In press, *SAGE Handbook of Evolutionary Psychology*

Edited by Todd K. Shackelford

Chapter Title:

BEHAVIOR GENETICS

Authors:

Nicole Barbaro[1] & Lars Penke[2]

[1]Oakland University, Department of Psychology

[2]University of Goettingen, Georg Elias Müller Institute of Psychology

Nicole Barbaro, Oakland University, Department of Psychology, Rochester, MI, 48309. United States.
n.barbaro.psych@gmail.com

Lars Penke, University of Goettingen Institute of Psychology, Gosslerstr. 14, 37073 Goettingen, Germany
lpenke@uni-goettingen.de

Abstract:

Behavior genetic research addressing how genetic variation and environmental variation explain complex phenotypes has instigated a dramatic paradigm shift in the field of psychological science in the last 50 years. The classic nature vs. nurture debate has slowly given way to a more nuanced view of nature and nurture that recognizes the complex interplay of genes and environment in explaining human psychology. More recently, genome-wide molecular genetic approaches have afforded more detailed and nuanced insight into how genetic variation can explain complex phenotypes. In this chapter, we provide an overview of behavior genetic methods, interpretations, and implications. We review an assortment of well-established behavioral genetic research designs used in the psychological and behavioral sciences across two broad categories of methods: classical family studies and genome-wide association studies. For each approach, we review the basic research design details, underlying assumptions of the approach, and proper interpretation of results produced by each approach. We conclude with discussion of three primary implications of behavior genetic methods, including how findings from behavior genetic research can inform causal inference, developmental processes, and understanding of the evolutionary history of traits. As advances continue in behavior genetic methodology, so too will our understanding of how genetics influences our psychology and behavior.

Key Words: behavior genetics, twin studies, genome wide association, heritability

BEHAVIOR GENETICS

Importance

Findings from the field of behavioral genetics has been a substantial paradigm shift for the study of psychological science (Plomin et al., 2013, 2016). It is now beyond reasonable debate that understanding the role of genetic variation for explaining complex phenotypes is necessary for informed study of psychology and behavior. Genetic variation is essential to the process of evolution by natural selection. The idea that heritable variation provides the raw material upon which selection can act has long been acknowledged in the biological life sciences. Humans – and their brains – are not exempt from evolutionary processes. Any approach to understanding psychology that fails to take into consideration that genetic variation underpins the foci of psychological study is, therefore, untenable and not a viable approach to the conduct of psychological science (see Penke, 2011).

The claim that human psychological traits are heritable is supported on a substantial scale (e.g., Plomin et al., 2016; Polderman et al., 2015). Although this basic understanding of the widespread heritability of psychological and behavior traits has been prominent since the rise of the field in the 1970s, the methods used to understand genetic (and environmental) influences on traits has substantially advanced in the era after sequencing the human genome, beginning in the early 2000s (International Human Genome Sequencing Consortium, 2001; Venter et al. 2001). With these methodological advancements came a better understanding of the complex interplay of genes and environment that produce phenotypes. As methods and interpretations have become increasingly complex, however, basic understanding of behavior genetics has become increasing difficult as well.

Our goal in this chapter is to provide readers with a broad overview of the popular methods within the domain of classical family studies (e.g., twin and adoption studies), the foundation of behavior genetics, and within the domain of contemporary approaches using genome-wide approaches, which are leading to exciting and informative insights into the evolutionary history of traits. For each method, when relevant, we provide an overview of the research design, key aspects and terms, and the major interpretations of the results produced by the method. Finally, we conclude with three major implications of behavior genetics for psychological science, including inferring causality, inferring developmental processes, and inferring the evolution of traits.

Methods

The field of behavior genetics uses myriad research designs. Classical behavior genetics, which emerged in the 1970s, uses natural quasi-experiments afforded by twinning and adoption to infer the extent to which observable phenotypes are explained by genic variation and environmental variation.

Contemporary behavior genetics also uses genome-wide approaches to identify specific genetic markers associated with phenotypes. The history of these methodologies occur in parallel with the history of genetics more broadly, beginning with the discovery of DNA structure in 1964, experiencing a paradigm shift with the sequencing of the human genome in 2001, and elevating to public discourse in the current era of personal genomics. Classical approaches and genome approaches to studying the complex relationships between genetic variation and observable phenotypes differ in fundamental aspects with regard to design, assumptions, and implications. Each approach, however, is aimed at the same overarching goal: to advance understanding of how our genes contribute to who we are. This section will focus on two major domains of behavior genetic methods –family studies and genetic association approaches – outlining basic design features, assumptions, and reasonable interpretation of results.

Classical Family Studies

Twins are one of nature's greatest experiments, and the foundation of classical behavior genetics. Twins come in two types: monozygotic (MZ) and dizygotic (DZ). Because MZ and DZ twins differ in their average degree of genetic relatedness, the two genetic types of twins allow for inferences about the sources of influence on phenotypes. MZ twins result from a fertilized egg splitting very early in development, which means that their DNA is identical. DZ twins, in contrast, develop from different fertilized eggs, meaning that they are no more alike or different than a random pair of siblings from the same parent; DZ twins are simply siblings born on the same day. MZ twins share 100 percent of their DNA, whereas DZ twins share approximately 50 percent of their DNA, on average. Knowing parameters of average genetic relatedness, in combination with natural quasi-experimental situations caused by adoption and twinning, allows for inferences about the contributions of genes and environment to phenotypes of interest.

The contribution of nature (genes) and nurture (environment) to observable phenotypes has been a point of debate for more than a century. Children that grow up in the same family tend to resemble one another in many aspects. For most of psychological science's history, most prominently in psychoanalytic and behavioristic traditions, similarities between family members were presumed to be due primarily to the fact that they shared a family environment—resources, experiences, parenting, or neighborhood, for example. Utilization of twinning and adoption led to simple extrapolation about the likely degree to which nature and nature could account for family resemblance. If family resemblance is largely due to growing up in the same household, then it can be reasoned that: (1) adoptive siblings, who are not genetically related, should be just as similar as genetic siblings that grow up in the same house, and (2) MZ twins that grow up in different homes due to adoption in infancy, should be no more similar than two people drawn at random who grow up in different families.

These predictions were not borne out by observation, however. Early studies of twinning and adoption found precisely the opposite of what is predicted based on family environment causing similarity. Twins that are reared apart are nearly as similar on psychological and behavioral measures as twins reared together, and by late adolescence adoptive siblings are no more similar than two randomly selected children reared in different homes. These findings suggest that family resemblance is at least in part due to the contributions of shared genes, rather than to the contributions of shared environment. Twinning and adoption studies comparing the relative similarity between children reared together and apart are useful for understanding in broad strokes the contributions of genes and environment, but more precise estimates are produced from quantitative models utilizing MZ and DZ twins.

A fundamental goal of psychological science is to explain variation. Twin studies explain variation in a phenotypic trait, but in a different and complementary way to standard psychological science methods. Twin, adoption, and family studies, collectively known as quantitative genetic studies, can produce two primary outcomes of interest: (1) partitioning the variation of a trait into (usually) the three components of genetic variation, shared environmental variation, and non-shared environmental variation; and (2) partitioning the covariation between two traits – referred to as a *phenotypic association* – into (usually) the three components of genetic variation, shared environmental variation, and non-shared environmental variation.

Partitioning Variance. Correlation comparisons and ACE models (A meaning genetic variance, C meaning shared environmental variance, and E meaning non-shared environmental variance) are common approaches used to partition sources of variance for a particular trait in a given population. That is, given a population of individuals, univariate models can estimate what proportion of individual differences within the population for a phenotypic trait—for instance, extraversion—are attributable to genetic variation within the population, and what proportion is attributable to environmental variation (shared and nonshared) within that population. Phenotypic variance is composed of three factors, two of which are distinct environmental factors. Genetic variance explaining phenotypic variance is referred to as heritability (h^2), which acts to make two individuals who share more genes more similar to one another than two individuals who share fewer genes. Environmental influences on phenotypic variation are comprised of two components: shared environment (c^2) and nonshared environment (e^2). The environmental components (collectively referred to as *environmentality*, $c^2 + e^2$) refer to phenotypic variance accounted for by nongenetic, or "environmental" experiences, in a broad sense. Shared environmental variation, such as family-level variables, are aspects of the environment that make siblings (or others) reared together more similar to one another. Nonshared environmental variation, such as unique peer groups or stochastic developmental variation, are aspects of the environment that make

siblings (or others) reared together dissimilar from one another (the nonshared component usually also includes measurement error).

The classical method to estimating the contributions of genetic and environmental variance to a phenotype is to compare correlations between MZ and DZ twins to produce estimates of variance components: genetic, shared environmental, and nonshared environmental. The only information required is the average correlation between MZ and DZ twin pairs for a particular trait. Calculations can then be made to estimate the degree to which each component explains the phenotypic trait (Purcell, 2016). To calculate the contribution of genetic variation, h^2, one calculates the difference between the MZ correlation and the DZ correlation and then multiplies by two, $h^2 = 2\ [r_{mz} - r_{dz}]$. The contribution of shared environmental variation can be calculated by subtracting the MZ twin correlation from the genetic contribution calculated above, $c^2 = r_{mz} - h^2$, or by subtracting half the genetic contribution from the DZ twin correlation, $c^2 = r_{dz} - [h^2/2]$. The nonshared environmental component is then calculated by subtracting the genetic and shared environmental values from one, $e^2 = 1 - [h^2 + c^2]$.

The above method based on the correlation coefficients between MZ and DZ twins will produce only one estimate. For example, a trait with $r_{mz} = .60$ and a $r_{dz} = .35$ will yield only $h^2 = .50$, $c^2 = .10$, and $e^2 = .40$. There is no other answer to the calculations. Structural equation modeling can also be used to estimate heritability and environmentality, applying the popular ACE models. The benefit of using structural equation models is that models can be compared to one another to examine, for example, if the $c^2 = .10$ value significantly contributes to the overall understanding of the trait of interest. Does removing the c^2 value from the model (the equivalent of fixing c^2 to zero) substantially affect the fit of the model to the data? Science favors parsimony, such that the simplest explanation should be accepted. ACE models typically test full models including each component, and additionally specify simpler combinations, AE, AC, CE, A, C, E, and compare the fit of those models to the full model. If, for example, the AE model is a better fit, or does not significantly impact the model fit compared to the full ACE model, the simpler model will be accepted, such that the contribution of shared environmental variance will be specified as 0, thus changing the relative contributions of genetic and nonshared variance.

ACE models, like any statistical model, rely on assumptions to produce estimates of heritability and environmentality for a particular trait. First, the models assume that MZ twins share 100 percent of their DNA and that DZ twins share 50 percent of their DNA. These assumptions of genetic relatedness are reflected in the model specifications, such that the A factors, or genetic latent factors, are set to correlate at 1.0 in MZ twins, and 0.50 in DZ twins. In theory, both of these assumptions are true, but in practice MZ and DZ twins genetic relatedness can vary around these assumed values. MZ twins, for example, are derived from the same fertilized egg, meaning that they start with exactly the same genome, but later mutations, genetic expression, and epigenetic modifications can result in minor genotypic differences

between MZ twins, though only in parts of the body (Charney, 2012). The assumption that DZ twins share 50 percent of their DNA reflects a population average, with specific samples, especially smaller samples, more likely to vary above or below the 50 percent average. Mathematical simulations to investigate the impact of deviations from genetic relatedness assumptions have shown that they have little impact on accurate estimation of heritability (Lui et al., 2018). Violation of the assumptions appear to have the largest effect on phenotypes of high true heritability, underestimating heritability estimates up to 10 percent, and correspondingly inflating nonshared environmental estimates; estimates of phenotypes of low true heritability are less likely to be affected (Lui et al., 2018).

Variance partitioning approaches also assume that the environmental experiences of MZ twins are no more alike than the environmental experiences of DZ twins, which is referred to as the *equal environments assumption* (Bhattacharjee & Sarkar, 2017; Scarr, 1968). Critics of twin studies argue that MZ twins are socialized to be more similar, and have more similar experiences because of their zygosity, which may bias heritability estimates upward (e.g., Joseph, 2004). Several types of studies, however, provide support for the equal environments assumption. Results of studies examining the similarities of MZ twins reared apart show that reared-apart MZ twins are just as similar as reared-together MZ twins. Other research has investigated whether true zygosity or perceived zygosity influences similarities between twins (Kendler et al., 1993), finding no association between perceived zygosity on trait similarity. Research utilizing doppelgangers (Segal et al., 2018), who are not genetically related, support this idea, indicating that people are treated similarly because of their heritable traits, not because of their zygosity. That is, people possessing particular heritable traits *evoke* similar reactions from other people. MZ twins are genetically identical and therefore evoke similar reactions from others, rather than others treating them similarly because of their perceived zygosity.

Many highly powered variance partitioning studies have produced robust and replicable results (Plomin et al., 2016; see also Stanley et al., 2018). These results from twin methods have culminated in what are referred to as the laws of behavior genetics (Turkhiemer, 2000). The first law of behavior genetics is that all complex phenotypes are, to some degree, heritable. The strongest evidence in support of this claim comes from a meta-analysis of twin studies on nearly 18,000 traits (e.g., intelligence, personality traits, health, relationships) – the largest behavioral genetic analysis on psychological phenotypes published to date (Polderman et al., 2015). Polderman et al. (2015) reported an average heritability of 49% across all complex human traits evaluated, supporting the first law. The results also showed that, on average, approximately 17% of variation across complex human traits is attributable to shared environmental variance. The finding that shared environment explains relatively less phenotypic variance than does genetic variation supports the second law of behavior genetics: that shared genes largely drive similarities between biological relatives, rather than shared environments. Finally,

Polderman et al. show that the remainder of the variance of phenotypic traits, on average, is attributable to nonshared environmental variance. Because shared genetic variation and shared environmental variation function to explain similarities between individuals, that a substantial proportion of phenotypic variation is nonshared implies that unique experiences and random stochastic developmental variation are what explains individual differences, which is known as the third law of behavior genetics.

Partitioning Covariance. A foundation of psychological science is to understand associations between two variables, or phenotypes, of interest; for example, understanding the association between early developmental stress and psychosocial outcomes. Such associations are referred to in behavior genetics as *phenotypic associations*. Phenotypic associations are correlational associations or regression coefficients reported in standard psychological research. What behavior genetics can add to standard phenotypic associations is the decomposition of the phenotypic association into the three components of genetic covariation, shared environmental covariation, and nonshared environmental covariation. That is, behavior genetic methods can identify the extent to which a reported correlation is explained by different factors, yielding informative insights into the nature of associations. Several statistical methods, often using structural equation modeling, can be used to partition covariance, and are explained in detail elsewhere (Turkheimer & Harden, 2014). Here, we will focus on the importance of this approach for psychological research, and the interpretation of partitioning covariance outcomes.

Establishing causal relations between traits or behaviors is a laudable goal of psychological research, yet one that is littered with myriad ethical and methodological obstacles. Given these obstacles, the social science model necessitates that three requirements are met to state that evidence is consistent with a causal interpretation: (1) two variables are related; (2) appropriate temporal sequence between the variables; and (3) relevant confounds are accounted for. Such approaches are motivated by a desire to establish a causal effect: that the change in one variable will cause a change in another variable. Behavior genetic methods are more aimed at identifying the causal structure of phenotypic associations (Briley et al., 2018): partitioning the covariance into different components of influence. Both causal goals are related. Once causal structure is identified, more precise investigation of causal effects can be pursued, for instance, by investigating genetic mechanisms or family-level factors.

Because of the first law of behavior genetics – that all complex phenotypes are heritable – it is reasonable to suspect that genes play some role in the architecture of phenotypic associations (Briley et al., 2018). Genetic covariation can affect two traits independently, producing a spurious phenotypic association. This particular problem, known as *genetic confounding*, is akin to the third variable problem in psychology. Identifying genetic confounds of phenotypic associations is necessary to accurately identify causal phenotypic effects. For example, if an association is found between depression and

anxiety, but the association is predominantly accounted for by shared genetic covariation as described above, changing levels of depression may not produce reliable changes in anxiety. Controlling for genetic covariation is necessary to identify true effects between phenotypic associations, as is often the goal of mainstream psychology. As the genetic covariation between two traits, r_g, increases, the more likely is it that genetic covariation may be confounding phenotypic associations (Barnes et al., 2014).

The findings of shared and nonshared environmental covariance, r_c and r_e respectively, are more nuanced in their interpretation (Turkheimer & Harden, 2014). A result of a significant proportion of a phenotypic effect being attributed to shared environmental covariance is demonstrating that between-pair differences are accounting for a proportion of the effect. A robust example of this is the positive association between spanking and physical abuse (Jaffee et al., 2004a). Covariance partitioning indicates that the majority of the phenotypic effect between spanking and physical abuse is attributable to shared environmental covariation, meaning that between-family differences, rather than within-pair (i.e., parent and child) differences account for the association. This contrasts with findings from the same data showing that the association between spanking and externalizing problems is attributable to genetic covariation and nonshared environmental covariation (Jaffee et al., 2004b). What this means is that after accounting for estimated genetic confounds, the remainder of the effect of spanking on subsequent externalizing problems is attributable to within-pair differences via nonshared environmental covariation, and not between-family differences. That within-pair differences, or nonshared environmental covariation, explains substantial proportions of phenotypic effects is, in fact, the strongest evidence for (quasi)causal effects that psychologists are often looking for (Turkheimer & Harden, 2014). One of the most important benefits of behavior genetic methods is the identification of environmental covariance to infer (quasi)causal phenotypic effects from correlational data.

Extended Family Designs. Extended family designs provide a method to identify environmental sources of variance on outcomes while controlling for genetic relatedness. Such designs allow for partitioning variance into genetic and environmental components to address questions of intergenerational transmission. Just as classical covariance partitioning methods, described above, can decompose the variance between two phenotypic traits, extended family designs have the same purpose, but are focused on phenotypic associations between generations. Phenotypic associations between generations, often between parents and children, are ubiquitous and foundational to many psychological theories of development, making extended family designs valuable to understand the causal architecture and directionality of intergenerational transmission (McAdams et al., 2014).

One popular extended family design is the Children of Twins (CoT) design, for which MZ and DZ twin parents and one child for each parent are assessed. CoT models allow for control of genetic

covariation between parents and offspring, with the primary outcome being a path estimate between the parent and offspring phenotype. A limitation of the one-child CoT design is a lack of an estimate for shared environmental covariation on offspring phenotypes given that the offspring are assumed to reside in different environment. A difficulty with the CoT (or, by extension, Children of Siblings) design is the lower power to detect the sources of variation in offspring traits (McAdams et al., 2018). The relatedness coefficients between the targets, the offspring, are cousin-level (0.25 and 0.125), rather than sibling-level (1.00 and 0.50) as in classical twin models, meaning that the coefficients themselves are lower and the difference between them smaller, which causes reduced statistical power to detect effects. Power analyses indicate that nearly 1,000 families are needed to accurately detect a phenotypic association between parent and offspring assuming moderately heritable traits (h^2 = .35; McAdams et al., 2018).

Newer models extending the one-child CoT designs can more accurately model relationships between parents and offspring. McAdams et al. (2018) proposes using a multiple CoT (MCoT) design, whereby at least two children of each twin parent are assessed. The addition of multiple children allows for modeling shared environmental variation on offspring phenotypes that is not available in the classic CoT model. Moreover, MCoT designs can appropriately model individual relationships between parent variables and offspring variables. Research on parent-child interactions has shown that children shape parenting behavior, therefore violating the assumption that parent variables that psychologists are most often interested in are invariant across children. For variables such as emotional sensitivity, each child may evoke more or less sensitivity from the same parent. MCoT designs (McAdams et al., 2018) can model such relationships, yielding more accurate estimates of parent-offspring associations. Another benefit of MCoT designs, relative to classic CoT designs, is that fewer families are needed (around 500) to achieve adequate power (McAdams et al., 2018). Although nearly as many individuals are needed (more children per family, for example), the fact that fewer families need to be recruited could be a benefit for executing high-powered extended family designs.

Longitudinal Designs. Behavior genetic models assume that genetic and environmental influences operate across development to give rise to observable phenotypes. Gene-environment interplay can take various forms across development (Briley et al., 2019; Scarr & McCartney, 1983). Take the associations between parenting and offspring outcomes. Cross-sectional designs, such as the CoT designs described previously, can partition covariance to identify the true phenotypic association between parent and offspring. The design, however, cannot inform the directionality of the association, or the nature of the association over time. Consider, for example, the association between harsh parenting and offspring externalizing behaviors. A cross-sectional CoT design can yield a phenotypic association between parenting and externalizing behaviors that controls for the genetic relatedness between the parent and

offspring, but it does not provide information as to whether the harsh parenting led to externalizing behaviors or whether the child's externalizing behaviors prompted harsh parenting (or both).

Longitudinal designs can offer insight into how parents affect children, and how children affect parents. Children are not born as blank slates on which parents unidirectionally shape child outcomes. Children are born with capacities that shape the parenting they receive. Children are also not randomly allocated to environmental experiences (Plomin et al., 1977; Scarr, 1982). Environments provided by parents are influenced by their heritable qualities, some of which are passed down to offspring (Kendler & Baker, 2007). Finally, mating is non-random, meaning that traits of reproducing partners can be correlated both phenotypically and genetically (Yengo et al., 2018).

That environments are nonrandomly distributed across individuals and are correlated among genetically similar individuals is referred to as *gene-environment correlation*, or *r*GE. *r*GE can take several forms across development: active, evocative (or reactive), and passive. *Active r*GE occurs when individuals actively seek out, avoid, or modify their environmental experiences that are nonrandomly influenced by their genotype. *Evocative r*GE occurs when organisms receive responses or evoke reactions from others in their environment that are nonrandomly influenced by their genotype. *Passive r*GE occurs when the environment that an individual inhabits—such as the neighborhood a child grows up in—is correlated with their genome. Parents endow offspring with an environment in which to live, and a genome comprised of half of each (biological) parent's genes, such that the environments children experience are correlated with the genotypes that they inherit from their parents (Kendler & Baker, 2007).

The dominant type of gene-environment correlation is proposed to change over the course of development (Scarr, 1992). Passive gene-environment correlation has greater explanatory power in infancy and early childhood. Because human infants are heavily dependent on caregivers during the first years of life, evident gene-environment correlations are most likely the passive type given the control caregivers have over children's environment. The importance of active gene-environment correlation increases with age, as individual decision making and environmental control also increase. The implications of this change in dominant gene-environment correlation type over development can, in part, explain the general increase in heritability estimates of a myriad of traits (most notably, intelligence) over development (Plomin et al., 2016).

The interplay of genes and environments is not simple or straightforward, but rather complex and dynamic (Briley et al., 2019). A comprehensive understanding of gene-environment interplay for understanding phenotypic associations therefore requires that genetic variance is accounted for in research designs, especially longitudinal designs in which genetically related individuals are the targets of focus. In much developmental psychology research, such targets are often parents and offspring, who share, on

average, 50% of their DNA. Failing to account for the genetic relatedness between family members may produce biased phenotypic effects of developmental processes across the lifespan.

Genetically-sensitive longitudinal designs can provide a nuanced understanding of purported parenting effects on child development, for example. The negative developmental impacts of spanking on child developmental outcomes is a contentious topic of research, of which most centers on the negative effects of parenting on children. A behavior genetic perspective, however, offers additional insights into understanding the dynamic nature of such phenotypic associations. A longitudinal twin study conducted by Cecil et al. (2012) is particularly suited to demonstrate evocative child effects of spanking, which are suggested by the results of the above genetically-informed studies. Using cross-lagged panel analyses, Cecil et al. examined whether harsh punishment (i.e., smacking and shouting) was associated with self-control difficulties from early childhood to adolescence. The results indicated bi-directional effects of harsh punishment and self-control; but, interestingly, between seven and 12 years, only evocative effects were found such that self-control difficulties predicted later harsh punishment, but harsh punishment did not predict later self-control difficulties. Cecil et al. did find long-term effects of harsh punishment on early adolescent conduct problems, but for only boys. These results are consistent with the general notion that children's individual behavior can evoke or exacerbate parental punishment, with the unique contribution that the effects of harsh punishment may be particularly relevant in early childhood, as opposed to later childhood and adolescence (Cecil et al., 2012).

Genetically-sensitive longitudinal designs can be utilized to understand how variance components change over time, and can provide insight into which components are responsible for stability and change over development (Briley & Tucker-Drob, 2017). Each trait investigated is unique in many regards, but common among well-studied traits is the finding that shared environmental variance decreases rapidly across development (Plomin et al., 2016; Briley & Tucker-Drob, 2017), whereas non-shared environmental variance and genetic variance increase or remain relatively stable. Longitudinal behavior genetic designs are also capable of decomposing phenotypic stability into genetic and environmental components. Essentially, these models identify whether the genetic or environmental variance of a phenotype at one time point is the same variance that influences the phenotype at another time point. Specifically, variance components are correlated across time points. If the genetic variance component for a trait at time 1 are highly correlated with the genetic variance component for a trait at time 2, this indicates that genetic variance is, in part, responsible for the observed phenotypic stability of the trait. Collectively, longitudinal behavior genetic designs can provide a wealth of information on phenotypes over development.

Genetic Association Approaches

Classical family studies dominated the behavior genetic literature until the 21st century, largely because such designs were the most accessible to researchers. The sequencing of the human genome in 2001 propelled a new era of behavior genetics research; an era that is rapidly advancing and changing. At the beginning of this transition, scientists worked under what we now know to be false assumptions of how genes are associated with complex phenotypes. The goal was to find "genes for" a particular trait, with the assumption being that traits would be underpinned by a few genes with relatively large effects, and that these genes would explain the genetic variance for traits that had been indicated by decades of family studies.

The focus on finding "genes for" a particular trait produced a cascade of research in the early 2000s. Linkage and candidate gene approaches were applied to complex behavioral traits because such approaches had been so successful at identifying genes that contributed to Mendelian disorders – disorders in which a single gene can be identified by following traditional inheritance patterns, which are more common for some medical diseases. The "genes for" approach to complex phenotypes was bolstered by early high-profile candidate gene studies, such as those on depression (Caspi et al., 2002, 2003). Enthusiasm for candidate gene approaches waned as the number of failed replications increased, including the high-profile candidate gene findings for depression (Border et al., 2019). It is now the consensus view that candidate gene approaches to understanding complex phenotypes are flawed and insufficient given what is now know about the genetic nature of complex phenotypes (Dick et al., 2014). Virtually all phenotypes of relevance for psychologists and behavioral scientists belong in the category of complex phenotypes.

Genome wide association studies (GWAS) became increasingly popular as their cost dropped through the 2000s (Mills & Rahal, 2019). GWAS identify associations between genes and phenotypes, as do candidate gene and linkage approaches, but do so across the entire genome. GWAS utilize genetic markers known as single nucleotide polymorphisms, or SNPs, across the genome to identify genetic regions that are associated with an outcome of interest. SNPs are variations in DNA building blocks (either A, C, T, or G), whereby, say, a C might be replaced by a T at a certain locus in a gene or stretch of DNA. These variations are normal in DNA, but particular variations may be relatively more or less common in populations of individuals. Catalogs of known SNPs (more than 100 million have been identified) are used as reference when correlating SNPs with phenotypes of interest. To date, nearly 4,000 GWAS have been conducted on thousands of traits (Mills & Rahal, 2019). The findings of which have greatly informed the field of behavior genetics.

Genetic Associations with Phenotypes. Family studies, described above, provided the foundation for what are known as the "three laws of behavior genetics" (Turkheimer, 2000). GWAS has

contributed to what is now known as the “fourth law” of behavior genetics: complex phenotypic traits are highly polygenic (Chabris et al., 2015). Polygenic means that the phenotypic trait of interest has hundreds to thousands of genetic associations. The thousands of GWAS that have been published over the preceding two decades have revealed approximately 10,000 robust associations between genetic markers and various traits and disorders (Visscher et al., 2017), using a purely exploratory approach. The polygenic nature of complex traits sits in stark contrast to historical assumptions of candidate gene and linkage approaches that predicted just a few genes would be found underpinning each trait.

This fundamental insight to the polygenic nature of complex phenotypes also changed the way in which the effects of genes on complex phenotypes are understood. Whereas candidate gene approaches, for example, assumed that identified genes would have substantial effects and would thus explain the genetic variance estimated from family studies, GWAS approaches were increasingly producing two novel insights. (1) Identified genetic variants have very small effects, often with each genetic variant explaining much less than one percent of the variance of a phenotype; and (2) GWAS approaches struggled to explain the genetic variance estimated from family studies, with such approaches yielding proportionately small heritability estimates, a problem referred to as *missing heritability*.

The small effects of individual genetic associations with phenotypes have implications for how behavior genetic research is conducted. Because many genetic loci contribute to the genetic variance of a trait, and thus the effect of each genetic variant is very small, it then implies that GWAS sample sizes must be very large to find reliable and accurate genetic associations. Indeed, the average sample size of GWAS has increased substantially since the first GWAS in the early 2000s, with some of the largest studies using data from over 1 million people (Mills & Rahal, 2019). Because larger samples have greater power to detect small effects, such as the type of effects GWAS are searching for, the number of associations found between genetic loci and complex phenotypes has rapidly increased over time as well. Analyses of the literature show that as sample sizes of GWAS have increased, so too have the number of associations found, and the number of traits studied. SNPs from GWAS can also be used to estimate a heritability based on this genetic data, free of the assumptions underlying family studies (as discussed above) (Yang et al. 2010). This SNP-based heritability forms the natural upper boundary for how much variance of a trait can be explained by SNP associations from GWAS. Currently, there seems to be no plateau of GWAS associations with increasing sample sizes, suggesting that continuing to increase sample sizes will continue to identify novel genetic associations and approach the level of explained variance set by the SNP-based heritability (Visscher et al., 2017).

However, estimates of SNP-based heritability from GWAS are often substantially smaller than heritability estimates from family studies (usually around half as large). This peculiar part of the missing

heritability is a gap that cannot be expected to close just by increasing GWAS sample sizes. It is also a consequence of the polygenic nature of complex traits: Because SNP catalogs that are used as reference panels for GWAS are by nature not complete, and extraordinary sample sizes are needed to detect effects, by design, substantial portions of heritability estimates from family studies will not be accounted for in average GWAS designs. As sample sizes have increased, and SNP catalogs have expanded, the proportion of additive genetic variance from SNP based GWAS is continuing to increase. New methods are also being developed to "find" the missing heritability.

Whole genome sequencing. SNPs used in reference panels for GWAS as described above do not cover the entire genome, with many SNPs inferred or imputed based on linkage disequilibrium, which is the correlation of a SNPs being associated with another SNP based on its location in the genome. Moreover, the SNPs used in reference panels tend to be "common" variants found in at least 1% of the population (Visscher et al., 2017). These design features therefore mean that many SNPs will not be analyzed, which includes rare genetic variants occurring in less than 1% (and often much smaller proportions) of the population. Rare variants contribute to the problem of SNP heritability estimates being smaller than family study heritability estimates.

Whole genome sequencing methods are being developed and refined to address these limitations of traditional SNP based GWAS to "find" the missing heritability of complex phenotypes. Many whole genome sequencing methods have been developed in recent years, with the utility of each depending on the specific genetic nature of the trait being assessed (Evans et al., 2018). Because whole genome sequencing methods can capture more SNPs at lower frequency levels than traditional GWAS approaches, whole genome sequencing studies have found that typical SNP heritability estimates may be underestimated by ~20% for most traits due to non-inclusion of rare variants (Evans et al., 2018).

Rare variants have become increasingly important for understanding and estimating the heritability of complex phenotypes and for explaining substantial amounts of heritability estimates found in family studies. Rare or low frequency variants, in contrast to the common variants tagged as SNPs in typical GWAS, have a greater impact on the outcome of interest, meaning that each rare variant explains a greater proportion of variance in the phenotype than does each common variant. Common variants, for example, often explain less than 0.5% to 1.0% of variance in a phenotype, whereas the effect of rare variants can be up to ten times larger than common variants in some cases (Manolio et al., 2009; Marouli et al., 2017). The heritability of traits such as intelligence and educational attainment has been much more accurately estimated using indirect approaches to capture rare variants, with family-specific (i.e., rare) genetic variants being implied in half of the heritability of these traits (Hill et al., 2018).

GWAS approaches, including whole genome sequencing methods, have unambiguously demonstrated the high polygenic nature of complex phenotypes, such that hundreds or thousands of genetic variants underpin traits of interest. Alternatively, some suggest that rather than polygenic, continued research could reveal that complex phenotypes will be more accurately described as "omnigenic" with essentially all active genes being associated with every complex trait (Boyle et al., 2017). Whole genome sequencing methods (see Evans et al., 2018) continue to utilize the full breadth of the information available to geneticists to uncover heritability of complex phenotypes and solve the missing heritability problem prompted by earlier genetic association methods (Génin, 2019). It is likely that methods will continue to improve resulting in continually increasing accuracy of heritability.

Polygenic Scores. GWAS has demonstrated that many genes of very small effects underpin phenotypes; therefore, using the genome to predict phenotypic outcomes necessarily needs to be based on the effects of multiple genes (Turkheimer, 2015). As GWAS discoveries increased, aggregate genetic "scores" started being used as predictors of phenotypes. Most commonly referred to as *polygenic scores*, these genetic predictors are aggregate genome-based calculations given to an individual based on their genotype. Polygenic scores are genome-wide weighted averages of significant SNPs from independent GWASs (Belsky & Harden, 2019). Although specific methods for constructing polygenic scores vary, each follow a general approach. First, significant SNP associations from a GWAS are identified and weighted, which is referred to as the discovery sample. Next, using participants that were not included in the discovery sample, individual polygenic scores are calculated by summing the weighted SNP alleles (Belsky & Harden, 2019).

Polygenic scores can be calculated for any phenotypic trait for which GWAS is available. Once polygenic scores are calculated for individuals within a sample, the polygenic scores can be used as a predictor variable in psychological research using standard statistical modeling techniques. Although polygenic scores can be useful at the population level, particularly for understanding how relatively high or low genetic "risk" for a trait is associated with outcomes, on average, polygenic scores are so far nearly useless for prediction at the individual level (Belsky & Harden, 2019; Turkheimer, 2015). Polygenic scores are currently capable of explaining a few percentage points of variance in most outcomes, with the highest ranges at 3%-15% variance explained depending on the trait (Visscher et al., 2017). Moreover, because polygenic scores are built from GWAS, the limitations of GWAS for finding genetic associations surrounding sample size, missing heritability, and power, are carried over into the limitations of polygenic scores – the score is only as good as the GWAS it is based on, at best.

Additional problems arise regarding the predictive utility of polygenic scores across heterogeneous populations. Allele frequencies between human populations can vary for many reasons

(e.g., genetic drift), leading to what is referred to as *population stratification*, or systematic differences in allele frequencies between populations. Because polygenic score construction relies on a discovery GWAS sample that is different than the test sample, and the majority of GWAS are from European populations (Mills & Rahal, 2019), using European-derived polygenic scores with diverse populations, such as those with Asian or African ancestry, results in poor predictive utility of polygenic scores in the test samples. For example, an analysis by Duncan et al. (2019) showed that the median effect size of polygenic scores in African ancestry samples was only 42% of that in a matched European ancestry sample, demonstrating the limitations of polygenic score generalizability to non-European samples. For polygenic scores to be of use in diverse human samples, GWAS must be representative of human populations.

Genetic Correlation between Phenotypes. Given that GWAS approaches have demonstrated that complex phenotypes are highly polygenic, there is the logical implication that genetic variants associated with one trait are likely to be associated with another trait. In other words, *pleiotropy* – genetic variants associated with more than one phenotypic trait – is pervasive (Visscher et al., 2017). Evidence for pleiotropy comes from a variety of behavior genetic methods estimating genetic correlations, r_g. Family studies, for example, can calculate r_g by correlating latent genetic factors for two phenotypes in twin models; but these correlations do not speak to the specific genetic variants underpinning the correlation. GWAS approaches can also calculate r_g similarly to family studies, but can also identify specific genetic variants that are associated with multiple phenotypes. As with other GWAS research goals, large sample sizes, approaching 100,000 participants depending on the precise method, are needed to achieve adequate power and produce accurate r_g estimates (Rheenen et al., 2019).

Additionally, whereas family studies require measurements on both phenotypes of interest from the same participants to calculate r_g, GWAS can estimate genetic correlations and identify underlying genetic variants when phenotypes are measured from different individuals (Visscher et al., 2017). Limitations of family studies for calculating r_g are especially challenging for low frequency phenotypes such as disease traits or psychopathologies. GWAS can overcome this limitation by combining information from large data sets that have been independently collected for specific disease traits. Such approaches to measure pleiotropy using GWAS, in fact, exclude closely related individuals (such as siblings or cousins); excluding close relatives from the datasets have the benefit of largely eliminating shared environmental confounds that are present in family studies, which may bias the genetic variance estimates (Rheenen et al., 2019).

Whereas GWAS methods for measuring pleiotropy avoid shared environmental confounding, researchers need to be cautious of other potential sources of confounding. Population stratification due to

genetic drift, non-random mating, and geographic isolation can lead to differences in allele frequencies between populations. Such population differences can bias genetic correlation estimates between traits. Measuring the same phenotypic traits in two different populations may yield different genetic correlations between two disease traits. Such a finding may be indicative of a gene-environment interaction, whereby genetic variants are expressed differently depending on the environment; or in contrast, the effect could be reflective of the population structure rather than anything to do with the functional associations between the traits themselves (Rheenen et al., 2019). For example, a disease that is relatively rare in one population may not show the same genetic association with other phenotypes as in a population where the disease is more prevalent. The discrepant genetic correlations, however, may not reflect genetic architecture differences between populations, but instead could simply be a result of the analysis not being able to accurately detect the low frequency genetic variants in the population where the disease is rare.

Genetic correlations provide useful information for researchers. Non-zero genetic correlations between two traits can help identify new risk factors for psychiatric conditions, for example, and yield insights into the genetic architecture between traits that may not have previously been obvious from standard social science methods that measure only phenotypes. For example, the understanding that Bipolar 1 disorder, characterized by mania, is more genetically similar to schizophrenia spectrum disorders than major depressive disorders (Coleman et al., 2019) was instrumental for changing the classification of Bipolar disorders in the fifth edition of the Diagnostic Statistical Manual of Mental Disorders to their own class, rather than a class shared with depressive disorders (APA, 2013).

Implications

Behavior genetic methods have substantively advanced in the preceding decades, most notably in the genomic era. This chapter is not an exhaustive description of all behavior genetic methods used in the social and behavior sciences, but is intended to be informative for a broad scientific audience interested in knowing what behavior genetic methods are available and, importantly, what the results of such methods mean for our understanding of psychology and behavior. In addition to the broader goals of behavior genetics – partitioning variance and covariance, and identifying genetic associations with phenotypes – these diverse methods can address pressing questions in the social sciences, especially regarding causality, development, and evolution, each of which will be briefly discussed.

Inferring causality.

Causal inference is a primary goal of social and behavioral science, although it is one that requires high standards of evidence to achieve. Behavior genetics, more broadly, contributes to improved

causal inference in psychological science (although, still imperfect). Inference of causality of an association between two traits, *X* and *Y*, can be achieved in two ways. First, by conducting a randomized controlled experiment to remove any confounding influences on the association between X and Y can instill greater confidence that X → Y. However, not all associations psychologists are interested in can be achieved by such methods. Also, not every individual is equally likely to naturally find themselves in each kind of situation that is simulated by experimental conditions (the very logic of gene-environment correlations), making random assignments to experimental conditions intrinsically artificial (Johnson & Penke, 2014). Causality can also be reasonably inferred by the removal of all relevant confounds of the association between X → Y (Pearl & Mackenzie, 2018). Whereas traditional psychological approaches often strive to include relevant controls in non-experimental designs, oftentimes genetic confounds are not included. Given that on average half of the variance in any trait of interest is genetic (Poldermann et al. 2015) and genetic correlations are ubiquitous, this is a major confound to ignore. Behavior genetic methods, such as family designs discussed above, offer a means to partition out genetic variance explaining the covariance between X → Y to better understand the phenotypic association and more accurately identify environmentally mediated effects, which are prime targets for interventions.

Behavior genetics knowledge and methods are also relevant for a broad understanding of the ways in which we think about causal models in psychological science. Because genes are the starting point from which phenotypes and behavior are descendant, temporal precedent is a reasonable *baseline assumption* for working with genetic association data (see Briley et al., 2018 for extensive discussion). That is, we can reasonably assume that with association between SNPs and a measured phenotype, generally the direction of causality is SNPs → phenotype, and not vice versa. This is an assumption that non-genetic psychological data cannot address (Briley et al., 2018). Temporal precedent of genes in relation to phenotypes is particularly relevant for hypothesized causal chains whereby genes → trait 1 → trait 2, or as in the case of pleiotropy, trait 1 ← genes → trait 2. However, not all genetic associations identified from GWAS are causal but, in a broad sense, temporal precedence of genes to phenotypes in a broad sense is an informative starting point. Psychologists interested in causal inference of phenotypic models would benefit from considering behavior genetic models to strengthen claims (Briley et al., 2018; Gage et al., 2016; Johnson & Penke, 2014; Turkheimer & Harden, 2013).

Inferring developmental processes.

Although it is a reasonable assumption that, broadly, genes cause phenotypes, genes are not the only cause of phenotypes. Phenotypes are intimately related to developmental processes; and behavior genetics assumes that phenotypes arise from complex interplay between genes and environment across development. Understanding developmental processes are therefore necessary for understanding

phenotypes of interest to psychologists and social scientists. Developmental processes underlying behavior genetic models are diverse and complex, making identification of any particular developmental process underlying a particular phenotype of interest a difficult task. Briley et al. (2019) provides extensive discussion of developmental processes underlying behavior genetic models, and the reader is encouraged to consult this text directly.

The most common developmental processes discussed in the behavior genetic literature are gene-environment correlations (see Kendler & Baker, 2007; Scarr & McCartney, 1983), which contribute to similarities between genes and environments, and to amplification of genetic effects overtime; gene-environment interactions (see Tucker-Drob & Bates, 2016), in that genetic or environmental effects can have differing effects given either one's genes or environment, and failure to model such interactions can inflate genetic or environmental estimates depending on the particular interaction type; and simultaneous gene-environment interplay, whereby both correlation and interaction processes (discussed above) are occurring.

Adult phenotypes, which are most often studied in behavior genetics, need to consider the range of developmental processes that can yield the adult phenotype (Briley et al., 2019) despite the immense difficulty involved in modeling such processes with real data. Put differently, knowing that a personality trait in a sample of adults is 40% tells scientists nothing about the developmental processes yielding that estimate, or how the estimate changed over time. Careful examination of potential developmental processes for a phenotype in the hope of eliminating certain processes can narrow the remaining possibilities for a phonotype. Despite the difficulty in modeling behavior genetic developmental processes, such understanding is necessary for a complete understanding of phenotypes, and empirical attention to developmental behavior genetics is an important step forward for the field.

Inferring evolution of psychological traits.

Although the fields of behavior genetics and evolutionary psychology both have their historical origins in the 1970s the fields have developed largely independently with little integration sought between the fields. One reason for this is the historical focus of evolutionary psychology to focus on species-typical adaptations (see Tooby & Cosmides, 1992) whereas behavior genetics has historically focused on sources of individual differences (see Plomin et al., 2013). Despite these disciplinary differences, behavior genetics can be informative for evolutionary psychology and offer important insights as to test evolutionary hypotheses (see Arslan & Penke, 2015, Penke et al. 2007, Penke & Jokela, 2016, and Zietsch et al., 2015 for extended discussions).

First, genetic correlations between traits (discussed above) can be informative for understanding the co-evolution of traits, such as those regarding sexual selection hypotheses (e.g., the link between female preferences and male ornamentation), and for supporting by-product hypotheses (e.g., if trait X is a byproduct of trait Y then they should predictably covary genetically). Behavior genetics, and particularly GWAS methods can importantly inform proposed evolutionary processes for traits such as whether the existence of a trait is most likely due to mutation-selection balance, neutral-mutation-drift, or balancing selection, which are each underpinned by differing genetic architecture (Arslan & Penke, 2015; Penke et al., 2007; Zietsch et al., 2015). Schizophrenia is a great example for which behavior genetic methods have yielded insights to the evolution of the disorder. Because the causal variants (i.e., variants found to be associated with schizophrenia in GWAS) underlying schizophrenia explain fractions of a percent of the variance in the disorder suggests that deleterious causal variants with large effects are selected against and are therefore rare (Ripke et al., 2014). Put differently, behavior genetic evidence suggest that schizophrenia is most likely under negative selection and its frequency maintained by mutation-selection balance, therefore rendering adaptationist hypotheses of schizophrenia unlikely to be true. As GWAS increase in size, frequency, and breadth, evolutionary psychology can greatly benefit from its findings.

Conclusion

Behavior genetics has revolutionized our understanding of psychology. The big insights from behavior genetics (see Plomin et al. 2016) have demonstrated that all complex phenotypes that psychologists and social scientists are interested in are to some degree heritable (Polderman et al., 2015). To ignore this foundational fact of psychological science is to fundamentally limit our understanding of who we are. Research methods available today, most notably classical family studies and rapidly advancing genome association and sequencing approaches, continue to deliver insights into psychology and behavior. As the field's methodological tool kit advances, however, understanding of behavior genetic findings by scientists and the public need to also advance. Behavior genetics findings are more nuanced in their interpretations than early family studies suggested. For example, interpretations such as 'the lack of substantial shared environmental effects implies that parents have little effect on their children's outcomes,' and 'substantial heritability of a trait naturally implies support for adaptationist models of traits' are no longer grounded in modern behavior genetics. Our intention with this chapter was to provide a broad introductory overview of the popular methods in the field, what can be inferred from these methods and, importantly, how complex and nuanced such interpretations are for understanding psychology and behavior. Behavior genetics has changed the way we understand psychology, and will continue to alter the landscape of psychological understanding.

References

Arslan, R. C. & Penke, L. (2015). Evolutionary genetics. In D. M. Buss (ed.): *Handbook of Evolutionary Psychology Vol. 2: Integrations* (2nd ed.) (pp. 1047-1066). New York: Wiley.

Barnes, J. C., Boutwell, B. B., Beaver, K. M., Gibson, C. L., & Wright, J. P. (2014). On the consequences of ignoring genetic influences in criminological research. *Journal of Criminal Justice, 42*, 471–482.

Belsky, D. W., & Harden, K. P. (2019). Phenotypic annotation: Using polygenic scores to translate discoveries from genome-wide association studies from the top own. *Current Directions in Psychological Science*, *28*, 82-90.

Bhattacharjee S. & Sarkar A. (2017). Equal environments assumption. In: Vonk J. & Shackelford T.K. (Eds.), *Encyclopedia of Animal Cognition and Behavior*. Springer.

Border, R., Johnson, E. C., Evans, L. M., Smolen, A., Berley, N., Sullivan, P. F., & Keller, M. C. (2019). No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *American Journal of Psychiatry*, *176*, 376-387.

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. *Cell*, *169*, 1177-1186.

Briley, D. A. & Tucker-Drob, E. M. (2017). Comparing the developmental genetics of cognition and personality over the lifespan. *Journal of Personality, 85*, 51–64.

Briley, D. A., Livengood, J., & Derringer, J. (2018). Behaviour genetic frameworks of causal reasoning for personality psychology. *European Journal of Personality, 32*, 202-220.

Briley, D. A., Livengood, J., Derringer, J., Tucker-Drob, E. M., Fraley, R. C., & Roberts, B. W. (2019). Interpreting behavior genetic models: seven developmental processes to understand. *Behavior Genetics*, *49*, 196-210.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., ... & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297, 851-854.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., ... & Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*, 301, 386-389.

Cecil, C. A., Barker, E. D., Jaffee, S. R., & Viding, E. (2012). Association between maladaptive parenting and child self-control over time: Cross-lagged study using a monozygotic twin difference design. *The British Journal of Psychiatry, 201*, 291-297.

Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., & Laibson, D. I. (2015). The fourth law of behavior genetics. *Current Directions in Psychological Science*, *24*, 304-312.

Charney, E. (2012). Behavior genetics and postgenomics. *Behavioral and Brain Sciences*, *35*, 331-358.

Coleman, J. R., Gaspar, H. A., Bryois, J., & Breen, G. (2019). The genetics of the mood disorder spectrum: Genome-wide association analyses of over 185,000 cases and 439,000 controls. *Biological Psychiatry*.

Dick, D. M., Agrawal, A., Keller, M. C., Adkins, A., Aliev, F., Monroe, S., ... & Sher, K. J. (2015). Candidate gene–environment interaction research: Reflections and recommendations. *Perspectives on Psychological Science*, *10*, 37-59.

Duncan, L., Shen, H., Gelaye, B., Meijsen, J., Ressler, K., Feldman, M., ... & Domingue, B. (2019). Analysis of polygenic risk score usage and performance in diverse human populations. *Nature Communications*, *10*, 1-9.

Evans, L. M., Tahmasbi, R., Vrieze, S. I., Abecasis, G. R., Das, S., Gazal, S., ... & Yang, J. (2018). Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nature Genetics*, *50*, 737.

Génin, E. (2019). Missing heritability of complex diseases: case solved? *Human Genetics*, 1-11.

Hill, W. D., Arslan, R. C., Xia, C., Luciano, M., Amador, C., Navarro, P., Hayward, C., Nagy, R., Porteous, D., McIntosh, A. M., Deary, I. J., Haley, C., & Penke, L. (2018). Genomic analysis of family data reveals additional genetic effects on intelligence and personality. *Molecular Psychiatry, 23, 2347–2362.*

Jaffee, S. R., Caspi, A., Moffitt, T. E., & Taylor, A. (2004a). Physical maltreatment victim to antisocial child: evidence of an environmentally mediated process. *Journal of Abnormal Psychology*, *113*, 44-55.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*, 860-921.

Jaffee, S. R., Caspi, A., Moffitt, T. E., Polo-Tomas, M., Price, T. S., & Taylor, A. (2004b). The limits of child effects: evidence for genetically mediated child effects on corporal punishment but not on physical maltreatment. *Developmental Psychology*, *40*, 1047-1058.

Johnson, W., & Penke, L. (2014). Genetics of social behavior. In B. Gawronski & G. Bodenhausen (eds.): *Theory and Explanation in Social Psychology* (pp. 205-223). New York: Guilford Press.

Joseph, J. (2004). *The gene illusion: Genetic research in psychiatry and psychology under the microscope*. New York: Algora.

Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: A systematic review. *Psychological Medicine, 37,* 615–26.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A test of the equal-environment assumption in twin studies of psychiatric illness. *Behavior Genetics, 23*, 21–27.

Liu, C., Molenaar, P. C., & Neiderhiser, J. M. (2018). The impact of variation in twin relatedness on estimates of heritability and environmental influences. *Behavior Genetics*, *48*, 44-54.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., ... & Cho, J. H. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*, 747-753.

Marouli, E., Graff, M., Medina-Gomez, C., Lo, K. S., Wood, A. R., Kjaer, T. R., ... & Rüeger, S. (2017). Rare and low-frequency coding variants alter human adult height. *Nature*, *542*, 186.

McAdams, T. A., Hannigan, L. J., Eilertsen, E. M., Gjerde, L. C., Ystrom, E., & Rijsdijk, F. V. (2018). Revisiting the children-of-twins design: improving existing models for the exploration of intergenerational associations. *Behavior Genetics*, *48*, 397-412.

Mills, M. C., & Rahal, C. (2019). A scientometric review of genome-wide association studies. *Communications Biology*, *2*, 1-11.

Pearl, J. & Mackenzie, D. (2018). The book of why: The new science of cause and effect. Basic Books: New York, NY.

Penke, L. (2011). Bridging the gap between modern evolutionary psychology and the study of individual differences. In D. M. Buss & P. H. Hawley (Eds.), *The evolution of personality and individual differences* (pp. 243–79). New York, NY: Oxford University Press.

Penke, L., Denissen, J. J. A., & Miller, G. F. (2007). The evolutionary genetics of personality (target article). *European Journal of Personality, 21*, 549-587.

Penke, L., & Jokela, M. (2016). The evolutionary genetics of personality revisited. *Current Opinion in Psychology*, *7*, 104-109.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *88*, 245–58.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science*, *11*, 3-23.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2013). *Behavioral genetics* (6th ed.). New York: Worth Publishers.

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*, 702–709.

Purcell, S. (2016). Statistical methods in behavioral genetics. In R. Plomin, J.C. DeFries, V.S. Knopik, & J.M. Neiderhiser, *Behavioral Genetics, Sixth Edition* (pp. 357-412).

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K. H., Holmans, P. A., ... & Pers, T. H. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*, 421.

Scarr, S. (1968). Environmental bias in twin studies. *Eugenics Quarterly*, *15*, 34-40.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development, 63*, 1–19.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*, 424–35.

Segal, N. L., Hernandez, B. A., Graham, J. L., & Ettinger, U. (2018). Pairs of Genetically Unrelated Look-Alikes. *Human Nature, 29*, 402-417.

Stanley, T. D., Carter, E. C., & Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin, 144*, 1325-1346.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19-136).

Tucker-Drob, E. M., & Bates, T. C. (2016). Large cross-national differences in gene× socioeconomic status interaction on intelligence. *Psychological science, 27*, 138-149.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science, 9*, 160-164.

Turkheimer, E. (2015). Genetic prediction. *Hastings Center Report, 45*, S32-S38.

Turkheimer, E., & Harden, K.P. (2014). Behavior genetic research methods: Testing quasi-causal hypotheses using multivariate twin data. In H.T. Reis & C.M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology*. 2nd Edition. Cambridge University Press.

van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H., & Wray, N. R. (2019). Genetic correlations of polygenic disease traits: from theory to practice. *Nature Reviews Genetics, 20*, 567-581.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., ... & Gocayne, J. D. (2001). The sequence of the human genome. *Science, 291*, 1304-1351.

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: biology, function, and translation. *The American Journal of Human Genetics, 101*, 5-22.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., ... & Goddard, M. E. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics, 42*, 565-569.

Yengo, L., Robinson, M. R., Keller, M. C., Kemper, K. E., Yang, Y., Trzaskowski, M., ... & Wray, N. R. (2018). Imprint of assortative mating on the human genome. *Nature Human Behaviour, 2*, 948-954.

Zietsch, B. P., de Candia, T. R., & Keller, M. C. (2015). Evolutionary behavioral genetics. *Current Opinion in Behavioral Sciences, 2*, 73-80.