# Predicting Cognitive-Ability Differences from Genetic and Brain-Imaging Data

—

*Emily A. Willoughby and James J. Lee*

## Contents

# 1 Introduction

Statistical genetics and brain imaging are together at the technological forefront of research into human intelligence. While these approaches have historically had little practical overlap, they are united both conceptually and in several broad methodological challenges. In concept, both areas attempt to explain complex human behavior by understanding its biological origins, and in doing so have faced the problems that arise from this complexity. The prospect of finding large-effect predictors, for example, has shaped both histories: statistical genetics with its study of candidate genes that were once thought to have outsized influence on the development of many traits, and neuroscience with its search for localized brain properties underlying complex behaviors. Both of these areas have then had to adjust their scope and methodology to address the issue of making valid and meaningful predictions from a large number of predictors with small effects. A key understanding is that larger samples of participants than originally employed may be necessary for these predictions to be accurate and useful.

This chapter is about our modern efforts to predict human intelligence and related outcomes from their biological antecedents. We will briefly review the discoveries and challenges that are shaping each discipline today, and some of the further insights that have been enabled by recent increases in predictive power.

If readers feel skepticism about the value of predictions based on a large number of features, each accounting for a small or even minuscule fraction of the variance, an analogy to natural selection may begin to overcome it. Genome-wide association studies (GWAS) have been so productive largely because of the enormous samples that have deployed to reliably find genetic variants of significant effect. In doing so, scientists are using more or less the same information available to nature when organisms undergo natural selection:

> ... whether a given allele is correlated with the phenotype (which, in this case, is fitness) to the resolution afforded by size of the population (Fisher, 1941). Nevertheless, looking around at the exquisite adaptedness of living things, we can be confident that Nature correctly picks out alleles for their causal effects on fitness often enough (Lee, 2012; Lee & Chow, 2013). If we live in a world that is simple enough for natural selection to be robust, than perhaps it is not surprising that we can make progress by duplicating Nature's strategy of using large sample sizes to detect small DNA-trait correlations. (Lee & McGue, 2016, p. 30)

# 2 Predictions from Genetic Data

The study of quantitative genetics has its roots in the ancient arts of agriculture and domestication. Today, breeders are highly effective in augmenting artificial selection for valuable traits in animals and crops with modern genetic techniques (Spindel & McCouch, 2016; Wray, Kemper, Hayes, Goddard, & Visscher, 2019). The success of both traditional breeding and its formal offshoot quantitative genetics is in spite of polygenicity: the phenotypes of interest being affected by many genetic variants, each of extremely small effect. Indeed, the field of livestock genetics has given rise to the concept of the polygenic score (PGS), which has crossed over to become a useful tool in human genetics.

## 2.1 The Polygenic Score

Something recognizably akin to the modern PGS was proposed nearly 30 years ago under the term "marker-based selection" (MAS; Lande and Thompson, 1990). According to this proposal, breeders can estimate the regression

coefficients of polymorphic markers correlated with the causal variants affecting a trait of interest, calculate scores for individuals with genotyping data, and thereby increase the efficiency of artificial selection. A substantial development of this fundamental insight was put forth by Meuwissen, Hayes, and Goddard (2001), who among other things proposed the use of Bayesian priors to overcome the problem of marker number being greater than the sample size. This approach was later proposed for use in human genetics to identify individuals at high risk for disease (Wray, Goddard, and Visscher, 2007).

To our knowledge, the first empirical use of PGS came in an early GWAS of schizophrenia (Purcell et al., 2009). This was a landmark study in a number of ways. First, Purcell and colleagues were the first to use the term "polygenic score" for a phenotypic prediction based on a linear combination of single-nucleotide polymorphism (SNP) genotypes, the weights in this combination being the regression coefficients estimated in a GWAS of the trait. Second, the higher mean PGS in schizophrenic cases, amounting to 3% of the variance (Nagelkerke's pseudo $R^2$ from logistic regression), provided proof of principle that GWAS can lead to predictive power even if relatively few SNP associations have been identified at a stringent threshold of statistical significance. These authors also demonstrated that variance explained in their validation sample was actually reduced if the terms in the PGS were restricted to only significant SNPs. In other words, a better PGS resulted when the effects of *all* SNPs were included—no $p$-value was too high. To many observers this was conclusive evidence that schizophrenia is a highly polygenic trait, with SNPs that affect it scattered throughout the genome.

Years of education was the first cognitive phenotype to be studied successfully in a GWAS. Today, we are able to predict upwards of 10% of the variance in various cognitive traits using PGS derived from 1.1 million individuals (Lee et al., 2018). This predictive power is the result of both the sample size and methodological developments that have improved the construction of PGS.

## 2.2 Methods of Construction

In a GWAS, a set of genetic markers is genotyped or imputed in a training sample, and estimates are obtained of each marker's association with the trait of interest. These weights are then used to construct a PGS for each individual in a replication sample that is independent from the initial training sample (Dudbridge, 2013). The estimated PGS, $\hat{S}$, of an individual is equal to the weighted sum of the individual's marker genotypes, $X_j$, at $m$ SNPs, such that

$$\hat{S} = \sum_{j=1}^{m} X_j \hat{\beta}_j. \tag{1}$$

Different methodologies for construction arise chiefly from two sources: how to generate the weights of the SNPs, $\hat{\beta}_j$, and how to determine which $m$ SNPs should be included.

### 2.2.1 Naïve Methods

In the most straightforward "naïve" method, weights are set equal to the coefficient estimates from univariate regressions of the phenotype on each variant $j$ (Equation 1). The $m$ SNPs may be selected via an algorithm that uses pruning to approximate independence of all markers used in the score. The motivation behind pruning is that genetic variants are often correlated with variants that are nearby in the genome, a phenomenon called linkage disequilibrium (LD). Failing to account for this nonrandom association—which arises chiefly from the shared evolutionary history of nearby variants that seldom experience recombination—might reduce the accuracy of the PGS if causal variants are in more or less LD with their neighbors than null variants. We can further restrict the included SNPs by omitting those which fail

to meet a certain significance threshold for association with the phenotype. A more stringent threshold might be thought to boost the signal-to-noise ratio by only including variants for which evidence of true association with the phenotype is particularly high, but it is not always clear how to optimize this threshold. This method and its variants are generally referred to as "pruning + thresholding."

#### 2.2.2 LDpred

LDpred estimation is a Bayesian approach that attempts to explicitly model and account for genetic architecture and LD. Other methods of PGS construction that prune markers based on LD may go too far and discard useful information. LDpred estimation explicitly assumes a prior for the distribution of effect sizes and sets the weight for each variant equal to the mean of its posterior distribution after accounting for LD. The theory underlying LDpred is derived assuming the covariance matrix of the genotype data in the training sample is known. In practice, this matrix is not known, so we replace the training-sample covariance matrix by an approximation that is estimated using observed LD patterns in a reference sample of conventionally unrelated individuals with the same ancestral background as the training and validation samples.

Theory and simulations show that LDpred outperforms pruning + thresholding, especially at large sample sizes. For example, the prediction $R^2$ increased from 20.1% to 25.3% in a large schizophrenia dataset. Unlike its simpler predecessors, an LDpred estimation converges to the heritable variance explained by the SNPs as a function of increased sample size, a desirable property of a PGS (Vilhjálmsson et al., 2015).

#### 2.2.3 Penalized Regression

The prior distribution of effect size in LDpred and the related methods of Meuwissen et al. (2001) can be regarded as supplying the additional information needed to estimate $m$ partial regression coefficients in a training sample consisting of fewer than $m$ individuals. An alternative approach toward this end is to "penalize" the regression model for having large coefficients and thereby shrink them conservatively (although such methods often also have a Bayesian interpretation). Ridge regression shrinks the prediction with a term penalizing the sum of the squared coefficients (de Vlaming and Groenen, 2015); LASSO does the same with the sum of absolute coefficients (Vattikuti, Lee, Chang, Hsu, & Chow, 2014). It is possible to apply these methods to univariate summary statistics, employing a reference panel to estimate the covariances between SNPs as in LDpred (Mak, Porsch, Choi, Zhou, & Sham, 2017). Some preliminary results obtained with LASSO in particular—including results with respect to IQ and educational achievement—suggest that this form of penalization can lead to a prediction $R^2$ at least as large as that of LDpred (Lello et al., 2018; Allegrini et al., in press).

If the fraction of common SNPs that must be given a nonzero weight is relatively small, say less than 1%, then on theoretical grounds we might expect LASSO to converge most efficiently on the full heritable variance. This is an issue worth following as research unfolds.

### 2.3 Empirical Applications

One of the outcomes that has been studied most successfully with PGS is educational attainment (measured in total years of education, *EduYears*). The first GWAS of educational attainment by the Social Science Genetic Association Consortium (SSGAC) found three SNPs reaching genome-wide statistical significance in a sample of ∼100,000 individuals and produced a PGS that accounted for 2% of the variance in educational attainment in independent samples (Rietveld et al., 2013). Three years later, a GWAS with ∼300,000 participants (EA2) found 74 loci (Okbay et al., 2016), and in the most recent SSGAC study of educational attainment (EA3), a GWAS of 1.1 million individuals iden-

tified 1,271 independent lead SNPs. A PGS constructed from this large sample is now able to predict 11–13% of the variance in educational attainment and 7–10% of the variance in IQ, providing further evidence that educational attainment is a viable proxy for cognitive ability in genetic research.

But a PGS can do more than simply explain a portion of variance in educational attainment, IQ, or other cognitive phenotypes. Even though these scores have not yet amassed the sample sizes necessary for explaining all heritable variance, they are reliable enough for many substantive research purposes, including the study of the effects that parents can have on the outcomes of their children, and how far a person rises through the ranks of society.

#### 2.3.1 Genetic Nurture

PGS have been used to predict outcomes in offspring consistent with a causal role of the environment fostered by the parents. The first study to demonstrate this sort of "genetic nurture"—or "passive gene-environment correlation," in the terminology of Plomin, DeFries, and Loehlin (1977)—found that only 70% of the correlation between the *EduYears* PGS and educational attainment is due to the effect of the offsprings' own scores on their own attainments. The remainder is due to the parent PGS acting as a confounder, affecting both the PGS and educational attainment of the offspring. This was inferred from a significant effect of the nontransmitted portion of the parent PGS on offspring attainment (Kong et al., 2018). A number of subsequent studies have replicated this finding (Belsky et al., 2018; Liu, 2018; Bates et al., 2019).

What is the heritable parental characteristic affecting offspring *EduYears*? Bates et al. (2019) found that parent socioeconomic status (SES) completely mediates the effect of the nontransmitted parent PGS. In a recent study at the University of Minnesota, we have examined this issue as well and also found that parent SES appears to be a complete mediator (Willoughby, McGue, Iacono, Rustichini, & Lee, under review). Parent IQ and *EduYears* by itself (rather than as part of the SES composite) also either substantially or completely attenuate the effect of parent PGS.

#### 2.3.2 Social Mobility

PGS can also help to guide our understanding of the role of genetics in societal success, especially since the prediction $R^2$ of the EA3 PGS exceeds that of parent income (Lee et al., 2018). Because children inherit both genes and socially transmitted advantage from their parents, it is plausible that an observed association between genes and social outcomes could be spurious. In a sample of over 20,000 participants in five longitudinal studies across the United States, Britain, and New Zealand, Belsky et al. (2018) found that individuals with higher *EduYears* PGS also tended to accumulate more wealth, more education, and greater success in their careers. Furthermore, however, they employed a within-family design to ask the key questions: Do offspring with the *higher* PGS, compared to their parents, tend to climb the social ladder beyond their parents' achievements? And, in families with multiple offspring, does the sibling with the higher PGS also tend to achieve more than his or her other siblings?

The answer to both questions, it turns out, is yes: Whether relative to the parents or to the other sibling, the individual with the higher PGS tends to be more upwardly mobile in career status, education, and wealth. These findings contradict the notion that GWAS results are nothing more than correlates of privilege, but rather affirm that at least some portion of this genetic endowment is likely to be causal.

#### 2.3.3 Assortative Mating

Assortative mating refers to any kind of mating preference leading to a correlation between the trait values of mothers and fathers. The psychological basis of a given preference might

be somewhat obscure. Do short people tend to have short spouses because they prefer a mate of similar stature? Or does everyone prefer taller mates, and short people simply have to settle for what remains after the taller have chosen? Regardless of the answer, Fisher (1918) noted that any such preference can have profound consequences for the genetic composition of the population, increasing the magnitudes of both the correlations between relatives and the additive genetic variance.

In a very clever study, Yengo et al. (2018) detected the signature of such assortative mating with respect to *EduYears* by finding a significant correlation between the PGS calculated over only the odd chromosomes and that calculated over only the even chromosomes. In the absence of assortative mating, this correlation is expected to be zero, essentially because of Mendelian independent assortment. But assortative mating will lead to the contributions of the different chromosomes to the PGS of a given gamete to be positively correlated, because knowing that a parent who contributed one set of chromosomes was phenotypically above average means that the parent contributing the complementary set was likely also above average.

#### 2.3.4 Evolution

Although intelligence has presumably increased fitness throughout millions of years of our recent evolution, it has been less clear how natural selection is operating on intelligence today. Recent applications of PGS have helped shed light on this question by using them to investigate signs of recent selection. One approach is to test for an association between PGS for a cognitive phenotype and measures of Darwinian fitness, such as total number of children or lifetime reproductive success (LRS) relative to others of the same age and gender. This approach offers a powerful advantage over methods that lack GWAS data, in that a direct measure of the focal trait's additive genetic value (i.e., the PGS) enables the use of Robertson's (1966) Secondary Theorem of Natural Selection to calculate the amount of evolutionary change.

Using this method, Beauchamp (2016) found a negative association between *EduYears* PGS and LRS in a sample of ∼20,000 Americans, implying that natural selection is slowly favoring lower educational attainment at a rate of what amounts to −1.5 months of education per generation. Kong et al. (2018) presented corroborating evidence from a study of ∼100,000 Icelanders, which found *EduYears* PGS to be associated with delayed reproduction and fewer children overall. From this, they extrapolated that the mean *EduYears* PGS is declining at ∼0.01 standard units per decade. In other words, evolution *does* seem to be currently operating on human intelligence, but in the opposite direction from that which prevailed in the deep evolutionary past.

## 3 Predictions from Brain-Imaging Data

What types of predictions made possible by brain imaging have led to the largest amounts of explained variance in individual differences in human intelligence? In describing the current state of predictive utility in brain imaging and intelligence, we have deliberately chosen to focus only on the predictions, rather than the mechanisms that may underlie them. In doing so, we will also leave out many studies of historical interest. As neuroimaging studies continue to use larger samples, out-of-sample replication, more explicit best practices and methodological uniformity—removing researcher degrees of freedom—the predictions made by these techniques will continue to improve.

### 3.1 Imaging Techniques

Since their application to human intelligence beginning in the late 1980s, neuroimaging has provided several noninvasive methods of study-

ing the biological basis of cognition in humans. These techniques have historically been limited to relatively small samples of study, largely because they are time-consuming and expensive to conduct. There have been recent encouraging signs, however, that these limitations are starting to be overcome (e.g., Elliott et al., 2018).

#### 3.1.1 Positron Emission Tomography

Positron emission tomography (PET) scanning works by tracking the location of a radioactive tracer compound in a person's body. This tracer may be an isotope like fluorine-18. It is chemically bound to a biologically active molecule that is designed to be used by a specific organ or region of the body. The isotope emits positrons continuously, and annihilates the first electron it contacts to produce a pair of gamma photons. By detecting these gamma rays, a PET scanner tracks the production of positrons at relatively high resolution.

For studies of the brain, the fluorine-18 is attached to an analogue of glucose, producing a molecule called fluorodeoxyglucose (FDG). Since neurons need glucose to fire, the amount of FDG deposited in the brain varies depending on how rapidly the glucose is being used in that area. This enables the scanner to see when and where parts of the brain are working especially hard in response to a cognitive task or stimulus.

#### 3.1.2 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) uses a powerful magnetic field to align protons in water molecules. When pulsed with radio waves, these protons are rapidly knocked out of and then back into alignment. This causes them to emit their own radio frequencies as they change energy levels. When the process is applied continuously, the pattern of radio waves can be detected and converted into a three-dimensional representation of the density of water molecules in that area. This technique allows for the production of images of living tissue at exquisite spatial resolution.

Structural MRI is useful for accurately capturing brain structure and volume, but functional magnetic resonance imaging (fMRI) is needed for tracking function and temporal changes by detecting rapid changes in oxygen level during blood flow. The utility of fMRI to the study of intelligence, like PET scanning, relies on the principle that a working brain uses resources; in the case of fMRI, the resource is oxygen rather than glucose. The location, quantity and rate of use conveys information about when and where the brain is working hard.

### 3.2 Empirical Findings

#### 3.2.1 Brain Size and Structure

While older studies of IQ and head circumference or post-mortem brain weight have drawn criticism, a recent meta-analysis of the relationship between IQ and MRI-measured brain volume has produced an estimate of $r \approx .24$ (Pietschnig, Penke, Wicherts, Zeiler, & Voracek, 2015), although the low reliability with which $g$ is measured in many of the contributing studies may mean that this should be treated as a lower bound (Gignac & Bates, 2017). Nevertheless, this association remains robust across age, IQ domain, and sex, though its size may continued to be debated. Similar findings were recently reported by Cox, Ritchie, Fawns-Ritchie, Tucker-Drob, and Deary (2019), who investigated the association in a sample of several thousand individuals from the UK Biobank. Corrected for sex and age, the association between total brain volume and $g$ was found to be $r = .275$ (95% C.I. = [0.252, 0.299]).

By analogy to polygenic scoring in GWAS, one might expect that brain-based prediction of intelligence should proceed by analysis of the brain into distinct features and then data-driven learning of each feature's predictive weight. The simplest possible such analysis may be the factorization of brain volume into surface area and thickness, and indeed several recent studies have attempted to determine the contribution of one or both factors to intelligence (Schnack

et al., 2015; Vuoksimaa et al., 2015; Walhovd et al., 2016; Schmitt et al., 2019). These studies differ in some of their findings, perhaps as a result of developmental complexity, but do agree that surface area is consistently correlated with intelligence.

A greater predictive power of surface area is not inconsistent with Lee, McGue, Iacono, Michael, and Chabris (2019), who found roughly speaking that any SNP affecting intracranial volume is very likely to go on to affect IQ, *EduYears*, and so on. This is because surface area accounts for far more of the total variance in volume than thickness.

### 3.2.2 The P-FIT Model

Brain efficiency has held true as a property associated with intelligence. PET scanning, for example, has shown that protracted practice at the video game Tetris led to decreased, rather than increased, activation (i.e., glucose use). Furthermore, the rate of decreased activity as a function of practice appeared to be related to IQ scores, suggesting that smarter people are able to consolidate learning at a quicker rate, freeing up other resources—doing "more with less." Perhaps more remarkably, it appeared that smarter brains were particularly efficient in certain regions and pathways (Haier, 2011).

In the early 2000s, in an era of small samples and inconsistent methodology, researchers began noticing that despite these hurdles, the body of research on brain imaging and intelligence had produced substantial overlap in the associations between intelligence and certain brain regions. These areas appeared to be distributed throughout the brain—which is perhaps to be expected from the results of Lee et al. (2019)—but were most prominently represented in the parietal and frontal areas of the brain, and the connections among them. This led to the development of the *parieto-frontal integration theory* (P-FIT) model of intelligence (Jung & Haier, 2007).

Since its inception, the P-FIT model has enjoyed much success and validation in modern neuroimaging research. For example, although cortical thickness as a whole may not be strongly associated with IQ, cortical thickness in certain regions along the P-FIT track have been shown to contain some signal (Karama et al., 2009). The P-FIT model has also helped to clarify the role of brain efficiency in intelligence, as studies have suggested that IQ is associated negatively with total length of the network paths connecting functional areas (van den Heuvel, Stam, Kahn, & Hulshoff Pol, 2009; Li et al., 2009) and consumption of both glucose and oxygen in critical regions (Neubauer & Fink, 2009).

### 3.2.3 The Connectome

Largely informed by and consistent with the P-FIT model as well as the composite and polygenic nature of psychometric $g$ and its correlates, the search for neurobiological substrates of intelligence has turned to more complex distributed models. The Human Connectome Project (HCP), which has produced maps of complete structural and functional neural connections—"connectomes"—both within and across individuals, has opened the doors to more precise studies of connectivity in large samples. Dubois, Galdi, Paul, and Adolphs (2018), for example, have recently shown in an HCP sample of ∼800 individuals that fMRI measures of activity in resting-state connectivity matrices are able to predict 20% of the variance in general intelligence. Remarkably, this prediction value is *after* controlling for brain volume, which has a substantial association with IQ on its own, as well as sex, age, and in-scanner motion. Over the connectome, connectivity among four resting-state networks emerged as carrying the most information about $g$: frontoparietal, cingulo-opercular, default mode, and visual—in good agreement with the P-FIT. It is of interest that this study employed feature selection and weighting based on a combination of $p$-value thresholding and elastic net (which is, in turn, a

combination of LASSO and ridge regression)—techniques that we mentioned in our discussion of genetic prediction. This suggests a convergence of methodology in these respective fields as they continue to develop.

Fornito, Arnatkevičiūtė, and Fulcher (2019) have reviewed attempts to bridge the gap between connectome and transcriptome—the latter being a brain-wide gene expression atlas that documents the transcriptional activity of thousands of genes across many anatomical locations of the brain. It turns out that the spatial patterning of gene expression is closely linked to neuronal connectivity. The transcriptional landscape of the brain is dominated by broad spatial gradients, which represent variations in interregional connectivity, regional cellular architecture, and microcircuitry. While this family of methods seeking relationships between gene expression and brain connectivity has not to our knowledge been applied to human intelligence, many of the studies cited earlier have used data sources of one type or the other. The possibility of using both types for the purpose of prediction therefore seems promising.

# 4 Looking Forward

The future of predicting individual differences in cognitive ability from its biological substrate—both at the genetic and neuronal level—is bright, and new developments and insights are occurring weekly. For example, GWAS have been conducted of only IQ and *EduYears* but also self-rated math ability and highest math class ever taken (Lee et al., 2018), and these may point toward the prediction of more specific outcomes (Park, Lubinski, & Benbow, 2007).

There is another dimension to genetic and neurobiological approaches to the prediction of cognitive ability, and it is on that final note that we stress both hope and caution. The process of natural selection that gave rise to a brain capable of studying itself in these ways has also now led it to the cusp of profoundly altering its future. The predictive power of PGS may eventually cross a threshold enabling an acceleration of human evolution that has not been possible until now. As embryo selection and genetic engineering become more feasible and affordable, they have the potential to significantly influence national competitiveness, human capital, and global economic and scientific progress (Shulman & Bostrom, 2014). They also have the potential to increase class divides and perhaps to change what it means to be human altogether. It is difficult to anticipate whether our long-term interest will be served by uncontrolled use of genetic engineering. How to use this technology wisely is a matter that will affect the world our posterity will inhabit for generations to come.

# 5 References


Allegrini, A. G., Selzam, S., Rimfeld, K., von Stumm, S., Pingault, J.-B., & Plomin, R. (in press). Genomic prediction of cognitive traits in childhood and adolescence. *Molecular Psychiatry*. doi:10.1038/s41380-019-0394-4

Bates, T. C., Maher, B. S., Colodro-Conde, L., Medland, S. E., McAloney, K., Wright, M. J., ... Gillespie, N. A. (2019). Social competence in parents increases children's educational attainment: Replicable genetically-mediated effects of parenting revealed by non-transmitted DNA. *Twin Research and Human Genetics*, *22*(1), 1–3. doi:10.1017/thg.2018.75

Beauchamp, J. P. (2016). Genetic evidence for natural selection in humans in the contemporary United States. *Proceedings of the National Academy of Sciences*, *113*(28), 7774–7779. doi:10.1073/pnas.1600398113

Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi,

A., ... Harris, K. M. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*, *115*(31), E7275–E7284. doi:10.1073/pnas.1801238115

Cox, S. R., Ritchie, S. J., Fawns-Ritchie, C., Tucker-Drob, E. M., & Deary, I. J. (2019). Brain imaging correlates of general intelligence in UK Biobank. *bioRxiv*. doi:10.1101/599472

de Vlaming, R. & Groenen, P. J. F. (2015). The current and future use of ridge regression for prediction in quantitative genetics. *BioMed Research International*, *2015*, 1–18. doi:10.1155/2015/143712

Dubois, J., Galdi, P., Paul, L. K., & Adolphs, R. (2018). A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philosophical Transactions of the Royal Society B*, *373*(1756), 20170284. doi:10.1098/rstb.2017.0284

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLOS Genetics*, *9*(3), e1003348. doi:10.1371/journal.pgen.1003348

Elliott, L. T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K. L., Douaud, G., ... Smith, S. M. (2018). Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature*, *562*(7726), 210–216. doi:10.1038/s41586-018-0571-7

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, *52*, 399–433.

Fisher, R. A. (1941). Average excess and average effect of a gene substitution. *Annals of Eugenics*, *11*(1), 53–63. doi:10.1111/j.1469-1809.1941.tb02272.x

Fornito, A., Arnatkevičiūtė, A., & Fulcher, B. D. (2019). Bridging the gap between connectome and transcriptome. *Trends in Cognitive Sciences*, *23*(1), 34–50. doi:10.1016/j.tics.2018.10.005

Gignac, G. E. & Bates, T. C. (2017). Brain volume and intelligence: The moderating role of intelligence measurement quality. *Intelligence*, *64*(May), 18–29. doi:10.1016/j.intell.2017.06.004

Haier, R. J. (2011). Biological basis of intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 351–368). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511977244.019

Jung, R. E. & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(2), 135–154. doi:10.1017/S0140525X07001185

Karama, S., Y, A.-D., Haier, R. J., Deary, I. J., Lyttelton, O. C., Lepage, C., ... Brain Development Cooperative Group. (2009). Positive association between cognitive ability and cortical thickness in a representative US sample of healthy 6 to 18 year-olds. *Intelligence*, *37*(2), 145–155. doi:10.1016/j.intell.2008.09.006

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjálmsson, B. J., Young, A. I., Thorgeirsson, T. E., ... Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, *359*(6374), 424–428. doi:10.1126/science.aan6877

Lande, R. & Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, *124*(3), 743–756. doi:10.1046/j.1365-2540.1998.00308.x

Lee, J. J. (2012). Correlation and causation in the study of personality (with discussion). *European Journal of Personality*, *26*, 372–412. doi:10.1002/per.1863

Lee, J. J. & Chow, C. C. (2013). The causal meaning of Fisher's average effect. *Genetics Research*, *95*(2-3), 89–109. doi:10.1017/S0016672313000074

Lee, J. J. & McGue, M. (2016). Why behavioral genetics matters: A comment on Plomin (2016). *Perspectives on Psychological Science*, *11*(1), 29–30. doi:10.1177/1745691615611932

Lee, J. J., McGue, M., Iacono, W. G., Michael, A. M., & Chabris, C. F. (2019). The causal influence of brain size on human intelligence: Evidence from within-family phenotypic associations and GWAS modeling. *Intelligence*, *75*, 48–58. doi:10.1016/j.intell.2019.01.011

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., ... Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*(8), 1112–1121. doi:10.1038/s41588-018-0147-3

Lello, L., Avery, S. G., Tellier, L., Vazquez, A. I., de los Campos, G., & Hsu, S. D. H. (2018). Accurate genomic prediction of human height. *Genetics*, *210*(2), 477–497. doi:10.1534/genetics.118.301267

Li, Y., Liu, Y., Li, J., Qin, W., Li, K., Yu, C., & Jiang, T. (2009). Brain anatomical network and intelligence. *PLOS Computational Biology*, *5*(5), e1000395. doi:10.1371/journal.pcbi.1000395

Liu, H. (2018). Social and genetic pathways in multigenerational transmission of educational attainment. *American Sociological Review*, *83*(2), 278–304. doi:10.1177/0003122418759651

Mak, T. S. H., Porsch, R. M., Choi, S. W., Zhou, X., & Sham, P. C. (2017). Polygenic scores via penalized regression on summary statistics. *Genetic Epidemiology*, *41*(6), 469–480. doi:10.1002/gepi.22050

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. doi:11290733

Neubauer, A. C. & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience and Biobehavioral Reviews*, *33*(7), 1004–1023. doi:10.1016/j.neubiorev.2009.04.001

Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., ... Benjamin, D. J. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, *533*(7604), 539–542. doi:10.1038/nature17671. arXiv: NIHMS150003

Park, G., Lubinski, D., & Benbow, C. P. (2007). Contrasting intellectual patterns predict creativity in the arts and sciences: Tracking intellectually precocious youth over 25 years. *Psychological Science*, *18*(11), 948–952. doi:10.1111/j.1467-9280.2007.02007.x

Pietschnig, J., Penke, L., Wicherts, J. M., Zeiler, M., & Voracek, M. (2015). Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean? *Neuroscience and Biobehavioral Reviews*, *57*, 411–432. doi:10.1016/j.neubiorev.2015.09.017

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*(2), 309–322. doi:10.1037/0033-2909.84.2.309

Purcell, S. M., Pato, M. T., Williams, N. M., Scolnick, E. M., Van Beck, M., O'Donovan, M. C., ... Holmans, P. A. (2009). Common polygenic variation con-

tributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*(August), 748–752. doi:10.1038/nature08185

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., ... Koellinger, P. D. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, *25*(4), 57–82. doi:10.1257/jep.25.4.57

Robertson, A. (1966). A mathematical model of the culling process in dairy cattle. *Animal Production*, *8*(1), 95–108. doi:10.1017/S0003356100037752

Schmitt, J. E., Neale, M. C., Clasen, L. S., Liu, S., Seidlitz, J., Pritikin, J. N., ... Raznahan, A. (2019). A comprehensive quantitative genetic analysis of cerebral surface area in youth. *Journal of Neuroscience*, 3028–3040. doi:10.1523/JNEUROSCI.2248-18.2019

Schnack, H. G., Van Haren, N. E. M., Brouwer, R. M., Evans, A., Durston, S., Boomsma, D. I., ... Hulshoff Pol, H. E. (2015). Changes in thickness and surface area of the human cortex and their relationship with intelligence. *Cerebral Cortex*, *25*(6), 1608–1617. doi:10.1093/cercor/bht357

Shulman, C. & Bostrom, N. (2014). Embryo selection for cognitive enhancement: Curiosity or game-changer? *Global Policy*, *5*(1), 85–92. doi:10.1111/1758-5899.12123

Spindel, J. E. & McCouch, S. R. (2016). When more is better: How data sharing would accelerate genomic selection of crop plants. *New Phytologist*, *212*(4), 814–826. doi:10.1111/nph.14174

van den Heuvel, M. P., Stam, C. J., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience*, *29*(23), 7619–7624. doi:10.1523/jneurosci.1443-09.2009

Vattikuti, S., Lee, J. J., Chang, C. C., Hsu, S. D. H., & Chow, C. C. (2014). Applying compressed sensing to genome-wide association studies. *GigaScience*, *3*, 10. doi:10.1186/2047-217X-3-10

Vilhjálmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., ... Price, A. L. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *American Journal of Human Genetics*, *97*(4), 576–592. doi:10.1016/j.ajhg.2015.09.001

Vuoksimaa, E., Panizzon, M. S., Chen, C.-H., Fiecas, M., Eyler, L. T., Fennema-Notestine, C., ... Kremen, W. S. (2015). The genetic association between neocortical volume and general cognitive ability is driven by global surface area rather than thickness. *Cerebral Cortex*, *25*(8), 2127–2137. doi:10.1093/cercor/bhu018

Walhovd, K. B., Krogsrud, S. K., Amlien, I. K., Bartsch, H., Bjørnerud, A., Due-Tønnessen, P., ... Fjell, A. M. (2016). Neurodevelopmental origins of lifespan changes in brain and cognition. *Proceedings of the National Academy of Sciences*, *113*(33), 9357–9362. doi:10.1073/pnas.1524259113

Willoughby, E. A., McGue, M., Iacono, W. G., Rustichini, A., & Lee, J. J. (under review). The role of parental genotype in predicting offspring years of education: Evidence for "genetic nurture".

Wray, N. R., Goddard, M. E., & Visscher, P. M. (2007). Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Research*, *17*(10), 1520–1528. doi:10.1101/gr.6665407

Wray, N. R., Kemper, K. E., Hayes, B. J., Goddard, M. E., & Visscher, P. M. (2019). Complex trait prediction from genome data: Contrasting EBV in livestock to PRS in humans. *Genetics*, *211*(4), 1131–1141. doi:10.1534/genetics.119.301859

Yengo, L., Robinson, M. R., Keller, M. C., Kemper, K. E., Yang, Y., Trzaskowski, M., ... Visscher, P. M. (2018). Imprint of assortative mating on the human genome. *Nature Human Behaviour*, *2*, 948–954. doi:10.1038/s41562-018-0476-3