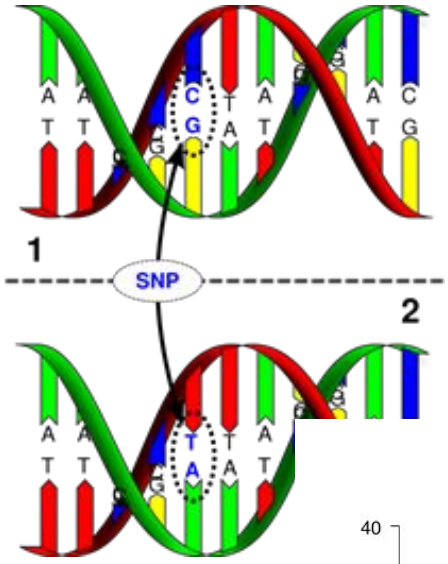
With this caveat in mind, let's talk about a little biology.

SNPs and GWAS



- SNP: single letter change in DNA sequence
- Millions of SNPs in human DNA
- Not “genes” (a gene is a ‘chunk’ of DNA)

SNPs and GWAS



- SNP: single le
  sequence
- Millions of SN
- Not “genes” (a

GWAS (genome-wide association study)

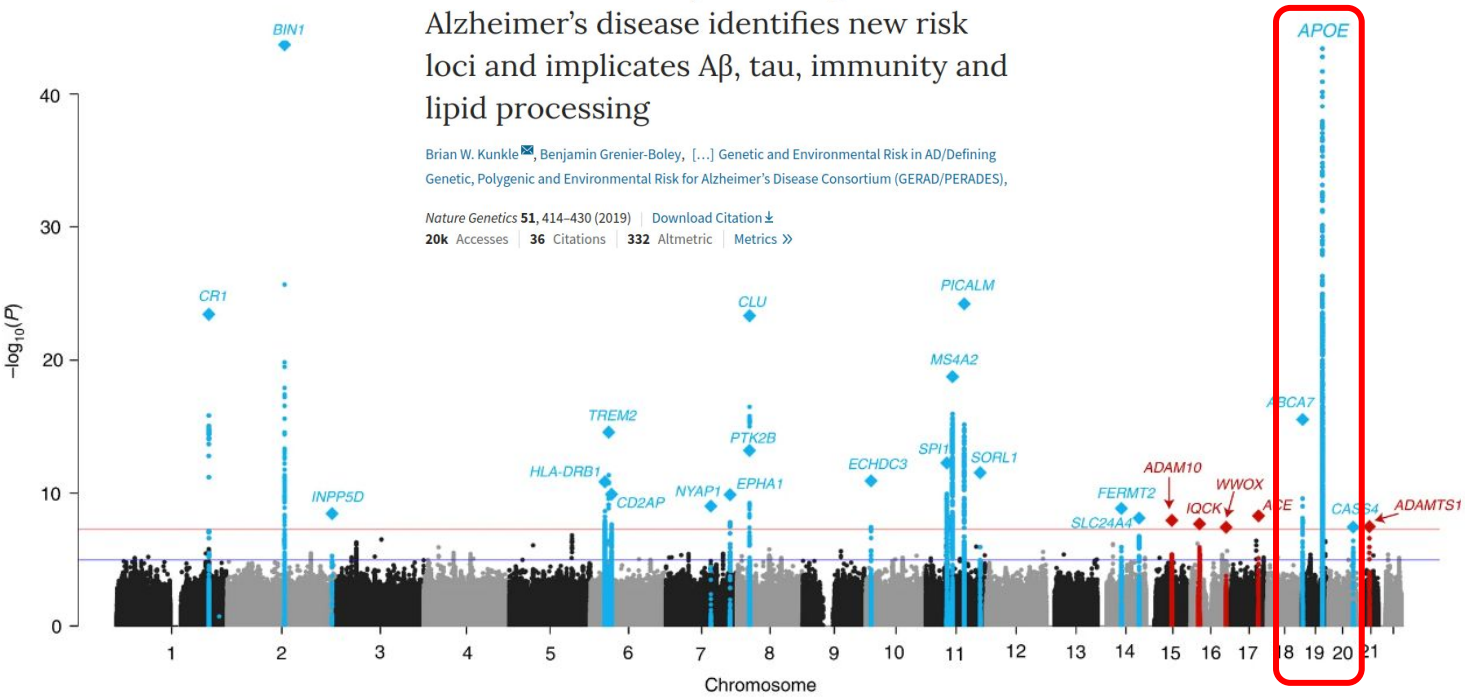
- Brute force means of analyzing associations between SNPs and phenotypes.
- Basic idea: regress phenotype on each SNP
- Can identify APOE like genetic predictors

Genetic meta-analysis of diagnosed Alzheimer’s disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing

Brian W. Kunkle, Benjamin Grenier-Boley, […] Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer’s Disease Consortium (GERAD/PERADES),

Nature Genetics **51**, 414–430 (2019) | Download Citation

20k Accesses | **36** Citations | **332** Altmetric | Metrics »



GWAS has been a huge success

- GWAS have been performed for many biomedical outcomes.
- Those findings are having impact in a variety of disciplines.
- But, why stop at biomedical outcomes?

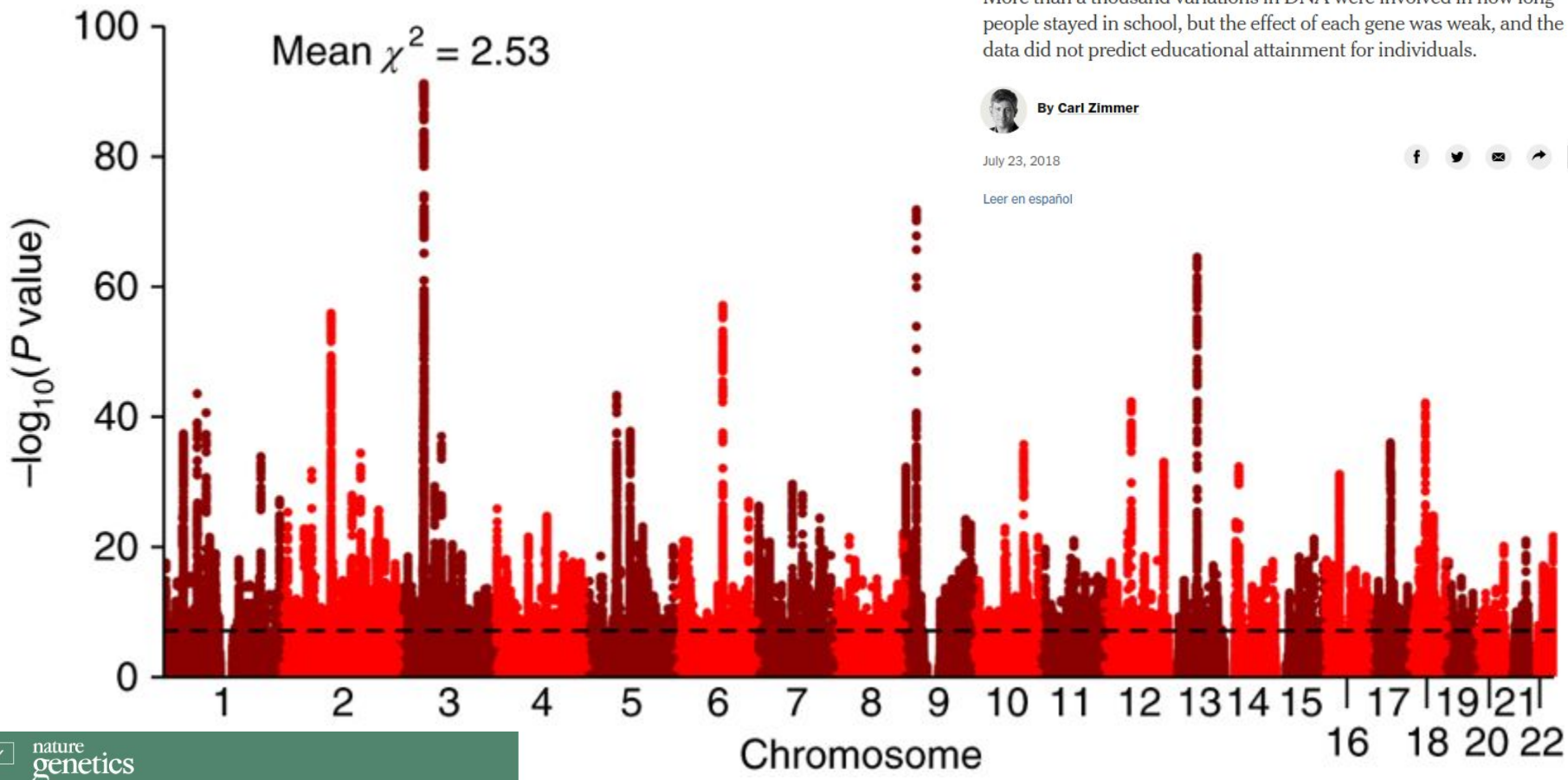
Am J Hum Genet. 2017 Jul 6; 101(1): 5–22. PMCID: PMC5501872
Published online 2017 Jul 6. doi: 10.1016/j.ajhg.2017.06.005 PMID: 28686856

10 Years of GWAS Discovery: Biology, Function, and Translation

Peter M. Visscher,[1,2,*] Naomi R. Wray,[1,2] Qian Zhang,[1] Pamela Sklar,[3] Mark I. McCarthy,[4,5,6] Matthew A. Brown,[7] and Jian Yang[1,2]

► Author information ► Copyright and License information Disclaimer

GWAS of educational attainment



nature genetics

Article | Published: 23 July 2018

Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals

James J. Lee, Robbee Wedow, [...] David Cesarini

The New York Times

MATTER

Years of Education Influenced by Genetic Makeup, Enormous Study Finds

More than a thousand variations in DNA were involved in how long people stayed in school, but the effect of each gene was weak, and the data did not predict educational attainment for individuals.

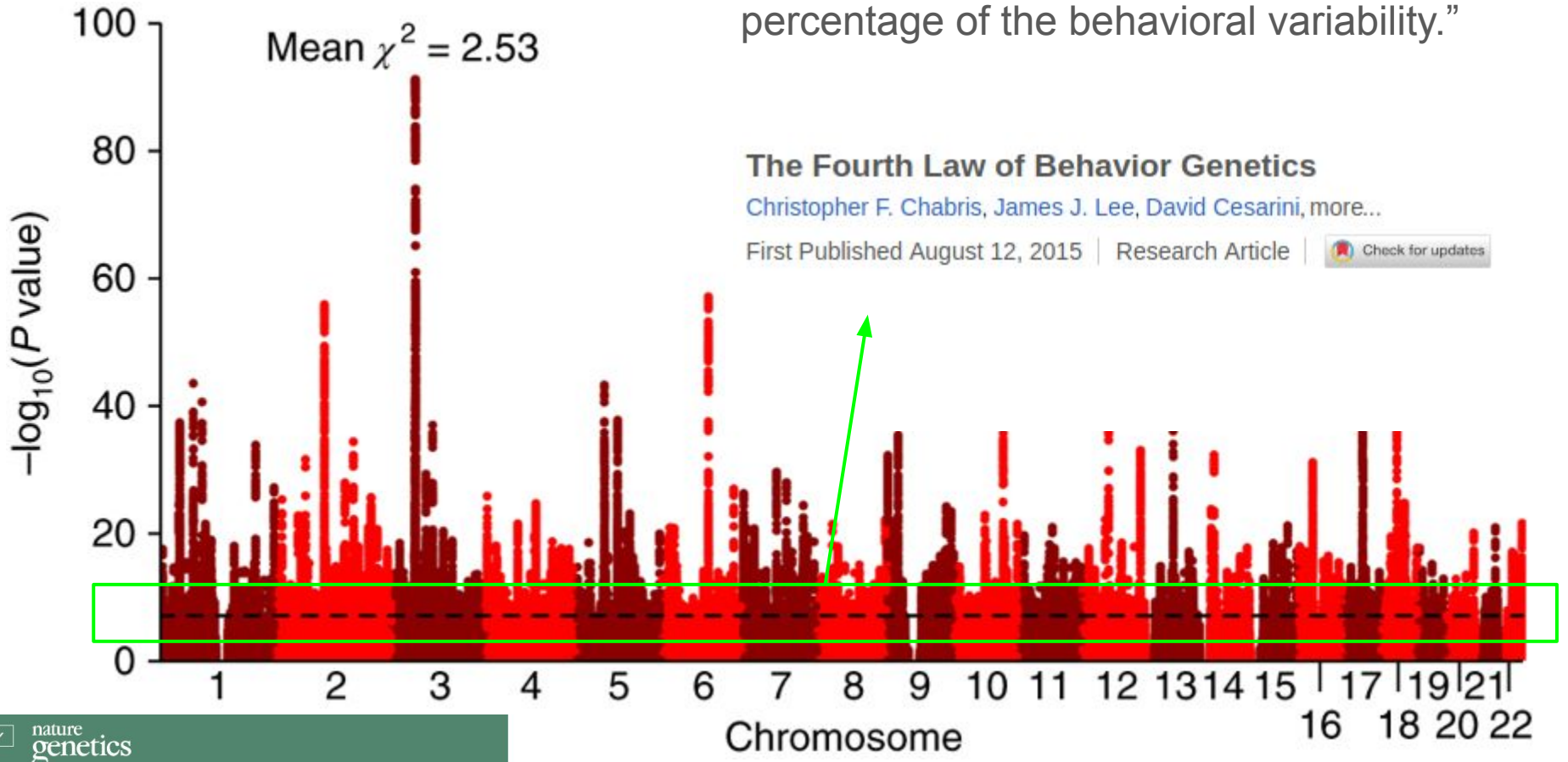
By Carl Zimmer

July 23, 2018

Leer en español

GWAS of educational attainment

"A typical human behavioral trait is associated with very many genetic variants, each of which accounts for a very small percentage of the behavioral variability."



The Fourth Law of Behavior Genetics

Christopher F. Chabris, James J. Lee, David Cesarini, more...

First Published August 12, 2015 | Research Article | Check for updates

nature genetics

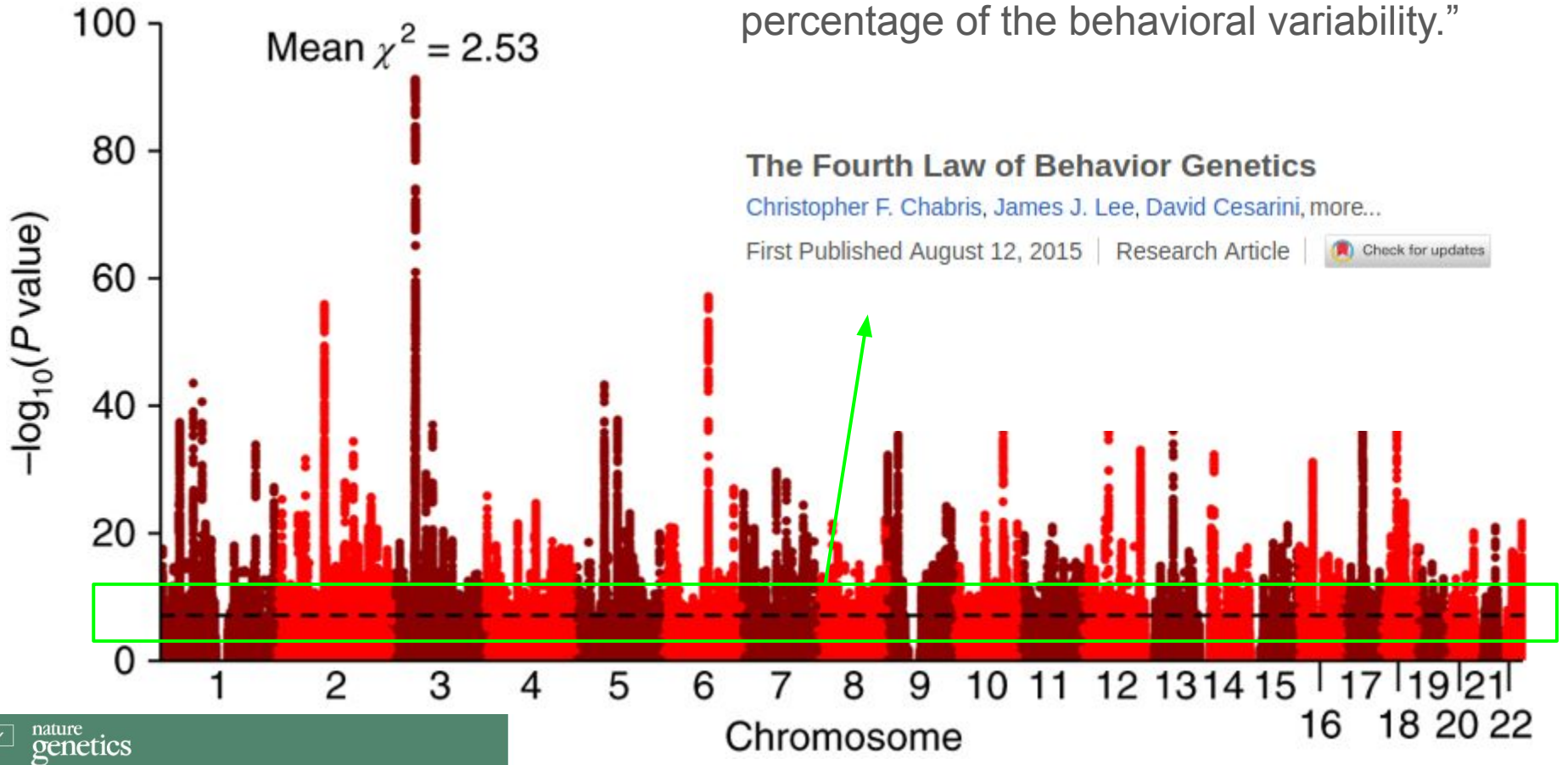
Article | Published: 23 July 2018

Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals

James J. Lee, Robbee Wedow, [...] David Cesarini

GWAS of educational attainment

"A typical human behavioral trait is associated with very many genetic variants, each of which accounts for a very small percentage of the behavioral variability."



The Fourth Law of Behavior Genetics

Christopher F. Chabris, James J. Lee, David Cesarini, more...

First Published August 12, 2015 | Research Article | Check for updates

nature genetics

Article | Published: 23 July 2018

Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals

James J. Lee, Robbee Wedow, [...] David Cesarini

- No APOE-like predictor here. Very diffuse signal.
- How are we going to incorporate all of this into our models?

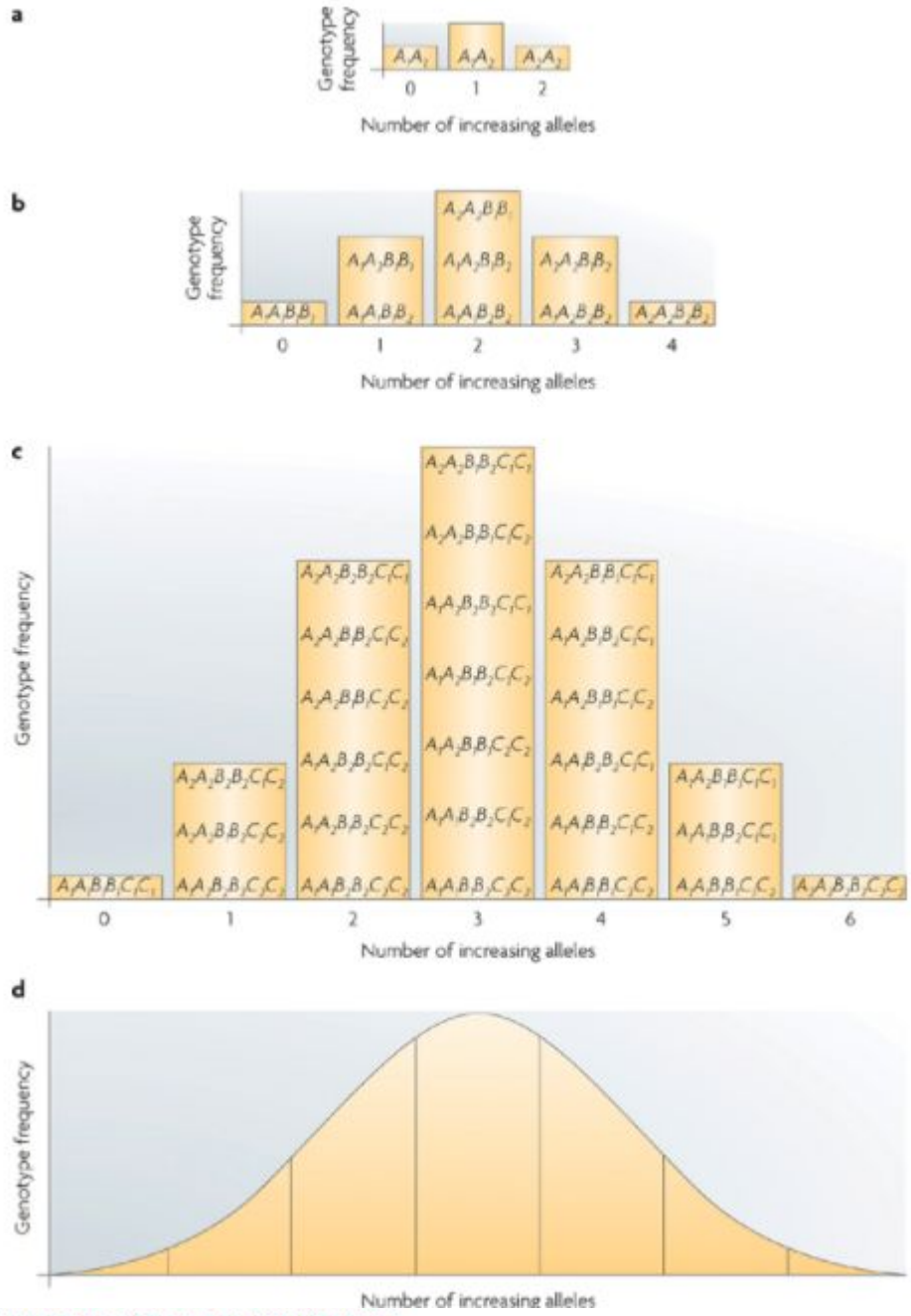
Polygenic scores

A genetic predictor

$$\sum_i \beta_i \cdot \#\,\text{Alleles}_i$$

Based on specific GWAS for a trait

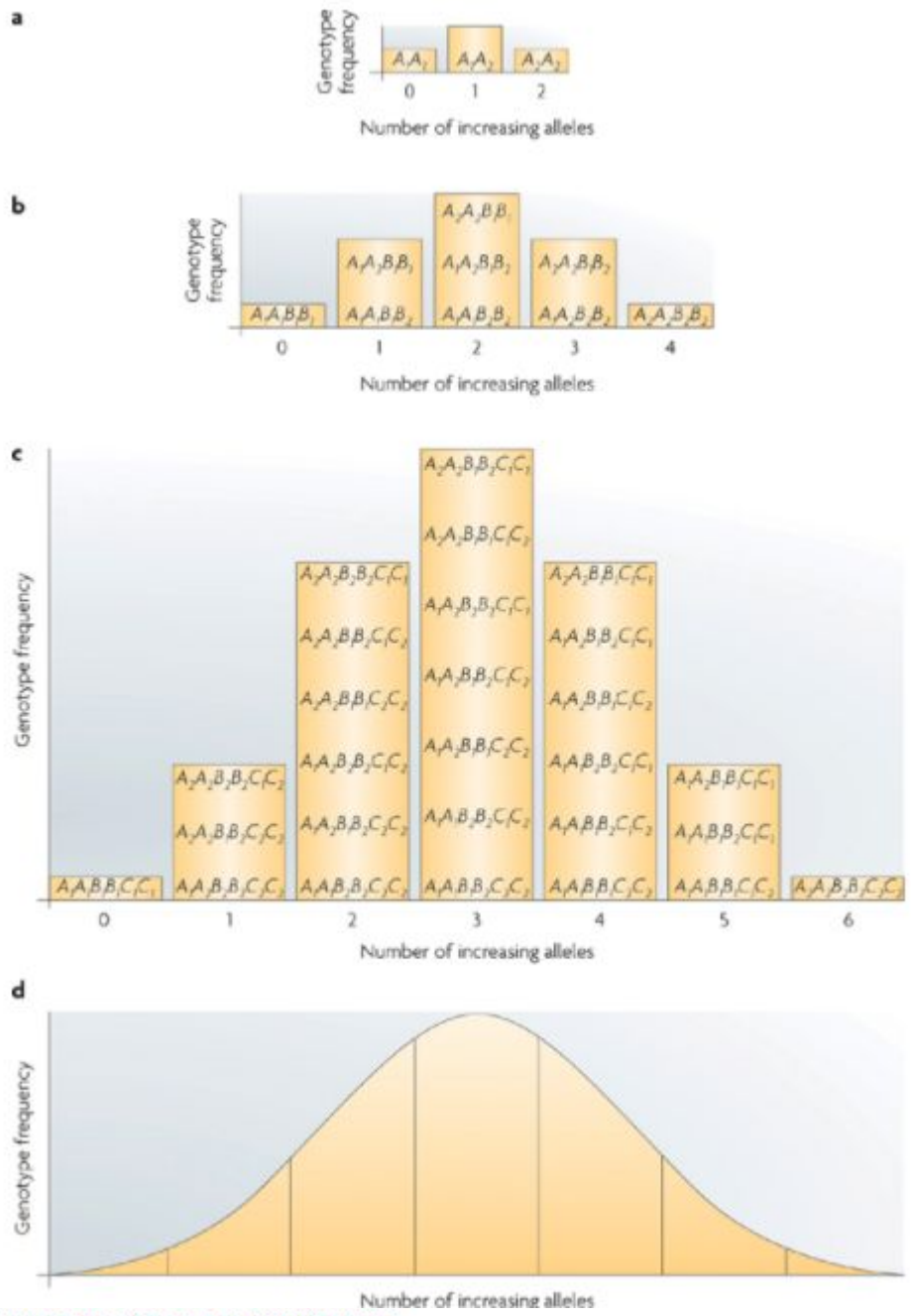
Based on person's genotype



Common disorders are quantitative traits

Robert Plomin, Claire M. A. Haworth & Oliver S. P. Davis

Nature Reviews Genetics **10**, 872-878 (December 2009)



Common disorders are quantitative traits

Robert Plomin, Claire M. A. Haworth & Oliver S. P. Davis

Nature Reviews Genetics **10**, 872-878 (December 2009)

Polygenic scores

A genetic predictor

$$\sum_i \beta_i \cdot \#\,\text{Alleles}_i$$

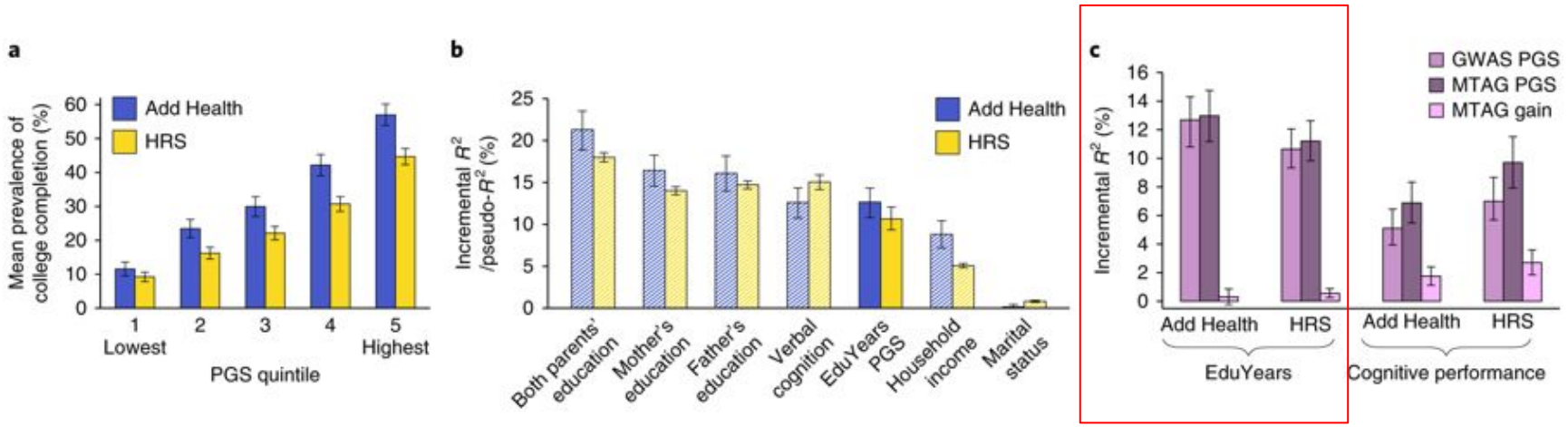
Based on specific GWAS for a trait

Based on person's genotype

Think credit score

But, derived from saliva

Predictive Power



Polygenic score for educational attainment predicts ~10% of variation in years of education.

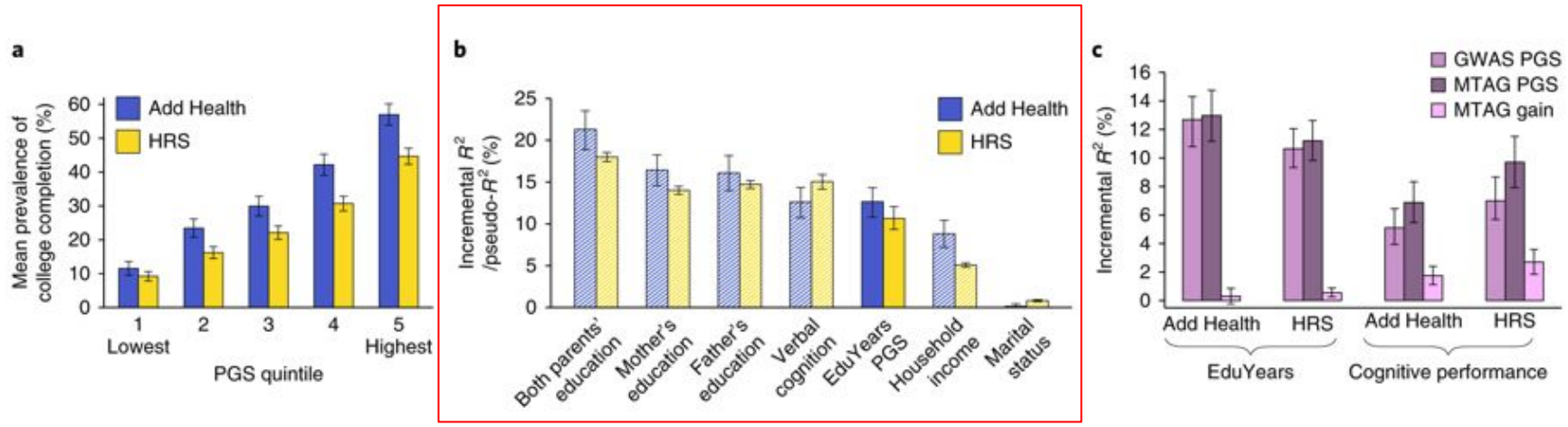
nature genetics

Article | Published: 23 July 2018

Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals

James J. Lee, Robbee Wedow, [...] David Cesarini

Predictive Power



Predictive power of polygenic score for educational attainment predicts out-of-sample as well as:

- Parental education
- Individual verbal ability

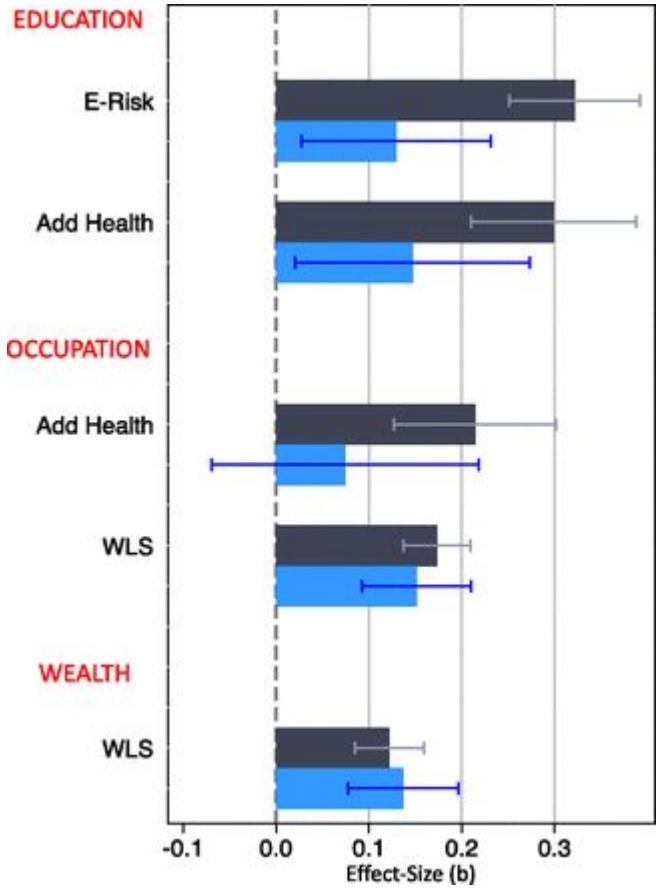
nature genetics

Article | Published: 23 July 2018

Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals

James J. Lee, Robbee Wedow, [...] David Cesarini

Crucially, also predict sibling differences



- The reasons that two siblings have different levels of education tend to be quite different as compared to two unrelated individuals.
- Rules out confounding forces that don't vary within-family.

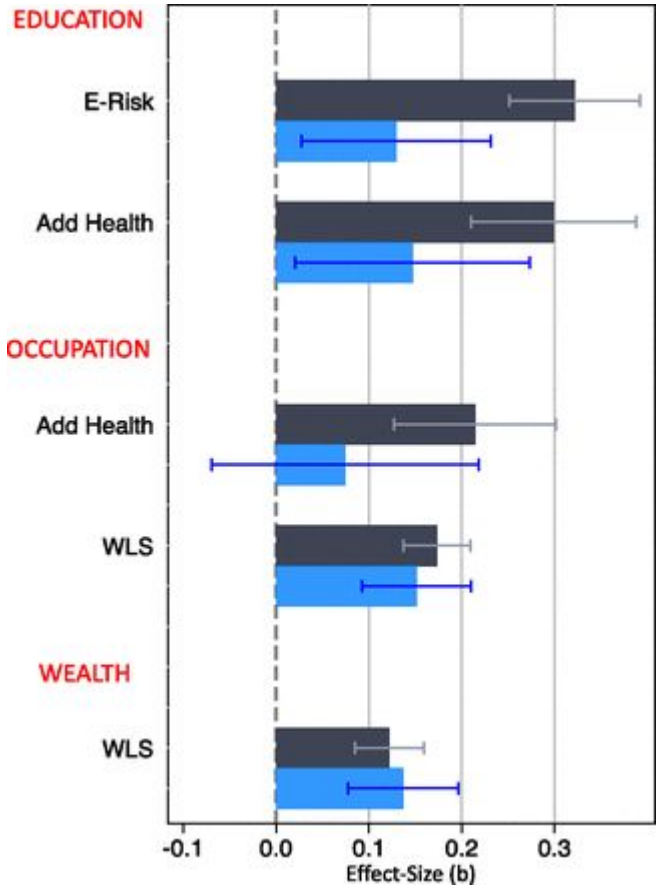
RESEARCH ARTICLE

Genetic analysis of social-class mobility in five longitudinal studies

Daniel W. Belsky, Benjamin W. Domingue, Robbee Wedow, Louise Arseneault, Jason D. Boardman, Avshalom Caspi, Dalton Conley, Jason M. Fletcher, Jeremy Freese, Pamela Herd, Terrie E. Moffitt, Richie Poulton, Kamil Sicinski, Jasmin Wertz, and Kathleen Mullan Harris

PNAS July 31, 2018 115 (31) E7275-E7284; first published July 9, 2018;
https://doi.org/10.1073/pnas.1801238115

Crucially, also predict sibling differences



RESEARCH ARTICLE

Genetic analysis of social-class mobility in five longitudinal studies

Daniel W. Belsky, Benjamin W. Domingue, Robbee Wedow, Louise Arseneault, Jason D. Boardman, Avshalom Caspi, Dalton Conley, Jason M. Fletcher, Jeremy Freese, Pamela Herd, Terrie E. Moffitt, Richie Poulton, Kamil Sicinski, Jasmin Wertz, and Kathleen Mullan Harris

PNAS July 31, 2018 115 (31) E7275-E7284; first published July 9, 2018;
https://doi.org/10.1073/pnas.1801238115

- The reasons that two siblings have different levels of education tend to be quite different as compared to two unrelated individuals.
- Rules out confounding forces that don't vary within-family.
- Light blue bars show within-family models.
  - Significant prediction but clear attenuation.
  - Related to "genetic nurture" issue that we will discuss.

Now what?

- Given this new predictor, what shall we do with it?
- An important question given that lots of studies are adding such genetic predictors.
  - Relatively cheap to add
  - Can be used to predict all kinds of traits.
- Being incorporated into clinic

Add Health
The National Longitudinal Study of Adolescent to Adult Health

HRS | HEALTH AND RETIREMENT STUDY

WISCONSIN LONGITUDINAL STUDY

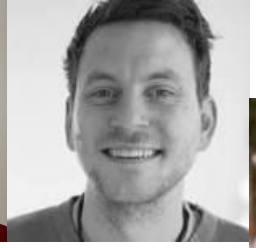
PGS Catalog

Pragmatic randomized trial of polygenic risk scoring for common diseases in primary care

Vassy, Jason L.

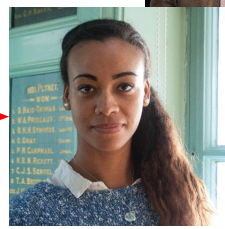
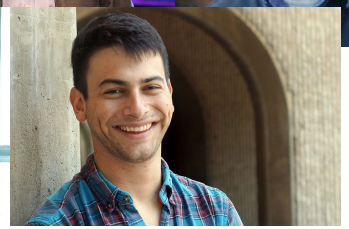
Harvard Medical School, Boston, MA, United States

- Sociology, Genetics, and the Coming of Age of Sociogenomics
- The Arrival of Social Science Genomics
- “Reports of My Death Were Greatly Exaggerated”: Behavior Genetics in the Postgenomic Era
- Phenotypic Annotation: Using Polygenic Scores to Translate Discoveries From Genome-Wide Association Studies From the Top Down
- Using genetics for social science
- The Genome Factor
- Genetics and Education: Recent Developments in the Context of an Ugly History and an Uncertain Future

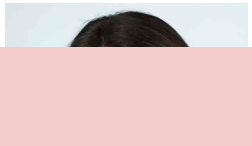




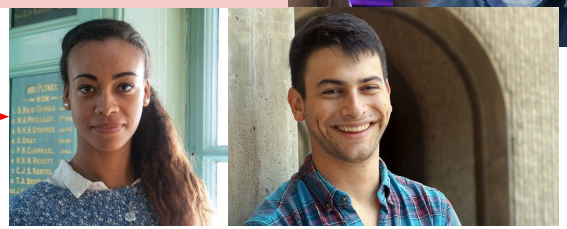




- Sociology, Genetics, and the Coming of Age of Sociogenomics
- The Arrival of Social Science Genomics
- "Reports of My [illegible] Exaggerated": B[illegible] the Postgenomi[illegible]
- Phenotypic Ann[illegible] Polygenic Score[illegible] Discoveries Fro[illegible] Association Stu[illegible] Down
- Using genetics f[illegible]
- The Genome Fa[illegible]
- Genetics and Education: Recent Developments in the Context of an Ugly History and an Uncertain Future









At this point, I see:

- Opportunities for genetics to throw novel light on old problems.
- Reasons for caution in application/interpretation of genetic findings.

Will talk about both

[with overemphasis on work in which I was involved]

Illumination: Genetic associations may help us to interpret social change given that genes are static

- Genetics that predict social status following fall of communism in Estonia are unpredictive in Soviet era. [I was not involved]

Genetic influence on social outcomes during and after the Soviet era in Estonia

Kaili Rimfeld, Eva Krapohl, Maciej Trzaskowski, Jonathan R. I. Coleman, Saskia Selzam, Philip S. Dale, Tonu Esko, Andres Metspalu & Robert Plomin

Nature Human Behaviour **2**, 269–275(2018) | Cite this article

Illumination: Genetic associations may help us to interpret social change given that genes are static

- Genetics that predict social status following fall of communism in Estonia are unpredictive in Soviet era.
- As opportunity structures change across 20th century, we see shifting patterns of association between genetics and educational outcomes across gender.

Genetic influence on social outcomes during and after the Soviet era in Estonia

Kaili Rimfeld, Eva Krapohl, Maciej Trzaskowski, Jonathan R. I. Coleman, Saskia Selzam, Philip S. Dale, Tonu Esko, Andres Metspalu & Robert Plomin

Nature Human Behaviour **2**, 269–275(2018) | Cite this article

Genes, Gender Inequality, and Educational Attainment

Pamela Herd, Jeremy Freese, Kamil Sicinski, more...

First Published November 22, 2019 | Research Article | Check for updates

https://doi.org/10.1177/0003122419886550

Article information

Illumination: Genetic associations may help us to interpret social change given that genes are static

- Genetics that predict social status following fall of communism in Estonia are unpredictive in Soviet era.
- As opportunity structures change across 20th century, we see shifting patterns of association between genetics and educational outcomes across gender.
- In 20th century, genetic predictor of education is weakly associated with smoking whereas, by end, it is highly predictive of smoking.

Genetic influence on social outcomes during and after the Soviet era in Estonia

Kaili Rimfeld, Eva Krapohl, Maciej Trzaskowski, Jonathan R. I. Coleman, Saskia Selzam, Philip S. Dale, Tonu Esko, Andres Metspalu & Robert Plomin

Nature Human Behaviour **2**, 269–275(2018) | Cite this article

Genes, Gender Inequality, and Educational Attainment

Pamela Herd, Jeremy Freese, Kamil Sicinski, more...

First Published November 22, 2019 | Research Article | Check for updates

https://doi.org/10.1177/0003122419886550

Article information

Education, Smoking, and Cohort Change: Forwarding a Multidimensional Theory of the Environmental Moderation of Genetic Effects

Robbee Wedow, Meghan Zacher, Brooke M. Huibregtse, more... Show all authors

First Published July 20, 2018 | Research Article | Check for updates

https://doi.org/10.1177/0003122418785368

Article information

Altmetric 26

Illumination: Genetic associations may help us to interpret social change given that genes are static

- Genetics that predict social status following fall of communism in Estonia are unpredictive in Soviet era.
- As opportunity structures change across 20th century, we see shifting patterns of association between genetics and educational outcomes across gender.
- In 20th century, genetic predictor of education is weakly associated with smoking whereas, by end, it is highly predictive of smoking.
- Holding genotype constant, higher-status schools keep high school kids in math courses longer.

Genetic influence on social outcomes during and after the Soviet era in Estonia

Kaili Rimfeld, Eva Krapohl, Maciej Trzaskowski, Jonathan R. I. Coleman, Saskia Selzam, Philip S. Dale, Tonu Esko, Andres Metspalu & Robert Plomin

Nature Human Behaviour **2**, 269–275(2018) | Cite this article

Genes, Gender Inequality, and Educational Attainment

Pamela Herd, Jeremy Freese, Kamil Sicinski, more...

First Published November 22, 2019 | Research Article | Check for updates

https://doi.org/10.1177/0003122419886550

Article information

Education, Smoking, and Cohort Change: Forwarding a Multidimensional Theory of the Environmental Moderation of Genetic Effects

Robbee Wedow, Meghan Zacher, Brooke M. Huibregtse, more... Show all authors

First Published July 20, 2018 | Research Article | Check for updates

https://doi.org/10.1177/0003122418785368

Article information | Altmetric 26

Article | Open Access | Published: 05 February 2020

Genetic associations with mathematics tracking and persistence in secondary school

K. Paige Harden, Benjamin W. Domingue, Daniel W. Belsky, Jason D. Boardman, Robert Crosnoe, Margherita Malanchini, Michel Nivard, Elliot M. Tucker-Drob & Kathleen Mullan Harris

npj Science of Learning **5**, Article number: 1 (2020) | Cite this article

3029 Accesses | **4** Citations | **86** Altmetric | Metrics

Illumination: Genetic associations may help us to interpret social change given that genes are static

- Genetics that predict social status following fall of communism in Estonia are unpredictive in Soviet era.
- As opportunity structures change across 20th century, we see shifting patterns of association between genetics and educational outcomes across gender.
- In 20th century, genetic predictor of education is weakly associated with smoking whereas, by end, it is highly predictive of smoking.
- Holding genotype constant, higher-status schools keep high school kids in math courses longer.
- Those at highest risk for depression see largest jump in depressive symptoms following spousal death.

Genetic influence on social outcomes during and after the Soviet era in Estonia

Kaili Rimfeld, Eva Krapohl, Maciej Trzaskowski, Jonathan R. I. Coleman, Saskia Selzam, Philip S. Dale, Tonu Esko, Andres Metspalu & Robert Plomin

Nature Human Behaviour **2**, 269–275(2018) | Cite this article

Genes, Gender Inequality, and Educational Attainment

Pamela Herd, Jeremy Freese, Kamil Sicinski, more...

First Published November 22, 2019 | Research Article | Check for updates

https://doi.org/10.1177/0003122419886550

Article information

Education, Smoking, and Cohort Change: Forwarding a Multidimensional Theory of the Environmental Moderation of Genetic Effects

Robbee Wedow, Meghan Zacher, Brooke M. Huibregtse, more... Show all authors

First Published July 20, 2018 | Research Article | Check for updates

https://doi.org/10.1177/0003122418785368

Article information

Article | Open Access | Published: 05 February 2020

Genetic associations with mathematics tracking and persistence in secondary school

K. Paige Harden, Benjamin W. Domingue, Daniel W. Belsky, Jason D. Boardman, Robert Crosnoe, Margherita Malanchini, Michel Nivard, Elliot M. Tucker-Drob & Kathleen Mullan Harris

npj Science of Learning **5**, Article number: 1 (2020) | Cite this article

3029 Accesses | **4** Citations | **86** Altmetric | Metrics

> Am J Psychiatry. 2017 Oct 1;174(10):963-970. doi: 10.1176/appi.ajp.2017.16111209. Epub 2017 Mar 24.

Genetic Heterogeneity in Depressive Symptoms Following the Death of a Spouse: Polygenic Score Analysis of the U.S. Health and Retirement Study

Benjamin W Domingue [1], Hexuan Liu [1], Aysu Okbay [1], Daniel W Belsky [1]

In summary

- These studies illustrate ways (at micro- and macro-scales) in which we may use genetics as a fixed point from which to understand external changes and their interactions.
- There are caveats internal to these studies, but of fundamental import to this potential of genetics is the fact that genetics are “inherently prospective” (fixed at conception, unchanging).
- That’s a big advantage!

In summary

- These studies illustrate ways (at micro- and macro-scales) in which we may use genetics as a fixed point from which to understand external changes and their interactions.
- There are caveats internal to these studies, but of fundamental import to this potential of genetics is the fact that genetics are "inherently prospective" (fixed at conception, unchanging).
- That's a big advantage!

BUT, I'm going to now turn to caveats. I'll focus on one big one (what if our predictor is confounded?) but I also want to acknowledge a second major issue (and implications):

1. Aren't yet ready to be deployed across all people

Article | Open Access | Published: 25 July 2019

Analysis of polygenic risk score usage and performance in diverse human populations

L. Duncan ✉, H. Shen, B. Gelaye, J. Meijsen, K. Ressler, M. Feldman, R. Peterson & B. Domingue

Nature Communications **10**, Article number: 3328 (2019) | Cite this article

2. May exacerbate health disparities.

Clinical use of current polygenic risk scores may exacerbate health disparities

Alicia R. Martin ✉, Masahiro Kanai, Yoichiro Kamatani, Yukinori Okada, Benjamin M. Neale & Mark J. Daly

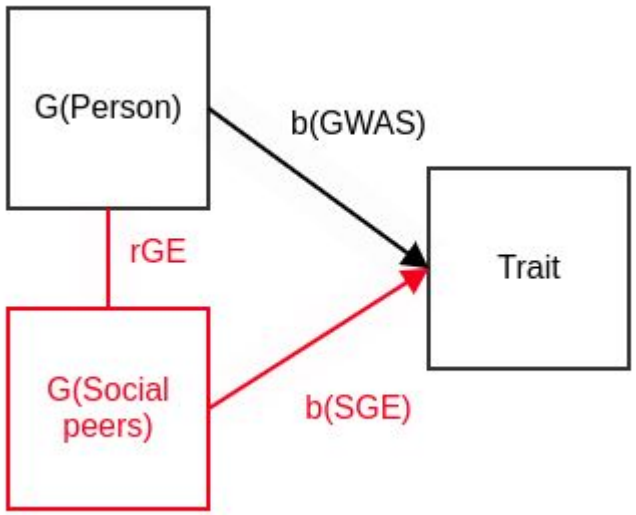
Nature Genetics **51**, 584–591(2019) | Cite this article

Caution: social genetic fx

In last few years, there has been a lot of attention paid to the problem of confounding in GWAS.

- GWAS estimates associations between black boxes.
- But, connections between an individual's genomes and their social peers could lead to indirect effects through red pathway.

This turns out to be an old problem in the animal literature (I learned about it when asked to review a paper about mice and how the genomes of their cagemates may matter).



The social genome: Current findings and implications for the study of human genetics

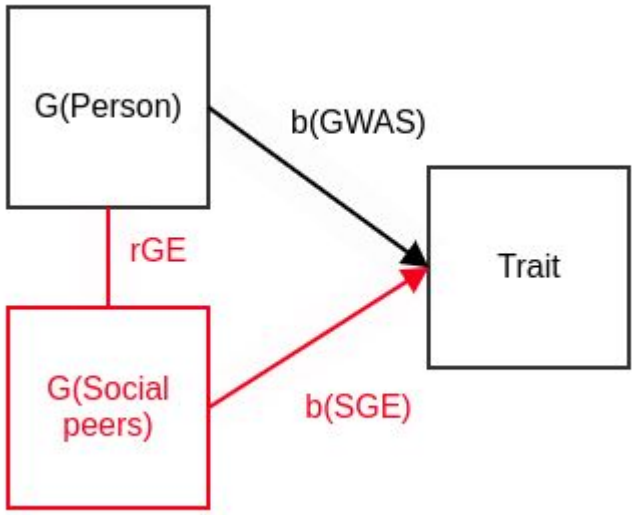
Benjamin W. Domingue, Daniel W. Belsky

Published: March 16, 2017 • https://doi.org/10.1371/journal.pgen.1006615

Caution: social genetic fx

As an analogue to the cagemate of a mouse, we aimed to study an adolescent's friends and classmates in the Add Health study.

- We showed connection between genomes of an individual and that of their social peers.
- And some connection between genomes of social peers and an individual's outcomes.



The social genome of friends and schoolmates in the National Longitudinal Study of Adolescent to Adult Health

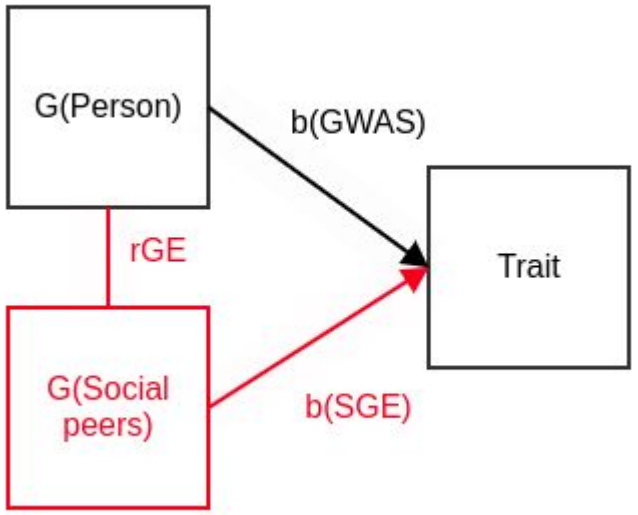
Benjamin W. Domingue, Daniel W. Belsky, Jason M. Fletcher, Dalton Conley, Jason D. Boardman, and Kathleen Mullan Harris

PNAS January 23, 2018 115 (4) 702-707; published ahead of print January 9, 2018
https://doi.org/10.1073/pnas.1711803115

Caution: social genetic fx

As an analogue to the cagemate of a mouse, we aimed to study an adolescent's friends and classmates in the Add Health study.

- We showed connection between genomes of an individual and that of their social peers.
- And some connection between genomes of social peers and an individual's outcomes.



Kong et al. asked same question with respect to parents (as social peers).

- You don't have to worry about whether there is genetic similarity between an individual and their parents.
- They were able to show that the *non-transmitted* genetic information held by Icelandic parents predicted child outcomes.
- 'Genetic Nurture'

The social genome of friends and schoolmates in the National Longitudinal Study of Adolescent to Adult Health

Benjamin W. Domingue, Daniel W. Belsky, Jason M. Fletcher, Dalton Conley, Jason D. Boardman, and Kathleen Mullan Harris

PNAS January 23, 2018 115 (4) 702-707; published ahead of print January 9, 2018
https://doi.org/10.1073/pnas.1711803115

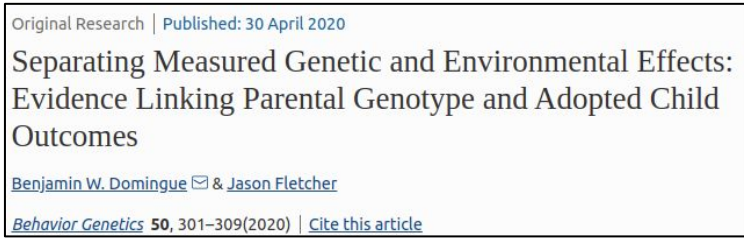
The nature of nurture: Effects of parental genotypes

Augustine Kong[1,2,3,*], Gudmar Thorleifsson[1], Michael L. Frigge[1], Bjarni J. Vilhjalmsson[4,5], Alexander I. Young[1],...
+ See all authors and affiliations

Science 26 Jan 2018:
Vol. 359, Issue 6374, pp. 424-428
DOI: 10.1126/science.aan6877

Genetic nurture

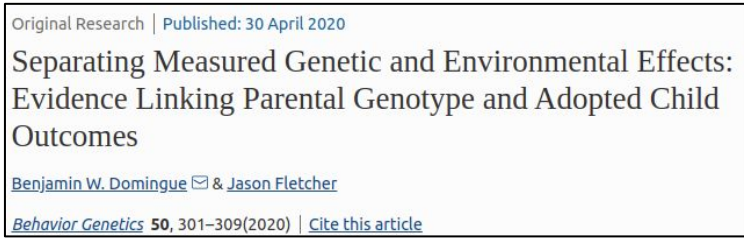
- As a different lens for understanding the problem, we showed that the genomes of adopted parents predict outcomes for non-biological offspring.



Genetic nurture

- As a different lens for understanding the problem, we showed that the genomes of adopted parents predict outcomes for non-biological offspring.
- We showed that parental genetics are predictive of maternal environments when child is in utero and also predictive of early academic outcomes when child is 4-7y.

Very challenging to disentangle!!



Original Research | Published: 30 April 2020

Separating Measured Genetic and Environmental Effects: Evidence Linking Parental Genotype and Adopted Child Outcomes

Benjamin W. Domingue & Jason Fletcher

Behavior Genetics **50**, 301–309(2020) | Cite this article

› Psychol Sci. 2020 Jul;31(7):781-791. doi: 10.1177/0956797620917209. Epub 2020 Jun 2.

The Earliest Origins of Genetic Nurture: The Prenatal Environment Mediates the Association Between Maternal Genetics and Child Development

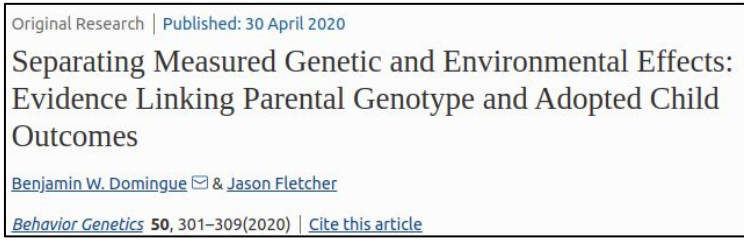
Emma Armstrong-Carter [1], Sam Trejo [1], Liam J B Hill [2 3], Kirsty L Crossley [3], Dan Mason [3], Benjamin W Domingue [1 4]

Genetic nurture

- As a different lens for understanding the problem, we showed that the genomes of adopted parents predict outcomes for non-biological offspring.
- We showed that parental genetics are predictive of maternal environments when child is in utero and also predictive of early academic outcomes when child is 4-7y.

Very challenging to disentangle!!

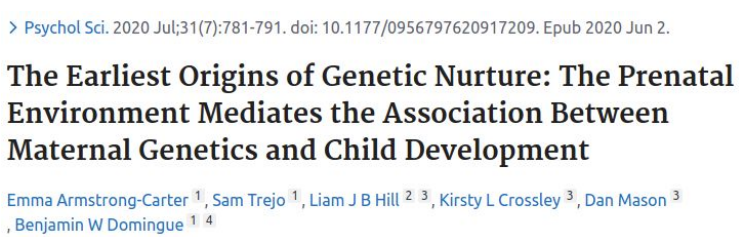
- But, we are able to specify a model that may be useful in helping us further understand how these associations come to be (by introducing novel parameters) that might further inform us about the reason for genetic nurture.


Original Research | Published: 30 April 2020

Separating Measured Genetic and Environmental Effects: Evidence Linking Parental Genotype and Adopted Child Outcomes

Benjamin W. Domingue & Jason Fletcher

Behavior Genetics **50**, 301–309(2020) | Cite this article


› Psychol Sci. 2020 Jul;31(7):781-791. doi: 10.1177/0956797620917209. Epub 2020 Jun 2.

The Earliest Origins of Genetic Nurture: The Prenatal Environment Mediates the Association Between Maternal Genetics and Child Development

Emma Armstrong-Carter [1], Sam Trejo [1], Liam J B Hill [2 3], Kirsty L Crossley [3], Dan Mason [3], Benjamin W Domingue [1 4]


› Biodemography Soc Biol. Jul-Sep 2018;64(3-4):187-215. doi: 10.1080/19485565.2019.1681257.

Genetic nature or genetic nurture? Introducing social genetic parameters to quantify bias in polygenic score analyses

Sam Trejo [1], Benjamin W Domingue [1]

- These results suggest that GWAS findings may be confounded by non within-person fx.
- This is a problem, but also reminds us that estimates of environments need not be free of confounding either.
- Other problems well-known to social scientists may be relevant
  - Selection Bias
  - Health disparities as a function of SES

Within-sibship GWAS improve estimates of direct genetic effects

Laurence J Howe, Michel G Nivard, Tim T Morris, Ailin F Hansen, Humaira Rasheed, Yoonsu Cho, Geetha Chittoor, Penelope A Lind, Teemu Palviainen, Matthijs D van der Zee, Rosa Cheesman, Massimo Mangino, Yunzhang Wang, Shuai Li, Lucija Klaric, Scott M Ratliff, Lawrence F Bielak, Marianne Nygaard, Chandra A Reynolds, Jared V Balbona, Christopher R Bauer, Dorret I Boomsma, Aris Baras, Archie Campbell, Harry Campbell, Zhengming Chen, Paraskevi Christofidou, Christina C Dahm, Deepika R Dokuru, Luke M Evans, Eco JC de Geus, Sudheer Giddaluru, Scott D Gordon, K. Paige Harden, Alexandra Havdahl, W. David Hill, Shona M Kerr, Yongkang Kim, Hyeokmoon Kweon, Antti Latvala, Liming Li, Kuang Lin, Pekka Martikainen, Patrik KE Magnusson, Melinda C Mills, Deborah A Lawlor, John D Overton, Nancy L Pedersen, David J Porteous, Jeffrey Reid, Karri Silventoinen, Melissa C Southey, Travis T Mallard, Elliot M Tucker-Drob, Margaret J Wright, Social Science Genetic Association Consortium, Within Family Consortium, John K Hewitt, Matthew C Keller, Michael C Stallings, Kaare Christensen, Sharon LR Kardia, Patricia A Peyser, Jennifer A Smith, James F Wilson, John L Hopper, Sara Hägg, Tim D Spector, Jean-Baptiste Pingault, Robert Plomin, Meike Bartels, Nicholas G Martin, Anne E Justice, Iona Y Millwood, Kristian Hveem, Øyvind Naess, Cristen J Willer, Bjørn Olav Åsvold, Philipp D Koellinger, Jaakko Kaprio, Sarah E Medland, Robin G Walters, Daniel J Benjamin, Patrick Turley, David M Evans, George Davey Smith, Caroline Hayward, Ben Brumpton, Gibran Hemani, Neil M Davies

doi: https://doi.org/10.1101/2021.03.05.433935

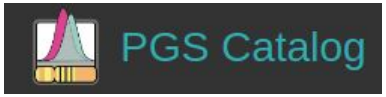
- These results suggest that GWAS findings may be confounded by non within-person fx.
- This is a problem, but also reminds us that estimates of environments need not be free of confounding either.
- Other problems well-known to social scientists may be relevant
  - Selection Bias
  - Health disparities as a function of SES

More on environments

- We've talked a little about environments
  - They may serve as confounders (i.e., genetic nurture, social genetic fx)
  - Genetic fx may vary across them (classic GxE)
- We're going to now discuss a specific source of confusion that might arise when we think about GxE

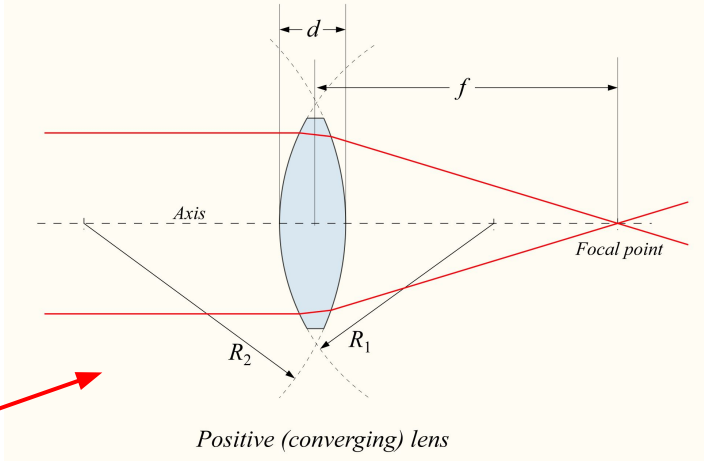
GxE research is hard.

- Widespread polygenic scores →more GxE research.
- Great interest in documenting ways in which genetic effects may be contingent on social environment.
  - Or, alternatively, ways in which social effects may be contingent on genotype.



GxE research is hard.

- Widespread polygenic scores →more GxE research.
- Great interest in documenting ways in which genetic effects may be contingent on social environment.
  - Or, alternatively, ways in which social effects may be contingent on genotype.
- Numerous challenges, paper at right began to explore some of them.
- A suggested typology:
  - Environment as LENS
  - Environment as DIMMER



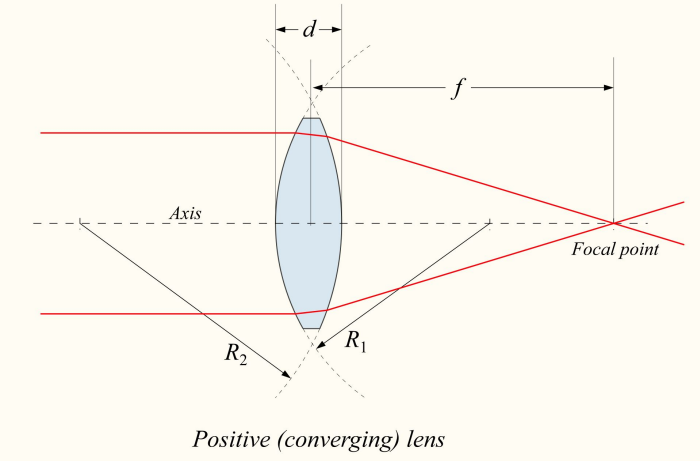
Interactions between Polygenic Scores and Environments: Methodological and Conceptual Challenges

Benjamin W. Domingue, Sam Trejo, Emma Armstrong-Carter, Elliot M. Tucker-Drob

Sociological Science September 21, 2020
10.15195/v7.a19

GxE research is hard.

- Widespread polygenic scores →more GxE research.
- Great interest in documenting ways in which genetic effects may be contingent on social environment.
  - Or, alternatively, ways in which social effects may be contingent on genotype.
- Numerous challenges, paper at right began to explore some of them.
- A suggested typology:
  - Environment as LENS
  - Environment as DIMMER
- Now going to focus on one problem that results from studying a particular environmental dimmer.



Heteroscedasticity

We're going to focus on one particular challenge: **heteroscedasticity of the error term**.

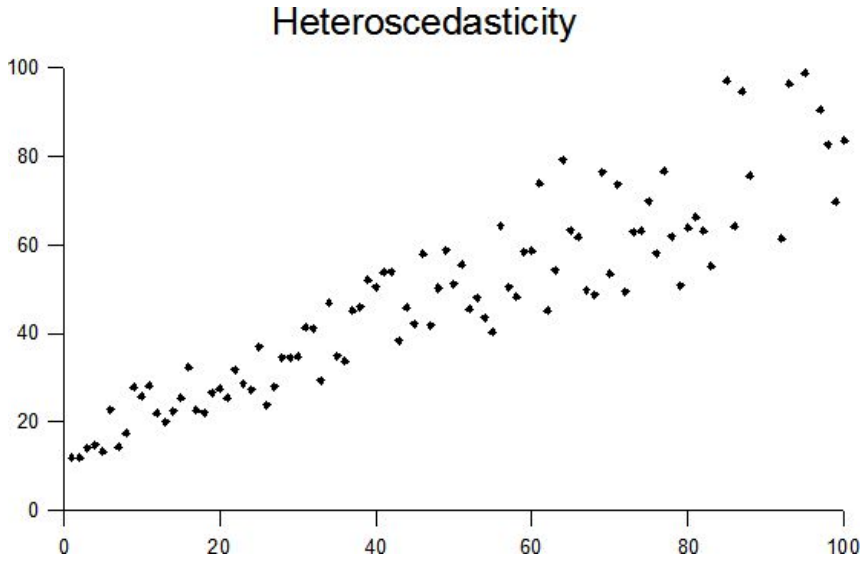
The error term: ϵ

- Typically assumed to be independent of predictors.
- And to have a normal distribution.

But life may be messier.

- Could be asymmetric.
- Could have shape that depends on the predictors.

Such heteroscedasticity can lead to confusion in GxE modeling.



Heteroscedasticity

We're going to focus on one particular challenge: **heteroscedasticity of the error term**.

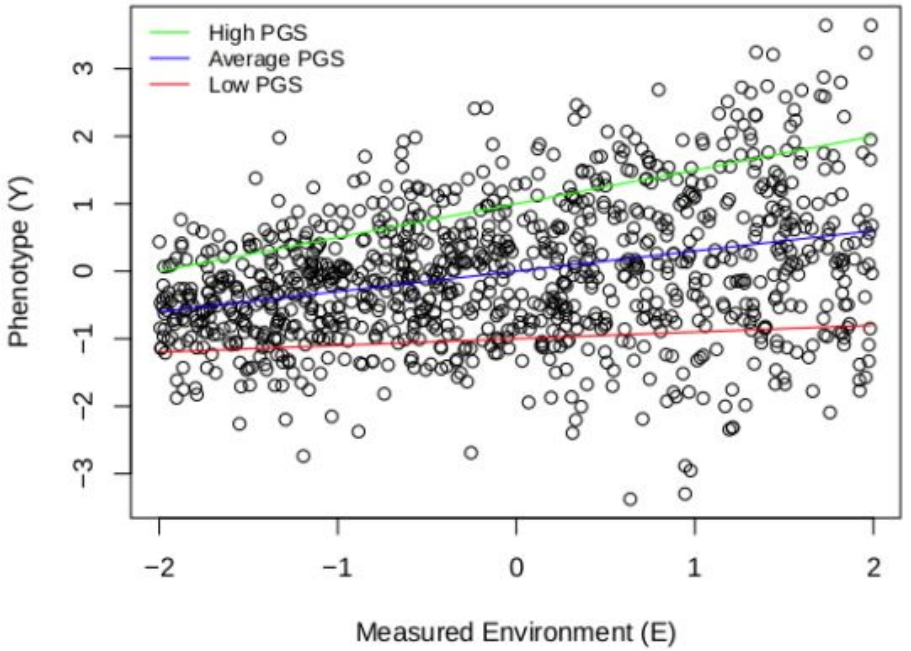
The error term: ϵ

- Typically assumed to be independent of predictors.
- And to have a normal distribution.

But life may be messier.

- Could be asymmetric.
- Could have shape that depends on the predictors.

Such heteroscedasticity can lead to confusion in GxE modeling.



Heteroscedasticity

We're going to focus on one particular challenge: **heteroscedasticity of the error term**.

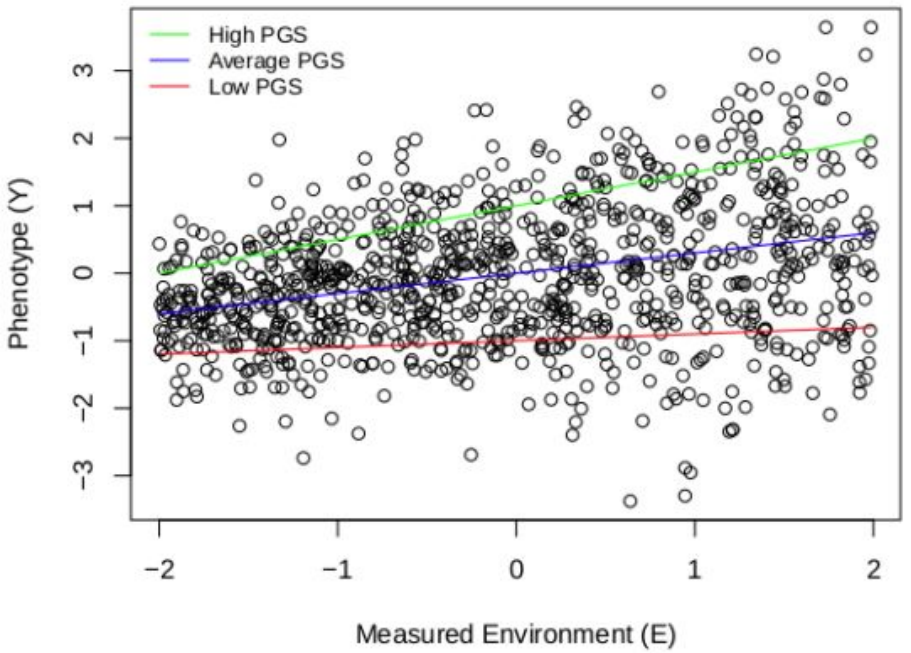
The error term: ϵ

- Typically assumed to be independent of predictors.
- And to have a normal distribution.

But life may be messier.

- Could be asymmetric.
- Could have shape that depends on the predictors.

Such heteroscedasticity can lead to confusion in GxE modeling.



We're interested in a particular scenario:

- A genetic predictor that **predicts a constant proportion of the outcome variance**.
- But we have an outcome whose variance shifts as a function of E (heteroscedasticity!)
- This heteroscedasticity can lead to GxE findings.

We will investigate this phenomenon and contrast it with traditional approaches via a comparison of 3 models.

Method overview

Distinguishing between interaction and dispersion effects in the analysis of gene-environment interaction

Benjamin W. Domingue, Klint Kanopka, Travis T. Mallard, Sam Trejo, Elliot M. Tucker-Drob

doi: https://doi.org/10.1101/2020.09.08.287888

We contrast three models:

1. Classic Interaction model
2. Scaling model (scale of the outcome changes but genetics explain constant proportion of variance)
3. Heteroscedasticity model (flexible model of which 1 & 2 are reduced forms)
    a. Note: not merely adjusting standard errors to account for heteroscedasticity, we model it directly.

Method overview

Distinguishing between interaction and dispersion effects in the analysis of gene-environment interaction

Benjamin W. Domingue, Klint Kanopka, Travis T. Mallard, Sam Trejo, Elliot M. Tucker-Drob

doi: https://doi.org/10.1101/2020.09.08.287888

We contrast three models:

1. Classic Interaction model
2. Scaling model (scale of the outcome changes but genetics explain constant proportion of variance)
3. Heteroscedasticity model (flexible model of which 1 & 2 are reduced forms)
    a. Note: not merely adjusting standard errors to account for heteroscedasticity, we model it directly.

We devise approaches to:

- Test whether data is consistent with #2 given estimates from #3.
- If not, ask about the role of scaling in the observed interaction.

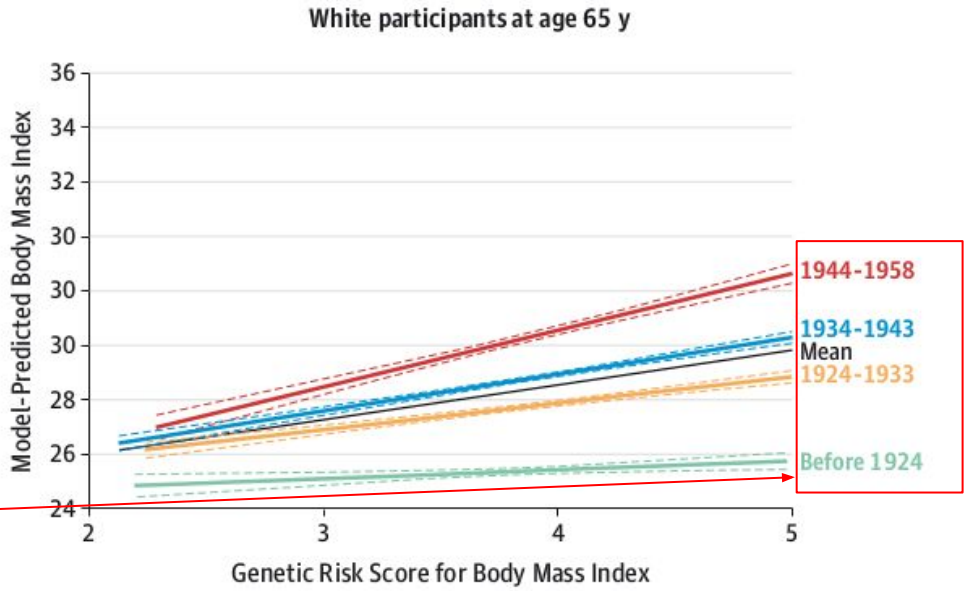
I'm obviously talking about a specific kind of interaction, but this idea could be applied in other settings where heteroscedasticity is a concern (i.e., nothing here specific to study of genotypes).

An empirical example: Birth year and BMI

- High-quality GWAS have produced relatively penetrant genetic predictors of BMI.
- Studies have examined shifting patterns of penetrance as a function of birth cohorts.

An empirical example: Birth year and BMI

- High-quality GWAS have produced relatively penetrant genetic predictors of BMI.
- Studies have examined shifting patterns of penetrance as a function of birth cohorts.
  - Polygenic score is more penetrant/increasingly predictive in more recent birth cohorts.
  - Note increasing slope for later birth cohorts.



Association of a Genetic Risk Score With Body Mass Index Across Different Birth Cohorts

Stefan Walter, PhD[1]; Iván Mejía-Guevara, PhD[2]; Karol Estrada, PhD[3]; et al

Author Affiliations | Article Information

JAMA. 2016;316(1):63-69. doi:10.1001/jama.2016.8729

An empirical example: Birth year and BMI

- High-quality GWAS have produced relatively penetrant genetic predictors of BMI.
- Studies have examined shifting patterns of penetrance as a function of birth cohorts.
  - Polygenic score is more penetrant/increasingly predictive in more recent birth cohorts.
  - Note increasing slope for later birth cohorts.
  - Observation made by several groups using data from the HRS.
    - Not meant as attack!

Am Sociol Rev. Author manuscript; available in PMC 2016 May 24.
PMCID: PMC4878452
Published in final edited form as:
NIHMSID: NIHMS780234
Am Sociol Rev. 2015 Aug; 80(4): 705–737.
PMID: 27231400
doi: 10.1177/0003122415590627

Lifetime Socioeconomic Status, Historical Context, and Genetic Inheritance in Shaping Body Mass in Middle and Late Adulthood

Hexuan Liu[1,3] and Guang Guo[1,2,3]

Sci Rep. 2016; 6: 30348.
PMCID: PMC4960614
Published online 2016 Jul 26. doi: 10.1038/srep30348
PMID: 27456657

Changing Polygenic Penetrance on Phenotypes in the 20th Century Among Adults in the US Population

Dalton Conley,[a,1] Thomas M. Laidley,[2] Jason D. Boardman,[3] and Benjamin W. Domingue[b,4]

▸ Author information ▸ Article notes ▸ Copyright and License information Disclaimer

An empirical example: Birth year and BMI

- High-quality GWAS have produced relatively penetrant genetic predictors of BMI.
- Studies have examined shifting patterns of penetrance as a function of birth cohorts.
  - Polygenic score is more penetrant/increasingly predictive in more recent birth cohorts.
  - Note increasing slope for later birth cohorts.
  - Observation made by several groups using data from the HRS.
    - Not meant as attack!

Am Sociol Rev. Author manuscript; available in PMC 2016 May 24.
PMCID: PMC4878452
Published in final edited form as:
NIHMSID: NIHMS780234
Am Sociol Rev. 2015 Aug; 80(4): 705–737.
PMID: 27231400
doi: 10.1177/0003122415590627

Lifetime Socioeconomic Status, Historical Context, and Genetic Inheritance in Shaping Body Mass in Middle and Late Adulthood

Hexuan Liu[1,3] and Guang Guo[1,2,3]

Sci Rep. 2016; 6: 30348.
PMCID: PMC4960614
Published online 2016 Jul 26. doi: 10.1038/srep30348
PMID: 27456657

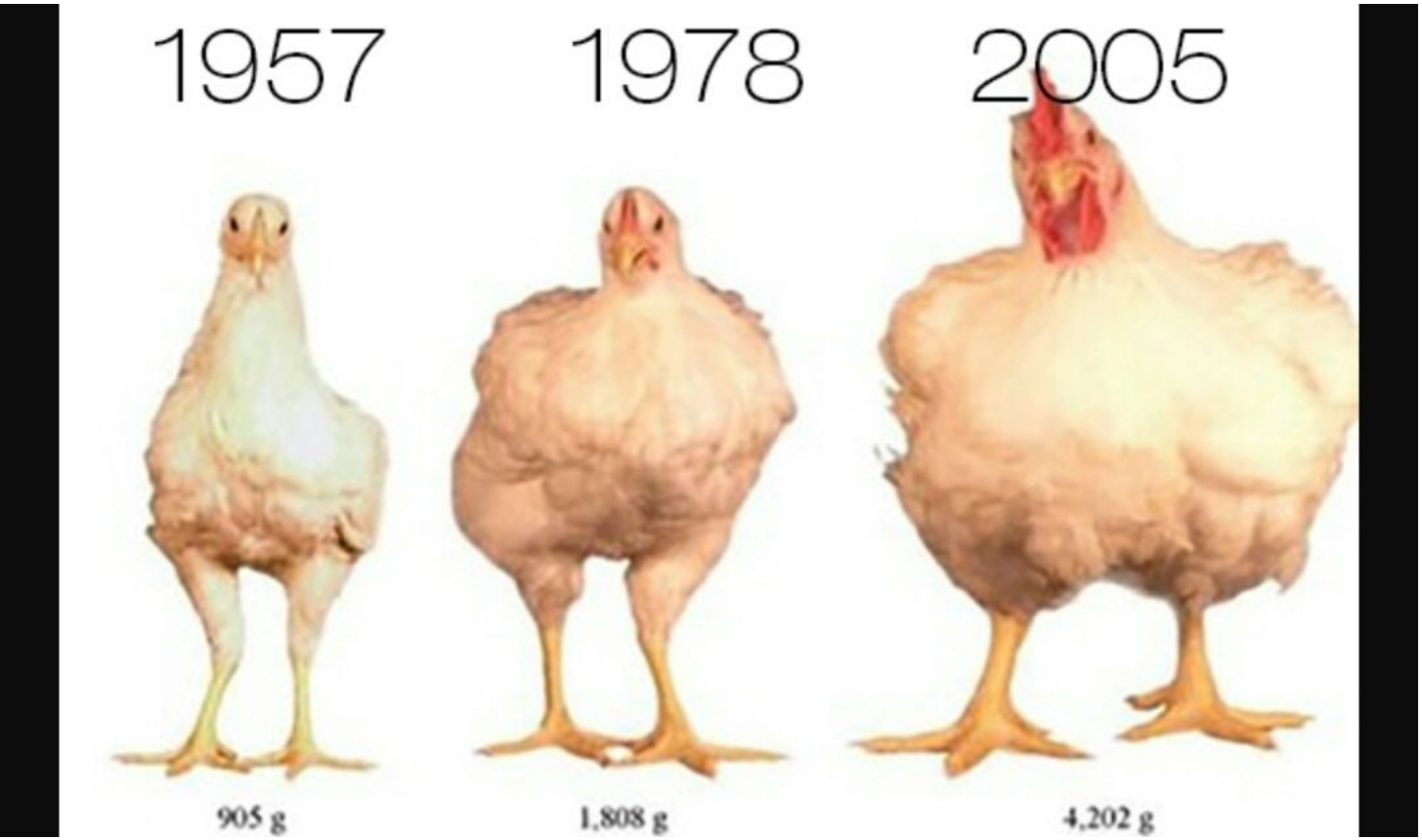
Changing Polygenic Penetrance on Phenotypes in the 20th Century Among Adults in the US Population

Dalton Conley,[a,1] Thomas M. Laidley,[2] Jason D. Boardman,[3] and Benjamin W. Domingue[b,4]

HRS | HEALTH AND RETIREMENT STUDY

An empirical example: Birth year and BMI

- But there is one very important fact about BMI: we've observed large shifts in the distribution over the last 50ish years.



An empirical example: Birth year and BMI

- But there is one very important fact about BMI: we've observed large shifts in the distribution over the last 50ish years.
  - For 50+ respondents, somewhere between 30-60% increase in class I obesity (BMI between 30 and 35) between 1960-1994.

International Journal of Obesity (1998) 22, 39–47

Overweight and obesity in the United States: prevalence and trends, 1960–1994

KM Flegal, MD Carroll, RJ Kuczmarski and CL Johnson

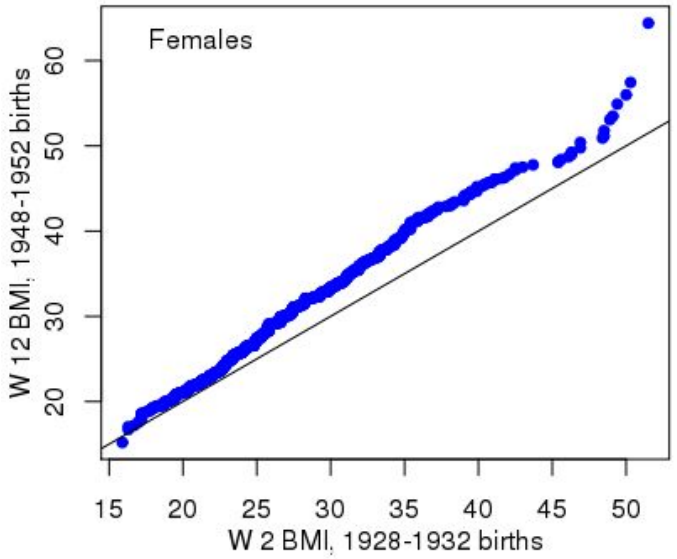
National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville MD 20782, USA

Table 2 Prevalence of class I obesity (30.0 ≤ BMI ≤ 34.9) by age group, gender and survey: United States, 1960–1994

	Age group (y)							Total	
	20–29	30–39	40–49	50–59	60–69	70–79[a]	≥ 80[b]	Crude[c] (≥ 20 y)	Age-adjusted (20–74 y)
Men:									
NHES I	6.6	9.1	10.7	12.1	7.0	7.0	–	9.1	8.8
NHANES I	5.9	10.6	12.5	12.0	8.6	10.4	–	9.7	9.5
NHANES II	6.0	9.6	13.7	11.2	12.2	11.9	–	10.0	10.0
NHANES III	8.4	11.6	16.5	22.5	20.4	15.2	6.6	14.3	14.6
Women:									
NHES I	4.4	8.1	11.6	14.0	18.8	16.1	–	11.2	10.4
NHANES I	4.9	8.9	10.8	13.9	18.4	15.9	–	10.9	10.5
NHANES II	6.0	9.5	10.3	14.0	14.2	14.0	–	10.4	10.2
NHANES III	8.6	14.1	15.5	20.2	17.4	15.9	11.2	14.5	14.2
Both genders:									
NHES I	5.5	8.6	11.1	13.1	13.2	11.8	–	10.2	9.6
NHANES I	5.4	9.7	11.6	12.9	14.1	13.6	–	10.3	10.1
NHANES II	6.0	9.6	11.9	12.7	13.3	13.1	–	10.2	10.1
NHANES III	8.5	12.9	16.0	21.4	18.8	15.6	9.6	14.4	14.4

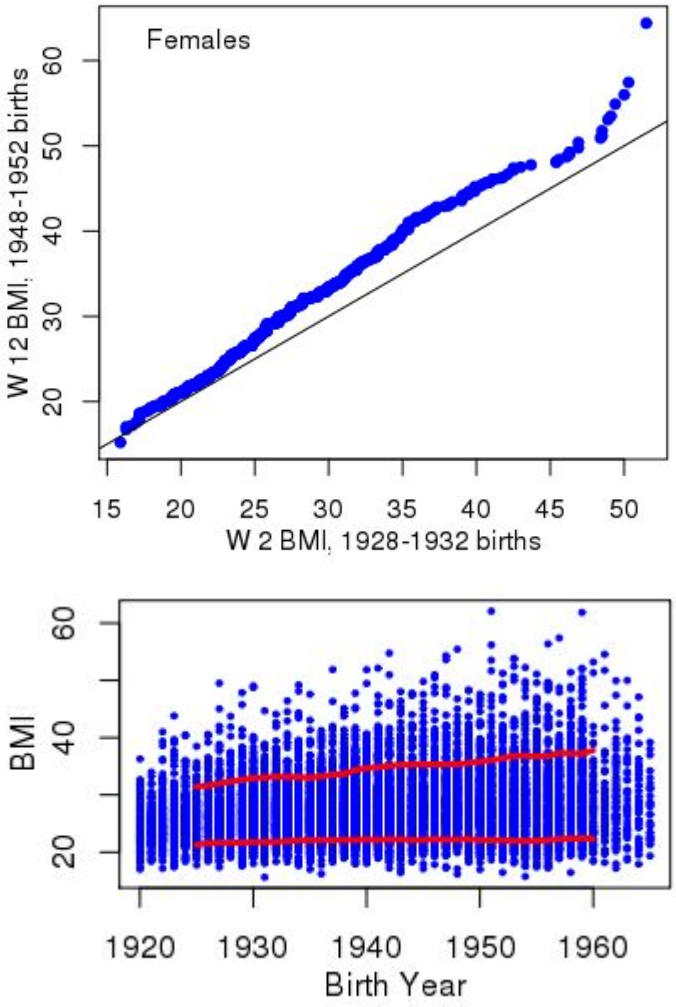
[a] For NHANES I and NHANES II, the estimates in this category are for persons aged 70–74 y only.
[b] Data for this age group only available from NHANES III.
[c] Crude (non-age-adjusted) total includes ages 20–79 y for NHES I, ages 20–74 y for NHANES I and NHANES II, and ages ≥ 20 y (no upper age limit) for NHANES III.

An empirical example: Birth year and BMI



- But there is one very important fact about BMI: we've observed large shifts in the distribution over the last 50ish years.
  - For 50+ respondents, somewhere between 30-60% increase in class I obesity (BMI between 30 and 35) between 1960-1994.
- A similar shift is readily observed in HRS as well.
  - Note that BMI quantiles for ~64y respondents are higher at Wave 12 as compared to Wave 2.

An empirical example: Birth year and BMI



- But there is one very important fact about BMI: we've observed large shifts in the distribution over the last 50ish years.
  - For 50+ respondents, somewhere between 30-60% increase in class I obesity (BMI between 30 and 35) between 1960-1994.
- A similar shift is readily observed in HRS as well.
  - Note that BMI quantiles for ~64y respondents are higher at Wave 12 as compared to Wave 2.
- This shift in distribution as a function of birth year leads to a concern: might heteroscedasticity result?
  - If so, could this drive the previous GxE findings?

HRS Results

Let's focus on the males.

Table 4: Estimates from parameters of model 2 in analysis of GxE for BMI as a function of birthyear in the HRS. We show probabilities for parameters when the maximal probability in a column is larger than $1e-6$.

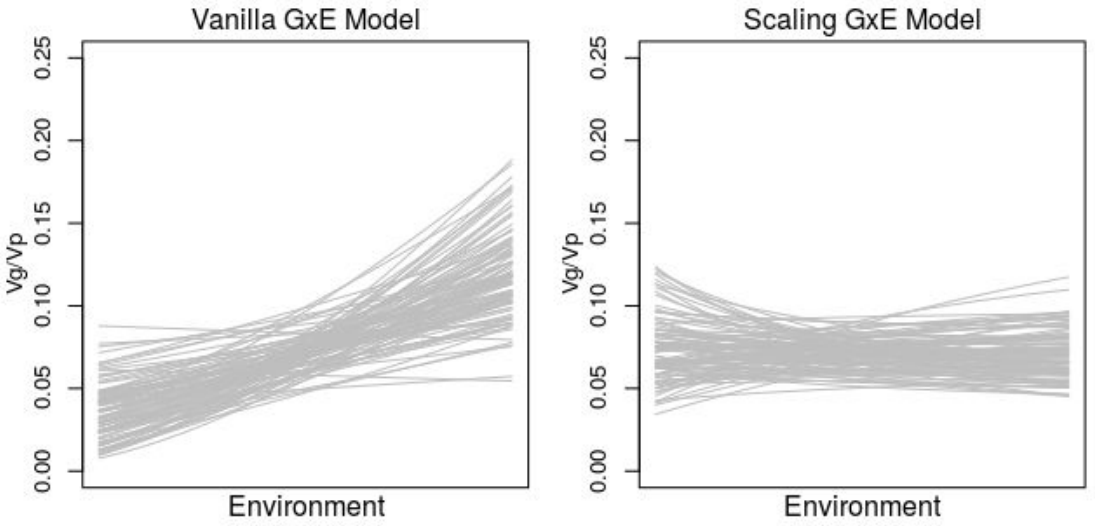
Gender	BMI[a]	N	τ_1	π_0	π_1	$\Pr_{\pi_1}$	λ_0	λ_1	ξ	$\Pr_{\xi}$
All	Std	11586	0.19	0.26	0.06	1.26e-10	0.94	0.13	-0.02	1.76e-02
All	Trans	11586	0.17	0.26	0.04	2.51e-05	0.95	0.06	-0.02	2.68e-02
M	Std	5022	0.19	0.26	0.04	6.12e-04	0.94	0.12	-0.01	3.91e-01
M	Trans	5022	0.17	0.26	0.03	2.35e-02	0.95	0.07	-0.01	3.83e-01
F	Std	6564	0.20	0.26	0.06	4.21e-08	0.93	0.13	-0.02	2.24e-02
F	Trans	6564	0.18	0.26	0.04	2.77e-04	0.95	0.06	-0.02	3.82e-02

[a] Std denotes the standardized mean BMI shown in the middle panel of Figure 2. Trans denotes the BoxCox transformation.

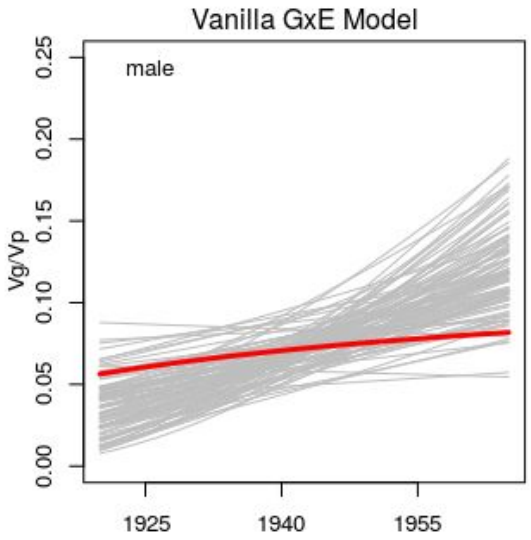
What does this imply for our understanding of the polygenic scores' behavior as a function of birth year?

- Posterior Predictive Check
  - We'll compare empirical results to data simulated under the vanilla/scaling GxE models
- We first estimates vanilla and scaling models in HRS.
  - We take estimates and simulate 100 new datasets.
- We'll then estimate the full model (#3) and see how it comports with simulated data.

$$\frac{Vg}{Vp} = \frac{(\pi_0 + \pi_1 E)^2}{(\pi_0 + \pi_1 E)^2 + (\lambda_0 + \lambda_1 E)^2}.$$



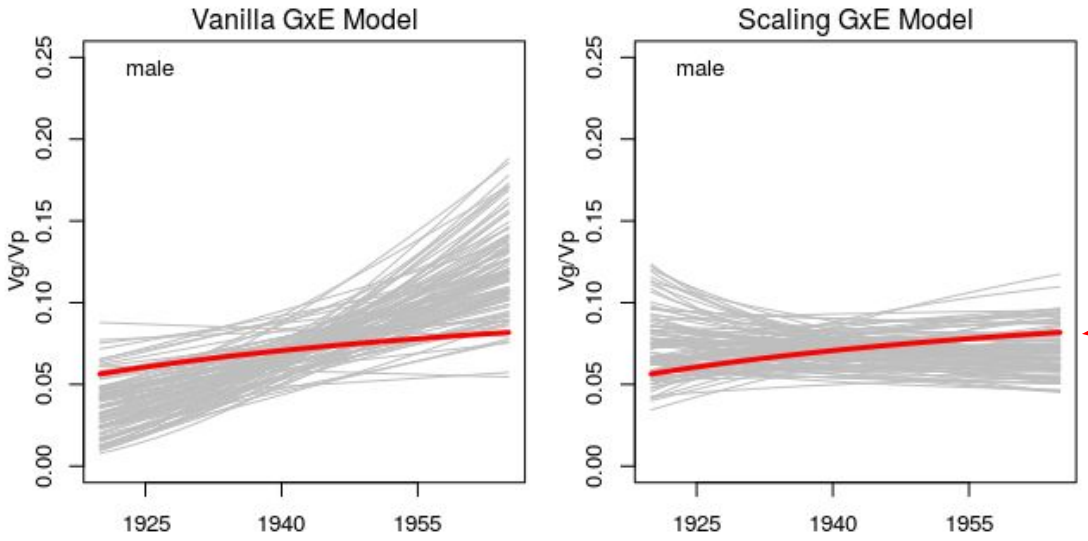
Comparing empirical data to simulated data



For males:

- Data are too ‘flat’ for vanilla GxE model.

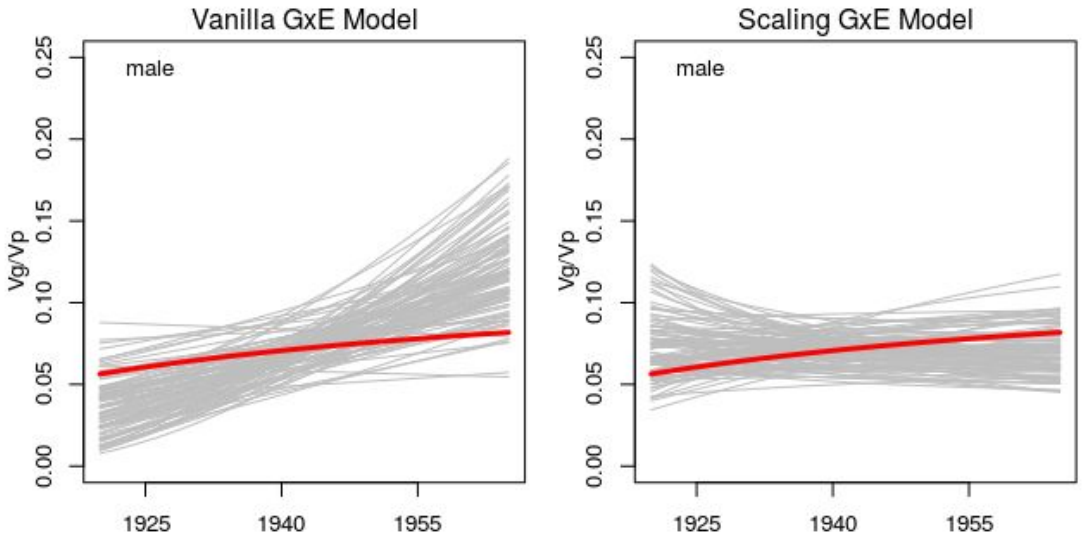
Comparing empirical data to simulated data



For males:

- Data are too ‘flat’ for vanilla GxE model.
- Scaling GxE produces data of ‘flatness’ comparable to that of our empirical data.

Comparing empirical data to simulated data

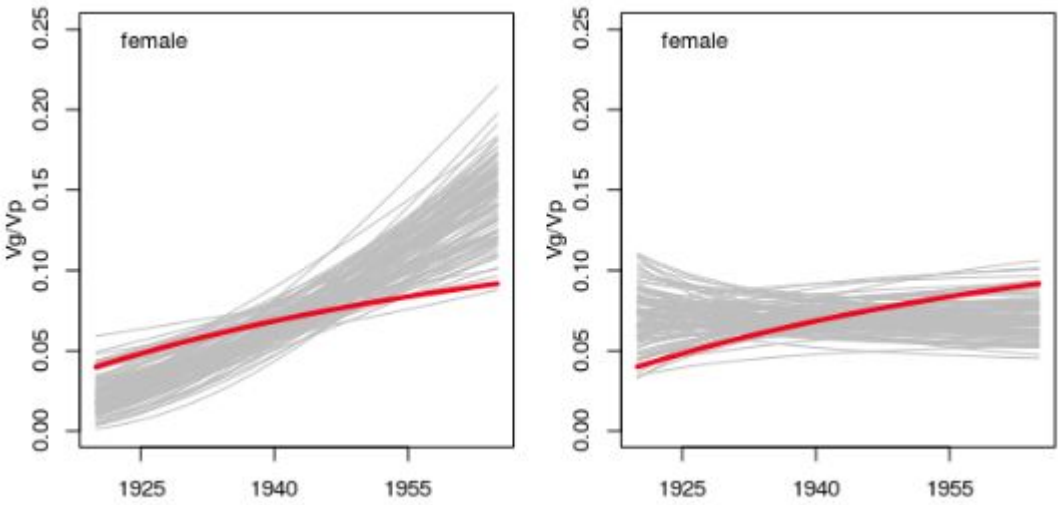


For males:

- Data are too ‘flat’ for vanilla GxE model.
- Scaling GxE produces data of ‘flatness’ comparable to that of our empirical data.

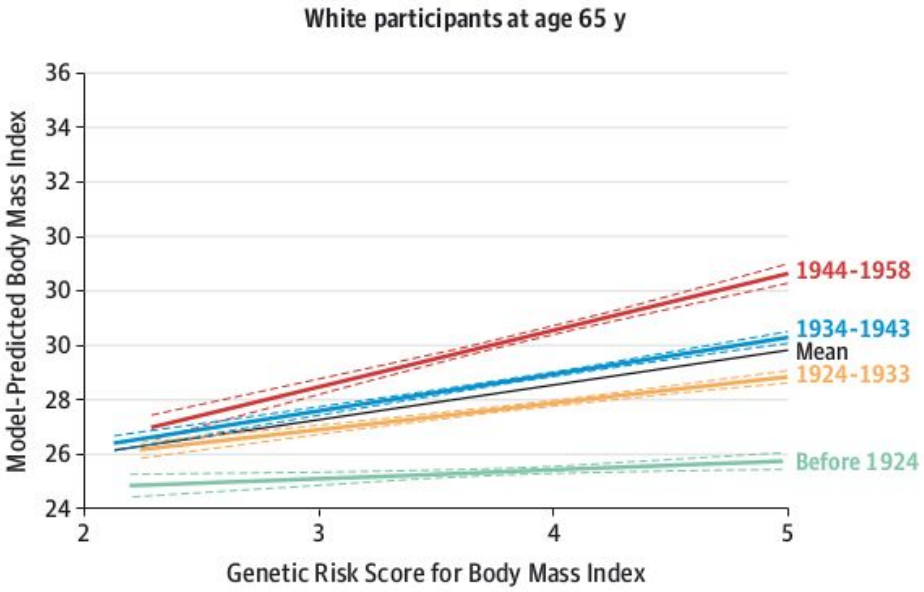
For males: PGS explains relatively constant amount of phenotypic variance (but phenotypic variance is changing).

For females: Not all scaling



For females: Messier Story. Data are not consistent with either reduced form model.

All respondents



Association of a Genetic Risk Score With Body Mass Index Across Different Birth Cohorts

Stefan Walter, PhD[1]; Iván Mejía-Guevara, PhD[2]; Karol Estrada, PhD[3]; et al

Author Affiliations | Article Information

JAMA. 2016;316(1):63-69. doi:10.1001/jama.2016.8729

If we analyze all respondents,

- We can decompose pi_1 into
  - 63% [CI: 50-96%] scaling
  - 37% non-scaling.
- My best guess: This finding in HRS has a lot to do with changing variation in BMI across birth year.

In closing

- We have described this kind of GxE as a ‘dimmer’ (E controls variation in outcome which induces GxE).
  - For conceptual reasons, we argue that this phenomenon may be relatively common.
  - In population-scale data, we have environments with big main effects.
- Code is available to conduct these analyses:
  - https://github.com/ben-domingue/scalingGxE
- GxE research could be illuminating but could also lead to needlessly complex mental models. Care is needed.

In closing

- We have described this kind of GxE as a ‘dimmer’ (E controls variation in outcome which induces GxE).
  - For conceptual reasons, we argue that this phenomenon may be relatively common.
  - In population-scale data, we have environments with big main effects.
- Code is available to conduct these analyses:
  - https://github.com/ben-domingue/scalingGxE
- GxE research could be illuminating but could also lead to needlessly complex mental models. Care is needed.
- Genetic data is increasingly available and likely to play a role in study of human behavior.
  - Ancillary information from the genome (e.g., methylation data)
- Care & caution are warranted
- Complex issues I haven’t discussed
  - Lots of thorny ethical problems in terms of their application
  - How the public deals with these scientific insights is also something we need to think more about.

NIH Eunice Kennedy Shriver National Institute of Child Health and Human Development
Health research throughout the lifespan

Collaborators:

- Emma Armstrong-Carter
- Dan Belsky
- Jason Boardman
- Dalton Conley
- Jason Fletcher
- Jeremy Freese
- Amal Harrati
- Paige Harden
- Pamela Herd
- Kathleen Mullan Harris
- Klint Kanopka
- Daphne Martschenko
- David Rehkopf
- Sam Trejo
- Elliot Tucker-Drob
- Robbee Wedow

(Postdocs & Grad Students)

RUSSELL SAGE FOUNDATION

JACOBS FOUNDATION
Our Promise to Youth

Distinguishing between interaction and dispersion effects in the analysis of gene-environment interaction

Benjamin W. Domingue, Klint Kanopka, Travis T. Mallard, Sam Trejo, Elliot M. Tucker-Drob

doi: https://doi.org/10.1101/2020.09.08.287888

THANK YOU!

bdomingue@stanford.edu