HERITABILITY STUDIES: METHODOLOGICAL FLAWS, INVALIDATED DOGMAS, AND CHANGING PARADIGMS*

CALLIE H. BURT
University of Washington

Citation: Burt, Callie H. Heritability studies: Methodological flaws, Invalidated dogmas, and changing paradigms. Advances in Medical Sociology: Health, Genetics, & Society 16, *forthcoming* 2015.

Running Head: HERITABILITY STUDIES IN THE POSTGENOMIC ERA

* Direct all correspondence to Callie Burt, Department of Sociology, University of Washington, Box 353340, Seattle, WA 98195; email: chburt@asu.edu.

HERITABILITY STUDIES: METHODOLOGICAL FLAWS, INVALIDATED DOGMAS, AND CHANGING PARADIGMS

ABSTRACT

Purpose: Heritability studies attempt to estimate the contribution of genes (vs. environments) to variation in phenotypes (or outcomes of interest) in a given population at a given time. The current chapter scrutinizes heritability studies of adverse health phenotypes, emphasizing flaws that have become more glaring in light of recent advances in the life sciences and manifest most visibly in epigenetics.

Design/methodology/approach: Drawing on a diverse body of research and critical scholarship, this chapter examines the veracity of methodological and conceptual assumptions of heritability studies.

Findings: The chapter argues that heritability studies are futile for two reasons: (1) heritability studies suffer from serious methodological flaws with the overall effect of making estimates inaccurate and likely biased toward inflated heritability, and, *more importantly*, (2) the conceptual (biological) model on which heritability studies depend—that of identifiably separate effects of genes vs. the environment on phenotype variance—is unsound. As discussed, contemporary bioscientific work indicates that genes and environments are enmeshed in a complex (bidirectional, interactional), dynamic relationship that defies any attempt to demarcate separate contributions to phenotype variance. Thus, heritability studies attempt the biologically impossible. The emerging research on the importance of microbiota is also discussed, including how the commensal relationship between microbial and human cells further stymies heritability studies.

Originality/value: Understandably, few sociologists have the time or interest to be informed about the methodological and theoretical underpinnings of heritability studies or to keep pace with the incredible advances in genetics and epigenetics over the past several years. The present study aims to provide interested scholars with information about heritability and heritability estimates of adverse health outcomes in light of recent advances in the biosciences.

Keywords: heritability study, twin study, epigenetics, plasticity, postgenomics, microbiome

HERITABILITY STUDIES: METHODOLOGICAL FLAWS, INVALIDATED DOGMAS, AND CHANGING PARADIGMS

"In the real world of humans, in a given context everything is heritable to some extent and environmental to some other extent, but the magnitudes of the proportions are variable from situation to situation, and have nothing whatsoever to do with the *causal* properties of genes and environment for the trait in question, unless one is interested in the pointless null hypothesis that one of the components is zero. This message has taken 100 years to soak into genetic social science, and it has not fully soaked in yet: not a month goes by without another outbreak of credulous surprise that one trait or another has turned out to be 50% heritable." (Turkheimer 2011: 598).

The nature vs. nurture debate continues in the social and behavioral sciences in the form of the heritability study. These studies aim to partition the variation in an outcome of interest (phenotype) into a proportion caused by genes (heritability) and a proportion caused by the environment. The general conclusion emerging from these studies is that variance in every examined characteristic—including preferences, social behaviors, and diseases—is significantly shaped (~50%) by genetic influences (Turkheimer 2000). For example, recent studies show substantial heritability of everything from breastfeeding (Colodro-Condre et al. 2013), breakfast eating patterns (Keski-Rahkonen et al. 2004), "treatment responses to tooth whitening" (Corby et al 2014), and sickness absences (Svedberg et al. 2012) to ADHD, autism, depression, hypertension, high cholesterol, and diabetes (e.g., Plomin, DeFries, Knopik, & Neiderhiser 2012).

While heritability studies continue to abound, the past decade has witnessed significant advances in our understanding of the function of "genes,"[1] including the nature of genetic influences on adverse health phenotypes. These advances in knowledge have not answered questions about how much variation is genetic versus environmental (or nature vs. nurture), but rendered such questions manifestly obsolete (Keller 2010). Recent advances in molecular genomics reveal that the causal interactions among genes, proteins, cells, and physical and social environments on health and development are entangled, dynamic, and context dependent. Such advances have exposed the conceptual framework of heritability studies—that of identifiably separate effects of genes vs. environments—as unsound (Burt & Simons 2014; Charney 2012; Charney & English 2013) and have fostered a

[1]Indeed, "several recent discoveries have cast serious doubt on the idea that there are coherent entities [clearly identifiable particulate units of inheritance] in our DNA that can unambiguously be called 'genes'" (Moore 2014: 46). Furthermore, although the gene was once considered the sole unit of inheritance, there is growing evidence to suggest that there are other mechanisms of biological inheritance, such as the epigenetic inheritance system (e.g., Charney 2012; Jablonka & Lamb 2006). Despite this, to maintain consistency with discussed studies and avoid confusion, I will continue to use the terms "genetic influences" and "genes" to refer to inherited DNA sequences.

growing scholarly awareness, including among former proponents of heritability studies, that heritability studies are superannuated (e.g., Rutter 2006; Turkheimer 2011). Yet, this misguided enterprise continues, primarily in the social and behavioral sciences, despite the fact that the biological model undergirding heritability studies is incompatible with contemporary bioscientific knowledge (Charney & English 2012; Duster 2003).

In the present chapter, I discuss heritability studies with two goals. First, I aim to assist readers in understanding heritability and making more judicious assessments of heritability estimates. Given the prevalence of heritability studies, the failure of some studies to clearly describe the shaky foundations of the model, and the potential misuse of these estimates, as well as the fact that these studies will continue (at least in the short term), it is vital that scholars are cognizant of the meaning and limitations of heritability estimates (e.g., "The heritability of ADHD in children is an estimated 75%"; Rietveld et al. 2004). Thus, I first discuss the concept of heritability and the technical/statistical limitations of the model, focusing on the most common method at present, the "twin study." Second, I critique the conceptual (biological) model of heritability studies, emphasizing fatal flaws that have become more glaring in light of recent advances in the life sciences, focusing on their relevance for studies of health. Echoing the arguments of prominent scholars (e.g., Charney 2012; Crusio 2012; Joseph 2004; 2006), I argue for an end to heritability studies as well as recognition of the problematic nature of existing heritability estimates.

In this chapter I argue against the continuation of heritability studies in the social sciences. I maintain that heritability studies are futile for two reasons. First, heritability studies suffer from serious methodological flaws with the overall effect of making estimates (potentially highly) inaccurate and biased toward inflated heritability and deflated shared environmental influences. Second, and more importantly, the conceptual (foundational) model on which heritability models depend—that of identifiably separate effects of genes vs. the environment on phenotype variance—is misguided. Although the conceptual model of heritability studies, like the technical model, has been criticized for more than half a century, profound scientific advances over the past decade, perhaps best exemplified in epigenetic research, provide clear evidence that the dichotomous view of "genetic" vs. "environmental" effects on phenotype variance assumed by heritability studies is untenable (Burt & Simons 2014;

2015; Charney 2012)."[D]evelopment is not merely a process involving a battle between nature (genes) and nurture (experience) but the interweaving of dynamic processes within a system that is inseparably both the organism and its environment" (Kaplan & Rodgers 2003: 5). Thus, regardless of their technical merits, the very goal of heritability studies is biologically impossible.

After laying out my argument that heritability studies should be discontinued, I conclude with a brief discussion of the need to move beyond heritability to focus on the plasticity of human development and the flexibility of genetic expression in response to environmental influences. As I hope is clear, I attempt to push the field forward, not by denying the role of genes or biological factors, but by recognizing the complexity of the biopsychosocial system. Many nuances and details cannot be covered in this chapter; fortunately, there is an extensive body of work that I point the reader to for more detail.[2] Before moving into the limitations and flaws in heritability studies, I provide a brief introduction to heritability, a concept that is frequently misunderstood.

HERITABILITY CONCEPT

The heritability construct was originally developed in agricultural research to assist animal and plant breeders to predict the outcomes of controlled breeding programs. Since the 1960s, this concept has been promoted as the "nature-nurture ratio," or the relative influence of heredity (through genes) and environmental experiences on trait variance within human populations (Joseph 2006; Lewontin, Rose, & Kamin 1984). Heritability is defined narrowly by Wahlsten (1994: 224) as "the proportion of variance in a measure of behavior or other phenotype in a breeding population that is attributable to genetic variation," and more loosely by Plomin and colleagues (2012:87) as "the proportion of phenotypic variance that is accounted for by genetic differences among individuals." The heritability coefficient is the numerical index of heritability that ranges from 0 (no genetic contribution) to 1.0. This estimate is derived from a model based on Mendelian principles as well as a number of assumptions, discussed below. Given that the construct of heritability is frequently misunderstood, it is useful to clarify some misunderstandings about heritability estimates.

[2] Notably, most of the arguments in this chapter are not original but are those of prominent scientists, many of whom I cite, whose criticisms accumulating over the past 80 years come together, in my view, to demonstrate the futility of the heritability study.

Perhaps most importantly, heritability estimates do not indicate the effect of genes on a particular trait. Instead, they index the genetic contribution to trait *variability* in a population. Heritability estimates are often wrongly interpreted as indicating the former by laypersons as well as some scholars (Hirsch 2004; Joseph 2006).[3] The methods of Mendelian genetics are responsive only to the tiny portion of genes that are polymorphic and make us different; they do not allow for conclusions about the role of heredity in general (Wahlsten 1990). "It is an illegitimate inference to assume that an assumption that variance for a trait is due largely to 'genes' as opposed to 'environment' entails that the trait itself is inherited (i.e., genetically transmitted from parents to offspring). Causes of a trait, and causes of trait variance are not the same thing" (Charney 2012:61). As Joseph (2004: 138-9) notes "This leads to a paradoxical situation in which a trait could be 100% inherited, yet have a heritability of zero—human beings having two eyes for example." Genetics, of course, explains humans' "eyedness"; we have two eyes because of our genetic endowment. Because everyone (or nearly everyone) who has fewer than two eyes is that way as a result of life experiences, the heritability of "eyedness" is *zero* (see also Lewontin et al. 1984). As such, heritability estimates do not indicate, for example, what proportion of a population's or an individual's risk for a disease is due to genes. Moreover, heritability estimates signify nothing about the number of genes involved nor do they contribute to the identification of specific candidate genes (e.g., Schielzeth & Husby 2014).

Notably, heritability is not a fixed estimation of the overall effects of genetics on trait variance but rather is time and population specific. As such, findings of genetic influences on differences within populations do not extrapolate into explaining differences observed in traits across populations (e.g., adults vs. children; rural vs. urban populations; low SES vs. high SES populations; Lewontin 1986). Indeed, even for trait variance that is entirely heritable within a population, the cross-population variance may be due completely to environmental factors (see Lewontin 1986; Joseph 2004). This aspect of heritability is wonderfully illustrated by Lewontin and colleagues (1984). Suppose one takes two handfuls of heterogeneous corn seed and plants one handful on a field of

[3] Much of the misunderstanding over the meaning of heritability has to do with the confusion of the words "heritable" and "inherited" or "heredity" (Joseph 2004; Stoltenberg 1997). The two concepts, according to Hirsch (1997:220) "have been hopelessly conflated…[b]ecause of their assonance." Indeed, this assonance and confusion has led Stoltenberg (1997) to propose that the word "heritability" be replaced by the word "selectability" so as not to confuse the scientific and folk definitions of the term.

nutrient-rich soil, and the other in a field of nutrient-poor soil. When the seeds have grown, we can see that there is variation within fields in plant height, as well as variation between fields, specifically with lower plant height for the poor quality soil field. Not allowing any environmental variation within the respective fields (equal light, water, etc.), differences in the resulting plant height *within* the fields is totally due to genetic factors (heritability = 1.0). Differences in plant height *between* the two fields, however, are completely due to environmental factors, specifically soil quality (heritability = 0).

Importantly, heritability estimates do not speak to the responsiveness of a phenotype to environmental interventions. Traits can be highly heritable and yet be drastically altered or eliminated by changes in the environment (e.g., Lewontin 1974; Plomin et al. 2012; see Joseph 2004, p. 140, for a discussion of the classic example of PKU). Thus, high heritability does not mean inevitability of a phenotypic outcome. In addition, the heritability of a trait can change as a result of changes in the environment. Returning to the cornfield example, if one were to plant trees around the corn field, thereby producing unequal sun exposure among the corn plants within a field, heritability would decrease as the resulting variation in sunlight exposure would account in part for differences in plant height. Thus, there *is not* and *cannot* be any static absolute value for the strength of genetic influences on trait variance.

HERITABILITY STUDIES

Grounded in the classical gene-centric paradigm (e.g., the central dogma of the genomic era), especially the assumption that genes and environments make identifiably separate contributions to phenotypes, heritability studies attempt to estimate the relative influence of genes on trait variance in a population. Under the prevailing methodology, the heritability of a given phenotype is usually estimated by comparing concordances and discordances between subjects relative to their presumed degree of genetic similarity. Various models have been used in traditional studies of heritability, including family, adoption, twins-reared-together, and twins-reared-apart studies. Although these different models are grounded in different assumptions, they are united by the fact that all compare phenotypes across varying degrees of genetic relationships and use these comparisons to estimate genetic and environmental influences without actually measuring either. (A newer method, GCTA, which does not rely on family comparisons and measures genetic variants, will also be briefly discussed later.) At present the

study of reared-together twins (referred to as the "twin study") is the primary workhorse of behavioral genetic studies of heritability; thus, I focus on twin studies.[4] Below I briefly discuss the set of technical/statistical assumptions and problems inherent in twin studies (for more detail on these technical flaws see, e.g., Jackson 1960; Joseph 2004; 2006; Lewontin et al. 1984; Wahlsten 1990).

THE TWIN STUDY (TWINS REARED TOGETHER)

INTRODUCTION TO AND FUNDAMENTALS OF TWIN STUDIES

Over the past several decades the "twin study" has become the primary method for estimating heritability.[5] As we know, identical (monozygotic or MZ) twins come from one fertilized egg and are assumed to share 100% of their genes (genetic clones), whereas fraternal (dizygotic or DZ) twins come from two fertilized eggs and are presumed to share on average half of their genes (the same amount as other non-twin biological siblings). Given this and along with several assumptions, researchers have used MZ and DZ twins as a natural experiment to estimate the relative influence of nature and nurture (Joseph 2004; Plomin et al. 2012).

Twin studies separate phenotypic variation into three components: genetics (*h*), shared environment (*c*), and unshared environment (*e*). Notably, the shared and unshared environments are misnomers (Burt & Simons 2014). The shared environment consists of non-genetic influences that make twins similar to each other, such as parenting, SES, and community factors. Although it is frequently assumed that the influences that cause trait similarity are in fact shared, technically, this component captures any factors (shared or not) that cause twins to be more similar to one another (Plomin 2011; Suhay & Kalmoe 2010). The unshared environment, on the other hand, consists of all non-genetic factors that create differences in traits among twins, such as different peers, classrooms, interactional experiences and the like, as well as model error. As with the shared environment, while it is often assumed that environmental influences that cause differences are "unshared" in the technical sense (e.g.,

[4] Space constraints do not permit an extensive examination of the statistical features and assumptions of the two other classic heritability models, adoption and twins-reared apart studies, which also rest on dubious empirical and conceptual foundations. Fortunately, excellent critical discussions of these models in relation to adverse health outcomes have appeared elsewhere (see, e.g., Joseph 2004; 2006; see Burt & Simons 2014 for a more concise overview).

[5] Some have misguidedly overgeneralized arguments against twin studies of heritability as arguments against the use of twin samples in research generally (e.g., Moffitt & Beckley 2015). To be clear, my critique is focused on twin studies of *heritability*, and it is silent on the use of twin-based research for other purposes (e.g., studies of MZ twin *dis*cordance).

being treated differently by parents), all factors that produce differences among twins are incorporated into the unshared environment, whether actually shared or not. Scholars not infrequently fail to describe what is meant by these terms, and some have made inappropriate and widely publicized conclusions about the irrelevance of parenting or community factors based on shared environmental estimates (e.g., Harris 1998; Rowe 1994).

The basic logic of the twin study model is to compare twin concordances for phenotypes and, based on several assumptions mentioned below, assign the greater similarity of MZ co-twins relative to DZ co-twins to their greater genetic similarity. The key assumption in this natural experiment is that everything is the same between the two types of twins *except* for the fact that MZ co-twins are genetic clones while DZ co-twins share on average 50% of their genes. In the simplest approach, heritability is usually estimated from twice the MZ-DZ differences in correlations. More recently, complex latent variable models have been utilized; however, the basic logic underlying these more sophisticated models is the same as that for the more naïve models.

The assumptions of the twin study have remained largely unchanged since the 1920s and remain crucial to the model given that neither genetic nor environmental influences are measured.[6] Some of these assumptions are relatively unproblematic, but others are quite dubious. Prominent questionable assumptions include the following: (Charney 2012; Joseph 2006; Plomin et al. 2012).[7]

1. The environments of MZ co-twins are no more similar than that of DZ co-twins for trait-relevant environments (trait-relevant equal environments assumption).

2. The relevant genes exert effects additively.

3. The risk of receiving a diagnosis for a phenotype (e.g., ADHD, depression, diabetes) is the same between MZ and DZ co-twins.

4. Phenotypic variation can be demarcated into genetic (*h*), shared environmental (*c*), or unshared environmental (*e*) components.

[6] The exception is the modification of the equal environments assumption to its trait-relevant form given clear evidence that the former was invalid (see Joseph 2004; 2006, for an overview of this shift.)

[7] As noted this is an incomplete list of assumptions, and excludes one, no assortative mating, whose violation has the effect of biasing heritability downward (e.g., Rutter 2006). Given space constraints combined with the fact that the no assortative mating assumption is more often noted in heritability studies (often implying or stating that the *h*2 estimate might be an underestimate), I focus on these four assumptions.

In general, few studies mention these assumptions, and even fewer describe them or the likely impact of their violation, which is generally inflated heritability estimates and deflated shared environmental influences. Below, I discuss (and question the validity of) these key assumptions from the classical genetic perspective that undergirds the heritability model, leaving aside the conceptual deficiencies that are glaring in light of recent advances in molecular genomics. I then turn to a critique of the model from a postgenomic perspective.

SOME METHODOLOGICAL LIMITATIONS OF TWIN STUDIES

Equal Environments Assumption (#1)

The trait-relevant equal environments assumption (hereafter EEA) is the assumption that the covariance between genetics and the environment is zero for environments that influence variance in the outcome under study. In other words, the model assumes that MZ co-twins experience the same degree of (trait-relevant) environment similarity as DZ co-twins (assuming, for example, that MZ co-twins are treated no more similarly than DZ co-twins). Without the EEA, the greater similarity of MZ co-twins could be due to genetics or more similar environments (see Suhay & Kalmoe 2010, for an algebraic demonstration of the importance of the EEA). The consequence of the EEA is that all greater phenotype similarity between MZ than DZ twin pairs is credited to genetics. As a result, EEA violations (greater trait-relevant environmental similarity among MZ than DZ twin pairs) serve to artificially inflate heritability and deflate shared environmental estimates.

Although the extent of EEA violations varies across different traits, in general, research evinces that MZ co-twins experience much more similar environments that DZ co-twins (see Joseph 2006 for an excellent review). Indeed, for years many scholars have argued that the more similar environments for MZ than DZ co-twins bias heritability upwards to an unknown degree for many, if not most, phenotypes (e.g., Horwitz et al. 2003; Joseph 2004; Lewontin et al. 1984). For example, research shows that MZ co-twins are more likely to be treated similarly by parents and others (Evans and Martin 2000), to have the same friends (Cronk et al. 2002; Horwitz et al. 2003), to share the same classroom (Cronk et al. 2002; Morris-Yates et al. 1990), and to spend time together (and therefore experience the same environments more frequently; Horwitz et al. 2003; Rende et al. 2004). Thus, MZ co-twins are more likely to experience the same physical environments (including exposure

to pollution and chemicals, food/nutrition, even traumatic experiences) and engage in the same activities (e.g., exercise; Carlsson et al. 2006), many of which can have profound and diverse effects on adverse health outcomes.

Evidence that EEA violations are consequential (i.e., inflate heritability) is found in studies that examine the association of greater environmental similarity with trait concordance (see e.g., Beckwith & Morris 2008; Joseph 2004). For example, Cronk and colleagues (2002) found that DZ co-twins who more often share classes, share friends, and dress similarly were more similar on health outcomes, such as ADHD and anxiety disorder. The authors also reported a significant influence of "perceived zygosity"[8] on conduct disorder and anxiety. Similarly, another study found that measures of environmental similarity predicted greater phenotype similarity in symptom scores for almost half of the tested health outcomes (Morris-Yates, Andrews, & Henderson 1990). Thus, evidence suggests that the more similar environments of MZ co-twins are trait-relevant (influence concordance) for adverse health outcomes and have the effect of upwardly biasing heritability estimates.

The EEA is even less reasonable for studies that include opposite-sex DZ twin pairs in their models. In these cases, the model assumes that MZ co-twins (all of whom are same sex) are treated no more similarly than opposite-sex DZ twins. Given the voluminous research on sex/gender differences in experiences, this assumption seems questionable. Evidence for this interpretation is found in Meier et al. (2010) who compared the correlation for childhood conduct disorder among opposite-sex and same-sex DZ twins and found that the opposite-sex correlation was significantly smaller (approximately half) than that of the same-sex DZ twins (see also Saudino, Ronald, and Plomin, 2005).

An even more questionable practice is the use of sibling data, which includes MZ twins, DZ twins, full siblings, half siblings, and sometimes cousins. In these models, kinship pairs are compared with their average level of genetic relationship, ranging from 1 for MZ twins to .25 for half siblings and .125 for cousins. These models assume that the environments of opposite-sex, different age cousins who reside in different homes are no less similar than that for MZ co-twins. The EEA seems absurd for studies of kinship pairs.

[8] See Beckwith and Morris (2008) for a discussion of "perceived zygosity" tests of the validity of the EEA.

Moreover, while the above-mentioned evidence focuses on the greater postnatal environmental similarity between MZ than DZ co-twins (and in some cases siblings), the greater environmental similarity for most MZ co-twins actually begins in the womb (see Charney 2008; Joseph 2006). Although this greater prenatal similarity may not be relevant for some traits, many adverse health phenotypes could conceivably be influenced by intrauterine exposure to chemical toxins, viruses, or the like.[9]

In sum, for many health outcomes, the greater prenatal and postnatal environmental similarity for MZ versus DZ co-twins (and especially when pairs of non-twin siblings or cousins are utilized) likely violates the trait-relevant EEA (e.g., Charney 2012; Suhay & Kalmoe 2010). To be sure, for most health-related phenotypes, there is much more work to be done in identifying pre- and postnatal environmental factors that contribute to the development of adverse health outcomes. As such, the twin study model rests on the heroic assumption of the equality of trait-relevant environments (many unknown) for MZ and DZ co-twins. To be sure, some studies recognize and make a laudable effort to control for the more similar environments for MZ co-twins (e.g., Boardman et al. 2011); however, capturing the manifold of ways in which MZ co-twins environments are more similar with survey items (especially for influences that are as yet unknown) is impracticable.[10] Importantly, even minor violations of the EEA can substantially inflate heritability estimates (thereby underestimating the influence of the shared environment; e.g., Suhay & Kalmoe 2010). Thus, it is a credible interpretation of twin-study findings that heritability estimates are upwardly biased due in no small part to MZ co-twins having more similar shared environments than DZ co-twins (e.g., Horwitz et al. 2003; Jackson 1960; Joseph 2006). Surprisingly, many recent twin studies of adverse health outcomes fail to even mention the EEA, and even fewer discuss the implications of EEA violations.

[9] While out of the scope of the present paper, scholars have noted that MZ and DZ co-twins experience prenatal environments that are more stressful than that of singletons, thus calling into question the generalizability of twin studies (Charney 2012a: 19).

[10] Some twin-study proponents have attempted to test the validity of EEA (at least postnatally) and/or the implications of EEA violations on heritability estimates using survey items about environmental similarity as well as cases of misidentified zygosity (e.g., Boardman et al. 2011). Certainly, contradictory evidence has emerged from these studies (discussed in the text); however, some studies report findings that are interpreted as bolstering the validity of the EEA. As discussed elsewhere, these tests are both limited and problematic (e.g., Beckwith & Morris 2008; Charney 2008; Joseph 2004; Suhay & Kalmoe 2010). For example, these tests frequently employ only a few broad (not necessarily trait-relevant) retrospective survey items about environmental similarity in attempt to capture similarity in environments. As discussed, there are numerous (and some unknown) trait-relevant environments, and these evade measurement in broad surveys especially with a limited number of items.

Genetic Additivity Assumption (#2)

Genetic influences can be separated into additive and nonadditive components. Nearly all twin studies assume an additive model of gene combinative effects such that genetic variants each contribute small individual effects that add up to shape variation in a phenotype. Nonadditive genetic variance is that which occurs as a consequence of interactions between genes such that the phenotype is different from the sum of the individual genetic effects (Stoolmiller 1999). Nonadditive variance can be one of two types. Dominance is that which arises when the alleles at a given locus (one from each parent) interact to produce a phenotype and occurs among genes that operate with a strict dominance-recessive mode of inheritance. Epistasis occurs when several genes (alleles at different loci) interact to produce a behavior (Stoolmiller 1999). Although the extent of nonadditive genetic variance for various health phenotypes is not known, there is good reason to believe that it is operative for all complex traits (Plomin et al. 2012).

As with the EEA, violations of the additivity assumption inflate heritability and deflate shared environmental estimates (see Grayson 1989; Stoolmiller 1999 for more detail). Basically, because MZ co-twins share all of their genes, genetic nonadditivity will reduce the genetic correlation for DZ co-twins but not MZ co-twins (because MZ co-twins have identical interacting alleles). Although dominance interactions can be modeled in twin studies, this is not often done and comes with problems of its own (e.g., shared environmental effects are ignored because they cannot be modeled simultaneously; Neale & Cardon 1992). Moreover, modeling epistatic interactions with human kinship data is generally considered impracticable (Eaves 1988). Regarding nonadditivity, renowned behavioral geneticists Plomin and colleagues (2012: 401) stated: "These types of effects complicate model fitting because there are many forms in which they could occur. Normal twin study designs do not offer much hope for identifying them." In short, dominance and epistasis almost certainly upwardly bias heritability estimates to an unknown degree (e.g., Burt & Simons 2014). Moreover, as with the EEA assumption, many twin studies of adverse health outcomes fail to explain the implications of assuming genetic additivity and the consequences of violation.

Nonblind Diagnosis (#3)

As noted, twin studies assume that MZ co-twins are not more likely than DZ co-twins to receive a diagnosis or label in response to their twin's diagnosis (or recognize their symptoms in the case of symptom reports). While space constraints and a lack of empirical research on this issue precludes a detailed discussion, I believe it is worth noting the possibility that nonblind diagnoses of co-twins with adverse health outcomes might bias heritability upwards.

It seems reasonable to expect that, in general, MZ co-twins are more likely to receive adverse health diagnoses in response to a diagnosis of their co-twin than are DZ co-twins, all else equal, given the belief in the heritability of many disease/disorder states, the recognition that MZ co-twins probably share more similar environments, and the tendency to treat them more similarly. Moreover, given their greater attachment, closeness, and feelings of similarity—as well as their recognition that they are genetic clones—MZ co-twins may be more likely than DZ co-twins to identify shared symptoms as problematic or seek medical attention for such symptoms in response to their co-twins diagnoses or symptom identification. Inasmuch as these practices are operative, heritability would be inflated and shared environmental effects would be deflated.

Separate, Identifiable Contributions of Genes and Environments (#4)

As noted, the goal of heritability studies is to parse the effects of nature and nurture on phenotypic variance. Of course, then, the model assumes that genes (G) and environments (E) have identifiably separate effects on phenotypes. However, evidence suggests this is a fallacious assumption.

From a classical genetic perspective, there are two issues that confound the partitioning of G and E effects: genetic-environmental covariance and gene environment interactions (G x E). *Genetic-environmental covariance* is the association of certain genotypes with particular environments. An example in the health realm could be individuals' genetically-influenced tastes (Breen, Plomin, & Wardle 2006) shaping preferences for highly processed, high sugar foods, which, in turn, could increase likelihood of diabetes, in combination with a host of environmental influences (SES, physical activity level, parental monitoring/regulation of food consumption). It is an ongoing matter of debate as to how to classify this covariance in the calculation of heritability. Perhaps most of us would agree that this process fits neatly in neither the G nor the E category. Several behavioral geneticists have

argued that such gene-environment correlations should be classified as genetic effects (e.g., Fowler, Baker, & Dawes 2008; Segal & Johnson 2009); however, as Rutter (2002: 4) noted, "it is misleading to suppose that just because genetic factors influence the occurrence of an environmental risk factor, this must mean that the risk process is genetically mediated. This assumption does not follow because there is no necessary connection between the causes of the origin of a risk factor [taste preferences] and its mode of risk mediation [metabolic effects]." I illustrate this point with the use of skin pigmentation.

In the U.S.A., a person genetically coded to have darker skin pigmentation will experience a different social environment, on average, than one with lighter skin. There is a wealth of evidence that health is influenced by environmental factors that are pervasively and systematically patterned along racial lines in the U.S. (Krieger 2000). Darker skin pigmentation is associated with exposure to environmental risk factors for adverse health outcomes, such as exposure to noxious chemicals, due in part to its association with socioeconomic status, community disadvantages, and environmental racism. Darker skin pigmentation is also associated with interpersonal racial discrimination, which has been linked to a variety of health outcomes, including depression, anxiety, hypertension, and the like (e.g., Burt, Simons, & Gibbons 2012; Brody et al. 2014; Williams 1999). Surely, we can all agree that classifying adverse health outcomes that result from the interaction of skin pigmentation with societal racism as due to genetic endowment is preposterous. However, even less manifestly spurious cases are problematic, as the twin-study model has to classify such complex, interactional relationships as *either* genetic or environmental. Such dynamic biopsychosocial associations defy such neat classification (Spencer & Harpalani 2004; Wahlsten 1990).

G x E interactions, situations in which the effect of genotype on phenotype is conditioned by environmental input (or environmental influences on phenotypes are conditioned by genotype), also thwart attempts to partition genetic from environmental influences. The traditional twin study requires that one assumes G x E interactions are nonexistent or that their effects are trivial (Wahlsten 1990). There is mounting evidence that this is not the case, and that G x E interactions are the rule rather than the exception (e.g., Bagot & Meaney 2010; Rutter 2007). For example, a wealth of research indicates that much genetic variation among individuals influences their

sensitivity to environmental influences, rather than genotype having an unconditional effect on phenotypes (e.g., Caspi et al. 2003; Belsky & Pluess 2009, see Boardman & Fletcher 2014). Although evidence of G x E interactions are not new (e.g., Hogben 1933), recent evidence documenting their prevalence is mounting with the advent of new technologies that make genotyping easier and more cost effective, especially for complex socially-mediated phenotypes (e.g., Belsky & Pluess 2013).

In the traditional way of partitioning, the effects of gene-environment correlations and interactions will be included in the heritability estimate if the environmental effect is shared (such as skin pigmentation) and in the nonshared environmental estimate if the gene-environment interplay operates in a twin-specific fashion. The rationale is that because origins of the environmental risk factor (e.g., taste preferences or racial discrimination) derive from genes (involved in taste buds or skin pigmentation), it is reasonable to attribute the whole of the environmental effect to genetics (Rutter 2002). Clearly, this argument is unsound. Such ubiquitous interactional relationships almost certainly render heritability estimates wildly inaccurate and deflate estimates of the shared environment (Burt & Simons 2014).

Additional Technical Issues

Confidence Intervals and Model Fitting. An additional reason to view twin-study heritability estimates with caution is that they tend to have large confidence intervals (Burt & Simons 2014; Rutter 2006). Frequently the point estimate is the focus of attention, a practice that serves to reify an inherently imprecise estimate. Perhaps more troublesome is the potential for bias as a result of these larger confidence intervals in the more sophisticated structural models that have been more widely adopted. These models operate on the principle of parsimony, and the usual approach involves a systematic comparison of different models with the aim of finding the simplest model that fits the data. When the elimination of a parameter does not worsen the fit of the model, it is standard practice to drop the parameter.[11] Importantly, the decision to drop a parameter is based on the lower end of the 95% confidence limits without reference to the upper end, and given the large confidence intervals could result in a parameter being dropped that in fact has a significant effect (Rutter 2006). Although many studies do not report

[11] This is customary practice in much structural equation modeling; however, it is important to remember that the lack of a significant effect of the shared environment is not tantamount to evidence of a zero effect (Burt & Simons 2014).

the estimates from the 'inferior' models, when these estimates are presented, it is not uncommon to see effects dropped that have confidence intervals that range from 0 to more than 30% (e.g., Boardman, Alexander, & Stallings 2011). In practice, this approach has often led to the elimination of shared environmental effects with the consequence of exaggerating genetic and nonshared environmental estimates (Rutter 2002).

Phenotype Specification. Heritability studies are, of course, dependent on the accuracy of the measure of the phenotype. Although phenotype ambiguity is perhaps more obvious in some domains, such as "liberal" and "conservative" political phenotypes and criminal phenotypes, phenotype ambiguity in adverse health outcomes is also an issue. Several scholars have written on this issue in relation to other phenotypes (e.g., Charney 2008; Duster 2003; Press 2006) so I will not belabor the issue, but a few deserve brief mention. Here I focus on two points that relate to the inherent imprecision and social construction involved in the creation of (most) health-related phenotypes: treating quantitative variables as dichotomous and treating ongoing processes as fixed.

Perhaps most importantly, none of the phenotypes that fall under the heading of health or illness are unequivocally biologically-given classifications. The classification of health and sickness is not a unique division of biological reality, but rather there is an inherent degree of social construction and vagueness in these classifications (e.g., Hubbard & Wald 1999). To be sure, some phenotypes clearly involve more social influence in categorization than others. ADHD, for example, is extremely elastic and relies on the presence of symptoms such as "Often talks excessively" and "Is often easily distracted." Not to be captious, but "often" "excessively" and "easily" are rather indefinite terms, and these broad statements conflate domain-specific attributes. However, other classification systems that may seem more given and unproblematic involve a degree of indeterminacy and arbitrariness, such as the classification of obesity (BMI >=30.0) and high blood pressure (systolic greater than 140 or diastolic greater than 90). Biology does not provide a neat cutoff for obesity at 30.0 and hypertension at greater than 140/90. This tendency to treat continuous variables as dichotomous or categorical and biologically given is thus unavoidable in health research and inevitably shapes results (Joseph 2004; Press 2006).

Social forces influence not only the creation of classification systems, but also their application and meaning. Where along the line that we (or others) decide we are "sick" or "healthy" depends on a number of

individual and social characteristics (Hubbard & Wald 1999; Rose 2006). Focusing on mental illnesses, for example, Kessler and colleagues (2005) concluded that "about half of Americans will meet the criteria for a DSM-IV disorder" in their lifetimes, and roughly 26% in the past year. Yet, less than half of individuals meeting diagnostic criteria for these disorders receive treatment and thus are (potentially) diagnosed (Kessler et al. 2001). For many adverse health outcomes, then, the focus of attention is on the *perception* by the individual or his/her medical team that symptoms are problematic and/or qualify as a disease or disorder.

Additionally, it is often the case that in measurement we treat categories as fixed variables (e.g., a "smoker"), neglecting the reality that these are often ephemeral states in a long developmental process. Individuals may often, at different points in their lives, experience episodic states of depression, anxiety, high cholesterol, hypothyroidism, alcoholism, and other conditions. Classifying such processes into static phenotypes ultimately myopically focuses on a tiny snapshot of an ongoing developmental reality.

To be sure, I am not arguing that because most health phenotypes are not given by biology and thus are in some degree social constructions, this implies that these categories are not based on something real, that they should not be used, or that they are not scientific.[12] The fuzziness and plasticity of classification does not invalidate them on their face. For example, the fact that the precise boundary between day and night cannot be demarcated does not mean that night and day cannot be meaningfully differentiated (Leahey & Leahey 1983). Rather, this discussion is intended to remind that phenotypes in heritability studies are not clear, biologically-given categories that are fixed within the individual or across time and space. Instead, phenotype classification necessarily simplifies a much more complicated biological *and* social process (Dupre 2012).

Although a thorough discussion of an additional point is not possible given space limitations, some phenotypes, in my view, are ill-suited for genetic investigation (e.g., "smoking desistance" or "age of first cigarette"), as they are manifestly the result of a multitude of individual traits that combine with numerous social influences in an ongoing process of development (much like "age of first game of golf" or "bacon-eating

[12] Recognizing that science is a social enterprise that is produced by human agents implies that science is to some degree a social construction. This does not imply an anti-realist position or that science is not about truths. Science is about truths, but schemes of classification are not inherent in the nature of most phenomena of interest to social and biological scientists (see Dupre 2012 for an excellent discussion of this point).

desistance"). Smoking studies often utilize samples of nonsmokers, former smokers, and current smokers thereby reifying these as "actual categories of human beings," and "neglect[ing] the profound and unique role of human cognition and self-reflexivity in regard to behavior" (see Press 2006: 143-44).

Summarizing Technical/Statistical Limitations

For at least 80 years scholars have criticized the technical limitations of heritability studies from a classical genetic perspective. Scholars have argued that given these flaws, which altogether have the potential to wildly bias heritability, render the estimates of heritability studies too problematic to be of scientific value (Charney 2008; Joseph 2004; 2006; Lewontin 1974). I concur with these scientists. Regardless of one's views on this issue, however, understanding the assumptions and inherent technical limitations of heritability studies is important given the prevalence of these studies and estimates.

Although I think understanding these methodological limitations is useful, as I discuss below, in my view the *fatal* flaw with heritability studies is conceptual. Although critics have denounced the conceptual model of heritability studies as flawed for decades (Hogben 1933; Lewontin 1974; Wahlsten 1979), recent evidence emerging from molecular genomics evinces the inseparability of genetic from environmental influences and has begun to identify mechanisms through which environmental factors influence genetic activity and thereby shape phenotype development. These recent advances have uprooted nearly all of the assumptions of the classic gene-centric paradigm undergirding heritability studies, including the notions that DNA is the sole biological agent of heritability and that genes are fixed entities that are encapsulated within the cell, produce proteins in a straightforward manner, and serve as the driver/executive of genetic activity within the cell (Charney 2012a; Griffiths & Stolz 2013; Keller 2010). Indeed, this new knowledge has transformed the scientific understanding of the gene and the relationship between genotype and phenotype and marks a shift from the classic genetic paradigm to a postgenomic view (Charney 2012a; Jablonka & Lamb 2006; Keller 2010). The postgenomic paradigm "is characterized by extreme complexity, variability, multilevel reciprocal interactionism, and stochasticity as an inherent property of biological systems, all of which contribute to what might be called the blurring of

boundaries, in particular, the boundary between genes and the environment" (Charney 2012a:332). As I discuss, heritability studies have no place in the postgenomic era.

POSTGENOMIC CHALLENGES TO HERITABILITY STUDIES

"In the emerging postgenomic paradigm, we are confronted with a biological world that is in many ways the opposite of that which has thus far enabled the methodologies of behavioral genetics" (Charney 2012a: 61).

When the Human Genome Project (HGP) was launched in 1990, many scholars anticipated finding a straightforward relationship between the complexity of an organism and its number of genes. With over 100,000 identified proteins in the human body, assuming each gene codes for a different protein, researchers projected that the human genome consisted of at least 100,000 genes. These projections were quite inaccurate. Completion of the HGP in 2001 revealed that the human genome has only roughly 23,000 genes (this number is continuously being revised), slightly more than the fruit fly and the roundworm and less than corn (Charney & English 2012; Claverie 2001; Venter et al. 2001). This finding was clear evidence that the predominant one gene—one protein assumption was incorrect (Silverman 2004). We now know that a single gene can code for multiple proteins, something that is estimated to occur in as much as 90% of all human genes (Charney 2008). These findings combined with the failure of genome-wide association studies (GWAS) to find strong main effects of genes on phenotypes, led to a rethinking about genetic functions, including questioning the utility and meaning of a "gene" (Jablonka & Lamb 2006; Keller 2010).

In response to the flaws inherent in the classic "gene-centric" paradigm, the new postgenomic paradigm has begun to emerge (Charney 2012a). Key to the postgenomic paradigm is the recognition of an interactional, bidirectional relationship among genes, cells, organisms, and environments as well as adaptive developmental plasticity—the capacity of organisms to modify physiological, morphological, or behavioral phenotypes in response to environmental conditions (Charney & English 2013). Rather than being a code, program, or blueprint for development, genes are now understood to be cellular resources (Gottlieb 1997; Keller 2010; Charney 2012a). The activity of the genome is regulated as part of the individual's general developmental-physiological adaptation to environmental signals—a "constant interplay between biology and experience" (Lickliter &

Honeycutt 2003: 820). Although these recent advances have become dominant/influential in biological science, regretfully they have not yet fully penetrated the social and behavioral sciences. Below I briefly discuss advances in molecular genetic knowledge that debunk the genetic model that undergirds heritability studies.

EPIGENETICS

"…[I]t is important to remember that by itself, DNA is an inert, sticky glop." (Hubbard 2013:22)

Genes, while certainly integral to life and variation, are now understood to be inert molecules incapable of doing anything on their own. For many years, genes were thought to be a self-activating code for the production of protein products (the "central dogma" of molecular biology); however, we now know this is not the case. Instead, we now understand that cells are the directors of genetic and other cellular activity, and genes are merely a resource upon which cells draw to respond to environmental signals. In thinking of this revised role of DNA and genetic function, Hubbard and Wald (1999) used an excellent analogy of DNA as a cookbook. Here, genes are recipes, and cells are the cook. In order to make a complex dish, a cook needs a recipe, which is contained in a cookbook (genome). Of course, the cookbook does not determine the recipe, and the recipe does not make the dish or determine how it will turn out. The dish depends on a number of external factors, including those that have nothing to do with the recipe itself or the cook. Such is the relationship among genes, cells, DNA, and the environment. Cookery is also an apt metaphor because it allows for a degree of flexibility that is also inherent in DNA action, as genes are utilized by the cell to respond to environmental changes (Hubbard & Wald 1999).

Although molecular epigenetic research is highly specialized and technical, understanding some basics is useful given that this provides evidence of the effects of environmental and behavioral influences on genetic activity. As I have noted, genes are not self-activating. Instead, DNA has to be transcribed to produce RNA and proteins,[13] but in order to be transcribed it has to be "turned on". The process of gene activation, often called gene expression, is regulated by the *epigenome*—the biochemical regulatory system that can turn on, silence (leave off), or change the transcriptional availability/activity of genes (Charney 2012a; Martiensen, Riggs, & Russo 1996). In a multitude of ways (see e.g., Charney 2012a; Jablonka & Lamb 2006), the epigenome regulates gene expression.

[13] In a process known as *transcription*, the DNA molecule is used by a cell to produce messenger RNA (mRNA), which in turn serves as a template for the synthesis of polypeptides, which form proteins (a process called *translation*).

Most gene regulation occurs in response to the immediate demands of the environment, takes place in time spans ranging from seconds to weeks, and differs between specialized cells (Francis 2011; Jablonka & Lamb 2006). For example, when eating a hamburger, our tongue, stomach, and pancreatic cells will react differently and many cells types will not react at all. With very few exceptions, these different cell reactions are due to epigenetic variation across the cells (Jablonka & Lamb 2006). In recent years, scientists have focused on gene regulation that takes place over much longer intervals. This rapidly advancing field, known as epigenetics,[14] focuses on the mechanisms of gene regulation implicated in changes in gene expression and phenotype, which can last for weeks, months, years, or across the life span and can even be transmitted onto future generations (epigenetic inheritance) (Bagot & Meaney 2010; Bollati & Baccarelli 2010; Charney 2012a; Francis 2011). Importantly for our purposes, epigenetic research reveals that the epigenome is responsive to environmental input (both internal and external to the cell); thus, the environment influences gene activation through the epigenome (Charney 2012a; Jablonka & Lamb 2006).

Although new epigenetic mechanisms are still being uncovered, the most well understood fall into a class known as chromatin markers, often referred to as "epigenetic markers." Two of the most well researched epigenetic markers are histone acetylation (Grunstein 1997) and DNA methylation (Bird 2002). These markers (posttranslational modifications) affect cellular activity and phenotypes by influencing the accessibility of DNA to transcription factors, thereby shaping whether (and to what extent) a gene is activated. Research documenting the effects of environmental factors on epigenetic markers is accumulating, thereby illustrating ways in which gene expression is influenced by the environment (Bagot & Meaney 2010; Charney 2012a; Francis 2011).

Environmental epigenetics focuses on the effects of social-environmental factors such as pollution, nutrition, parental care, and stress on gene regulation (Landecker & Panofsky 2013). Studies in this field (primarily on non-human animals) identify lasting effects of environmental factors on phenotypes through the epigenome. Research

[14] The term epigenetics has undergone many transformations in meaning since the term was coined, and even now continues to be employed loosely and interchangeably with other related concepts (see Jablonka et al. 2009). Conrad Waddington coined the term epigenetics in 1942 to refer to the developmental processes involved the development of a phenotype, recognizing that such a process is not a simple genotype →phenotype relationship. However, the modern use of the term epigenetics is usually narrower, and can be thought of as the study of the processes of gene regulation (e.g., Meloni 2014).

in the emerging field of neuroepigenetics reveals that adverse early environments and injurious social experiences generate epigenetic changes that shape brain architecture and consequently responses to later social experiences and adversities throughout the lifespan (e.g., Hoffmann & Spengler 2014; Sweatt 2013). Recent evidence from animal models suggests that such "experience-dependent DNA memories," inscribed epigenetically, can play key roles in development and stress responsivity (e.g., Hoffman & Spengler 2014; Murgatroyd et al. 2009) and can have transgenerational effects (i.e., environmentally-induced epigenetic changes can be transmitted to offspring; Franklin et al. 2010).

For illustration, in their now famous studies, Weaver and colleagues (2004; 2005; 2006) connected stress phenotypes (HPA responses to stress) in adult rats to early maternal care. Female rats who engaged in more frequent licking and grooming of their pups during the first days of life raised offspring that were less reactive to stress across their lifespan and, among the females, went on to become high-licking and high-grooming mothers themselves. Weaver and colleagues (2005; 2006) linked this early maternal care to epigenetic changes (DNA methylation) in the glucocorticoid receptor promoter, and further work has identified similar epigenetic changes linking early maternal care to hundreds of other genes in the rat pups, demonstrating the manifold biological effects of early care on the developing brain (at least in the rat). Importantly, while these results showed that early-life parenting experiences have a stable and broad effect on several health-related phenotypes, they also revealed that these epigenetic markers are reversible in adulthood (Weaver et al. 2005; 2006) consistent with the notion of organisms' adaptive phenotype plasticity in response to changes in social conditions (e.g., Ellis et al. 2012).

In addition to maternal care, studies of rodents have linked exposure to a number of substances, such as pesticides, fungicides, BPA, cigarette smoke, vehicle exhaust, and cocaine, to methylation or other epigenetic markers with concomitant changes in gene expression and phenotypes, including physiological function, behavioral outcomes, and propensity to cancers and metabolic disorders (e.g., Feil & Fraga 2011; Landecker & Panofsky 2013). For example, pesticide exposure has been linked to sperm defects and subfertility as well as decreases in fearful behaviors (see Skinner 2008). These changes occur without changes in DNA sequences, but

rather through changes in the epigenome (e.g., DNA methylation) associated with genes linked to these phenotypes (Charney 2012a).

While there are many more examples, these studies are sufficient to illustrate the fact that epigenetic research evinces that genetic and environmental effects are interactional and dynamic, and thus differences in phenotypes cannot be separated into genetic versus environmental influences. Epigenetic markers embody the blurred boundaries between genes and environments and the difficulty of classifying molecular mechanisms in gene expression in the G versus E dichotomy. For example, as Charney (2012b) questioned, how would heritability study proponents have us classify experience-dependent epigenetic marks (e.g., chromatin architecture) in the G vs. E dichotomy: As G because chromatin attaches to the genome and plays a role in silencing or turning genes on? Or as E given that chromatin structure is shaped by environmental input? What about the inheritance of environmentally-induced epigenetic reprograming in the absence of the original inducing environment (e.g., pesticide exposure, maternal care): As G because it is inherited? Or as E because it originated with an environmental stimulus? Or is E transmuted to G? These instances of the blurring of boundaries between genes and environments are a reoccurring theme in modern genomics research and point to the utter impossibility of separating "genetic" from "environmental" in development (Charney 2012a).

In short, the model of a distinct, particulate gene that directs cellular activity and performs one job (produces one protein product) independent of environmental input has been discredited in the past decade. The mere presence of a gene as part of a genotype does not ensure that it will be activated. Moreover, although evidence suggest that environmental events that occur early in life (e.g., prenatal care; early maternal care; exposure to toxic substances) tend to produce more pronounced epigenetic effects than those that occur later, methylation and other epigenetic processes continue throughout the lifespan (Francis 2011; Weaver, Meaney, & Szyf 2006) and are involved in (and necessary for) both healthy development as well as disease progression. Epigenetic research demonstrates that the genome is flexible in its expression, and this expression is responsive to context, experience, and developmental history (Griffiths & Stoltz 2013). As Dupré (2012: 3) noted: "One way of

beginning to think about epigenetics is to realize that the genome, as much as the organism, is a process rather than a static thing."

Several assumptions are necessary for heritability studies to be meaningful, but the most crucial of these is the (biological) assumption that genes and environments have identifiably separate effects on phenotypes (Charney 2012a). Biological evidence is clear that this is not how genes work.

ADDITIONAL GERMS FOR THOUGHT: WHAT ABOUT MICROBES?

As I noted, the finding from the HGP that the human genome consists of only ~23,000 genes was a surprise to many, as researchers projected roughly 100,000 genes. However, if the view of what constitutes the "human genome" is expanded, then 100,000 genes is a drastic underestimate (Turnbaugh et al. 2007). The human body is a complex, symbiotic whole comprised of many different microbial organisms (e.g., bacteria and fungi) of different kinds and genomes that live on or inside human tissues. Astonishingly, evidence suggests that within our bodies 90% of the cells are actually microbes, most of which inhabit the gastro-intestinal tract, and 99% of the genes on or under our skin are actually bacterial (Xu & Gordon 2003). Thus, the microbiome contains at least 100 times as many genes as our own genome (Gill et al. 2006).

Over the past decade along with (or in the wake of) new technologies for studying microbiota (a microbial population), interest in the role of the microbiome (the set of microbial genes) in human health has surged (Cho & Blaser 2012). Ambitious large-scale projects, including the U.S. Human Microbiome Project (Turnbaugh et al. 2007) and the European project MetaHIT (Erlich 2011), have begun to advance knowledge on the biological properties and medical significance of the human microbiome and its collective genes (metagenome; Cho & Blaser 2012). These projects in combination with smaller scale efforts have begun to elucidate the composition and development of individuals' microbiomes as well as their influential role in health and development. Several findings have emerged, including the density of the microbiome at various sites in the human body, the extent of inter- and intra-individual variation, and the dynamic and significant interplay between microbial and human cells. Acknowledging the significant role of the microbiome, Gill et al. (2006: 1354) noted that the microbiome "endows us with physiological properties that we have not had to evolve on our own."

Over the course of a lifetime, each individual develops a densely populated microbiome. Individuals, even from the same region and similarly aged, have quite different microbiomes, and differences between individuals appear to be much greater than intra-individual variation (at least over short periods of time; Cho & Blaser 2012).[15] The colonization of the human gut begins at birth, as infants pick up billions of microbes as they move from a sterile womb through a microbe-laden vagina (if a vaginal birth, which of course was the only form of birth throughout most of human history). Thus, an individuals' microbiome appears to be a consequence of this important initial composition from the mother as well as a dynamic interrelationship between environmental (e.g., diet, stress, exposure to antibiotics and probiotics; David et al. 2013; Jakobsson et al. 2010) and biological (likely including genetic) factors throughout development.

Much recent research has focused on the role of intestinal microbes in maintaining health and homeostasis. Scholars have explained that our gut microbiota represent a "virtual inner organ" (Forsythe et al. 2010: 9). Although we are only beginning to understand how microbiome composition and function influences human health, a spate of recent studies suggests that microbes have significant influences (in interaction with genomes and environments) in shaping health outcomes (e.g., Cho & Blaser 2012). In particular, we have clear evidence of the role of gut microbes in processing indigestible foods, such as plant polysaccharides, synthesizing essential amino acids, and therefore obtaining necessary key nutrients from our food (e.g., Gill et al. 2006). More recently, evidence of the influence of gut microbes through gut-brain signaling in homeostasis and nervous system functioning is beginning to accumulate (Mayer 2011; Forsythe et al. 2010). An expanding body of literature explores pathways of influence and links individuals' microbiota to adverse physical health outcomes through complex channels involving the immune system and host metabolism (Forsythe et al. 2010; Grice & Segre 2012). Although this work is still in its infancy, these nascent findings suggest "that there is much communication and interaction between microbial and human cells, and there is growing evidence that (gut) microbes have significant distal influences on health and disease through their effects, for example, on the immune system and on

[15] This is complicated by the fact that research suggests that bacteria from a specific body part have more in common than those from a specific person. Your armpit microbes have more in common with my armpit microbes than they do with your oral microbia (Ding & Schloss 2014).

neurotransmission" (Dupre 2012: 13). At present, gut microbes have been implicated in such diverse outcomes as chronic inflammation, anxiety, mood, aggressiveness, eating disorders, cognitive functioning, diabetes, psoriasis, autism, allergies, and atherosclerosis (Forsythe et al. 2010; Grice & Segre 2012; Mayer 2011; Turker et al. 2014).

This discussion of the role of microbes, which is necessarily abbreviated but hopefully stimulating, is intended to further cloud the reductionist, false genes versus environment dichotomy in shaping human phenotypes. Microbes exist within us and greatly outnumber our cellular and genetic material, and our microbiomes are a consequence of a complex interaction of biological and environmental factors, which consequently mediate and condition the effects of genetic and environmental influences on health. Indeed, emerging research indicates that microbiota influence *host* gene expression (for example, gut microbiota influence the expression of genes involved in host digestive functions and metabolism; Combe et al. 2014).

How would heritability study proponents have us classify the effects of microbes? As G because the microbiome, although not strictly human genes, fulfills functions that we have not had to evolve on our own? As G because microbiota can influence the epigenome and gene expression? Or as E because they are not simply human DNA sequences, even though many are intergenerationally transmitted (through the birth canal)? Moreover, although evidence suggests that individuals' microbiomes become increasingly stable over time (e.g., Forsythe et al. 2010), changes in one's lifestyle (diet, activity level) and medications (e.g., antibiotics) can alter the composition of one's microbiome, along with ontogenetic development (e.g., Cho & Blaser 2012; David et al. 2014). Humans are living systems that exist and develop in a synergetic, commensal relationship with other organisms (and their genomes), and such relationships defy classification into (human) genetic versus environmental influences.

EVEN IF WE COULD: HERITABILITY ESTIMATES LACK UTILITY

Quite aside from their many flaws, heritability estimates have little practical utility, and this reality is recognized at present even by prominent behavioral geneticists. For illustration, Eric Turkheimer, one of the most well-known behavioral geneticists, who developed the "three laws of behavioral genetics," provided a summary of why he thinks heritability studies lack utility: "…it is not about *cause*. Practitioners of the art wanted it to be about

cause, in the sense that the relative magnitudes of the various components were supposed to tell us something about the importance of genetic and environmental *causes* underlying a trait, but they do not" (Turkheimer 2011: 598; emphasis added).

Heritability estimates do not forecast the developmental endpoints for individuals or groups, the consequences of interventions, or the causal processes or mechanisms involved in phenotype variation (Burt & Simons 2014: 246; see also Rutter 2002). A high heritability coefficient says nothing about the amount of genetic influence on a trait (recall the 'eyedness' example), quantifying heritability does not contribute to the identification of specific candidate genes[16] (e.g., Schielzeth & Husby 2014), and heritability implies nothing about the potential for change through environmental alterations (see Joseph's (2004) PKU example).

Moreover, it bears repeating that heritability estimates are time, context, and population-specific averages. A heritability estimate could be analogous to an estimation of the average temperature of North America based on taking the average temperature from randomly selected cities on a given day (Kruger et al. 2008). The usefulness of such an estimate is scanty to naught. Importantly, and in contrast to the average temperature of North America, which has a true value, "[t]he proportion of variance attributable to A [additive genetic variance], C ["shared" environments] and E ["unshared" environments] depend on the variances of A, C and E, and the variances of A, C and E have no 'true' values to be estimated" (Turkheimer 2011: 598).

To be sure, this is not an argument that nothing positive has come out of heritability studies, because whatever their flaws (and there are many) they have likely significantly contributed to the acceptance of genetic influences on various phenotypes among social scientists. However, *at the present state of knowledge*, the numerical estimates of heritability, even if they could be accurately estimated, are neither of practical use nor do they advance scientific understanding of cause (see also, Turkheimer, Pettersson, & Horn 2014).[17] As Chaufan

[16] As Schielzeth and Husby (2014:45-6) note, "there is in general no reason to expect that traits with a high proportion of additive genetic variance should harbor large-effect variants rather than many loci with small effects…From a statistical viewpoint, what matters is a low amount of residual variance, not a high heritability per se…a direct link between heritability and the number and effect size of trait loci is not compelling."

[17] Indeed, Turkheimer (2011:598) noted that "Plomin and Daniels [1987] marked the beginning of the end" of heritability studies as a useful scientific endeavor.

(2008:24) noted, heritability estimates "provide no information relevant to … medical practices and public health policies that we do not already have."

Finally, given their problems, heritability estimates can (and likely have) misinformed and misguided efforts to understand the etiology of and therefore prevent or reduce adverse health outcomes (e.g., Lewontin 1974). For example, the absorption of environmentally-induced, intergenerationally-transmitted epigenetic markings into the G in heritability studies could lead scholars to conclude that the cause is inherited genetic defects, and hence environmental culprits (e.g., pesticides, maternal care) may not be considered (see Charney 2012b).

CUTTING-EDGE TECHNIQUES MEET OUTDATED NATURE VS. NURTURE DICHOTOMY

Biometrical Moderation Models

Advances in statistical modeling over the past decade have enabled heritability researchers to incorporate environmental contingencies into their variance partitioning analyses, "allowing the quantification of phenomena that have traditionally been characterized as gene-environment interaction and correlation" (Kruger et al. 2008: 1485). Several recent studies of adverse health outcomes have incorporated environmental contingencies (e.g., Boardman et al. 2012; Johnson et al. 2010; Kruger et al. 2008; Tuvblad, Grann, & Lechtenstein 2006). For example, Boardman and colleagues (2011: 1517; 2008) utilize biometrical moderation models with various smoking phenotypes, exploring, for example, "the effect of social policy on the extent to which genes influence [variance in] smoking desistance."

Although this might seem to be a useful strategy to overcome the G x E and *r*GE conundrum besetting heritability studies, here is why I think it is not. First, these models are still based on the flawed twin study model (with its various assumptions aforementioned). Second, these studies seek to identify one (or a select few) environmental contingency out of an infinite number of possibilities thereby having a significant risk of omitted variable (environmental contingency) bias. For example, in their study of body mass (BMI), Boardman and colleagues (2012: 370) compared the heritability of BMI across schools (using a twin-study of kinship pairs nested within schools) and examined the relevance of "school differences in both health-related policies and social norms

regarding body size" on heritability estimates using the Add Health data. The authors examined whether various school characteristics (school administrators' views of the severity of deviant behavior problems; punitiveness of school punishments, prevalence of weight loss resources), as well as school norms about body size (average school BMI for students who classify themselves as 'normal weight') and norm enforcement (school variance around this average for 'normal weight' self-classified students) influence the degree to which BMI is heritable. Boardman et al. (2012: 383) found that heritability estimates varied across several of these environmental characteristics, and they concluded that school social norms and institutional policies influence genetic contributions to weight gain and thus "provide new insights into social and institutional factors, while also helping to better implement social policies."

Certainly it is possible, indeed likely, that school-level factors have some effect on BMI and therefore the degree of genetic influences on variation in BMI. However, it is also plausible that wider contextual factors (e.g., community disadvantage, population SES, prevalence of healthy food options in the community) shape school norms. Moreover, the reverse causal pathway may be equally or more significant. In other words, students' BMIs may be driving the observed "school effect." For example, factors exogenous to the school (SES, geographic location, school funding (monies devoted to more expensive, healthier food options)) may be shaping average school BMI, and school policies may be endogenous to these factors. Concomitantly, a community (and its residents') history of disadvantage (or advantage) may have fostered unhealthy (healthy) eating patterns among a population, which, in turn, may engender epigenetic changes transmitted to future offspring that increased the likelihood of metabolic problems and a higher BMI. Analogously, school foodstuffs (e.g., cafeteria meals, snacks, soda availability) influence the food consumption of students, which over time likely influences the microbiomes of the population of students, and, in turn, interacts with genetic and environmental factors to shape metabolism and food preferences, and hence BMI. In short, there are many alternative explanations that are not examined (and some that cannot be examined at present), many of which are not neatly classified as genetic or environmental.

Notably, such estimates are not about mechanisms or cause. In contrast to what some studies argue or imply, identifying variation in heritability across various environments does not shed light on the genetic mechanisms underlying phenotypes (or phenotype variance) nor the potential efficacy of environmental interventions. Genetic mechanisms are not the focus of such studies. [Thus, to use the BMI example, such biometrical moderation models do not indicate "the extent to which the genetic influences on body mass are different," but rather differences in the extent to which such observed variation is credited to additive genetics (Boardman et al. 2012: 370)]. For elucidation, identifying the school factors that shape heritability of BMI can be analogized with the earlier cornfield example. Showing that cornfields with trees around them have lower heritability than those without tells us nothing about the causes of plant height, genetic or environmental. In the same way, identifying school factors tells us nothing about the *causes* of BMI, genetic or environmental. Furthermore, as noted above, a high heritability signifies nothing about the potential success of environmental interventions. Thus, identifying differences in heritability does not inform "debates about the most effective means to improve the health of the public" nor does a high heritability indicate that "policies may be ineffective for those who are more likely to gain weight because of very small differences across their genome" (Boardman et al. 2012: 371).

Moreover, and as should now be clear, my view is that this attempt to separate G from E is impossible. It is not merely that separate G and E effects interact to influence a phenotype in ways that are potentially separable at the population level; instead they do not have separate influences on phenotypes in the first place. As Turkheimer (2011: 600) notes, "individual differences in complex human characteristics do not, in general, have causes, neither genetic nor environmental. Complex human behaviour emerges out of a hyper-complex developmental network into which individual genes and individual environmental events are inputs. The systematic causal effects of any of those inputs are lost in the developmental complexity of the network."

Genome-Wide Complex Trait Analysis

Recently, scientists have developed a new methodology for identifying gene variants in human populations called genome-wide complex trait analysis (GCTA; Yang et al. 2010; 2011), and over the past four years its use for health phenotypes has burgeoned. In contrast to standard heritability estimates, which do not measure genetic

factors and instead rely on various assumptions around genetic relatedness, GCTA studies require that individuals are unrelated and rely on a genetic relationship matrix comprised of common single-nucleotide polymorphisms (SNPs) (Yang et al. 2011). GCTA studies avoid many of the abovementioned technical limitations that result from the fact that standard heritability models measure neither genes nor environments and thus rely on a host of assumptions. Basically, GCTA studies involve scanning hundreds of thousands of SNPs of thousands of (ostensibly) unrelated persons in a sample with the goal of determining whether phenotype similarity can be linked to various SNPs (Yang et al. 2010). The goal is an estimate of the heritability of a phenotype that is accounted for by the additive effects of shared SNPs (Charney 2013).

At present, GCTA results appear to support the arguments of twin and adoption study *critics* as they have yielded significantly lower and in some cases nonsignificant heritability estimates even when substantial twin study heritability estimates are found in the same sample (Charney 2013). For example, in their study of five anxiety-related traits (after a GWAS finding of "no common genetic variants of large effects that contribute to the heritability of these traits"), Trzaskowski and colleagues (2013) reported an average GCTA heritability estimate of 10%, which they noted is less than one fifth of the average twin-study heritability estimate of 55%. Similarly, comparing GCTA and twin study estimates, Trzaskowski, Dale, and Plomin (2013: 1048) concluded: "Behavioral problems in childhood [which included health-related phenotypes such as depression]—whether rated by parents, teachers, or children themselves—show no significant genetic influence using GCTA, even though twin study estimates of heritability are substantial in the same sample…" Thus, GCTA studies provide further evidence that something is awry in the high behavioral genetic heritability estimates of adverse health phenotypes (Burt & Simons 2015).

To be sure, the lower estimates from GCTA studies may be due in part to the fact that they only capture the additive effects of genetic variants that are in linkage disequilibrium with the common SNPs analyzed (rare SNPs are excluded) in genome wide association research (Yang et al. 2010). However, like standard heritability studies, GCTA studies rely on problematic assumptions whose violations can increase the likelihood of spurious

associations and hence flawed estimates, such as population stratification (PS) (Charney 2013).[18] PS raises the possibility of false associations due to gene-environment correlations (i.e., more genetically similar individuals are more likely to experience certain environments, discussed above with the example of skin pigmentation). Notably, research has demonstrated that GCTA studies are particularly vulnerable to PS confounding (e.g., Browning & Browning 2011; Janss et al. 2012). While the developers of GCTA are cognizant of the problems posed by PS and do make attempts to correct for it (Yang et al. 2011; in practice see Boardman et al. 2014), these corrections have been shown to be insufficient (for more information, see Charney 2013).

In sum, like other heritability studies, GCTA studies suffer from methodological limitations that undermine the legitimacy of their findings and augment the likelihood of spurious genetic associations. Ultimately, however, GCTA studies, like other heritability models, rest on an outdated, false gene vs. environment dichotomy. To repeat, the lack of scientific merit in heritability studies is not limited to technical/statistical problems that can be improved over time or can be addressed with new cutting-edge techniques that do not rely on the same dubious methodological assumptions. The problem is conceptual (biological): genes and environments do not have identifiably separate effects on complex phenotypes.

CONCLUSION

"One of the most striking features of the nature-nurture debate is the frequency with which it leads to two apparently contradictory results: the claim that the debate has finally been resolved (i.e., we now know that the answer is neither nature nor nurture, but both), and the debate's refusal to die." (Keller 2010: 1).

In this chapter, I have argued that there is compelling evidence that heritability studies are methodologically flawed, especially for complex adverse health phenotypes. I have also argued, drawing on recent advances in molecular genomics and epigenetics, that heritability studies are grounded on a specious conceptual foundation. Recent advances in molecular genomics have debunked nearly every assumption that underlies heritability studies. This new evidence manifestly supports, indeed proves, the arguments that critics of heritability studies have been

[18] Population stratification refers to the nonrandom distribution of polymorphisms in different populations (ethnic or geographical) due to unique ancestral patterns of migration, mating practices, and reproductive patterns (Charney & English 2013; Hunley, Healy, & Long 2009). Most nonfamilial populations exhibit PS, including those considered relatively homogenous (e.g., among Icelanders; Charney 2013).

making for decades that the goal of partitioning genetic and environmental effects on variance in phenotypes is unsound (e.g., Lewontin et al. 1984). In short, heritability studies attempt the impossible. Furthermore, that these methods are fatally flawed is no great loss to science; heritability estimates lack utility and are not about cause (Rutter 2002; Turkheimer 2011). Joining the arguments of others (e.g., Charney 2012; Lewontin 1974; Turkheimer 2011), I contend that further attempts to estimate heritability for complex phenotypes, given what we now know, are a manifest misuse of scholarly energy and attention (e.g., Chaufan 2008; Joseph 2004; Lewontin 1974). Moreover, given their many flaws, I urge scholars to recognize and acknowledge the problematic nature of existing heritability estimates and end the frequent use of the phrase: "We know from a wealth of behavioral genetic studies that the heritability of [insert health-related phenotype] is roughly xx percent." (Notably, this statement is misguided because heritability estimates are time, space, and population specific.) As I hope is now clear, these estimates are highly dubious at best and should not be presented as facts, given the flawed methodology and misguided conceptual model. No amount of quantitative genetic research can establish the validity of such heritability estimates. "Technically flawed and conceptually unsound models—no matter how often published or repeated—do not by virtue of their numbers make for sound evidence" (Burt & Simons 2014: 252).

Heritability studies rely on an outdated gene-centric biological model. Although such a deterministic model of genetic function may be ideally suited for heritability studies, it is chimerical (Charney 2012a). The remarkable advances in our understanding of genetic function that are most visible in molecular epigenetics debunk the oversimplified model of the genome and genotype-phenotype relationship on which such work relies (Charney & English 2013). Indeed, the more we learn about development, the less meaningful seems any attempt to estimate genetic (vs. environmental) contributions to phenotype variance—and the less important (Burt & Simons 2015).

Notwithstanding my strong objections to heritability studies, I am enthusiastic about biopsychosocial health research that recognizes the interactional, bidirectional relationship between genes, cells, organisms, and environments. The challenge is harnessing these rapid advances in molecular genomics to enhance our

understanding of the etiology of adverse health outcomes. In so doing, we need to leave the unproductive nature vs. nurture debate behind to focus on understanding the mechanisms and developmental processes of adverse health outcomes, "[starting] from the premise that biological and environmental systems are indivisible" (Braun 2004: 143). To understand how these factors dynamically interact requires study of the *processes* of development, rather than simply trying to link genomic variation to some sort of fixed end point (adverse health outcome). Attention should shift from heritability to humans' remarkable degree of adaptive phenotype plasticity (the capacity of a single genotype to support a range of phenotypes) and the biological pathways (e.g., epigenetic effects) that underlie such developmental adaptations.

Notably, in highlighting postgenomic findings and lines of research that recognize the "porousness of the biological with the social," it was not my intent to oversell the evidential strength of fields as nascent as epigenetics and medical microbiology (Meloni 2014: 6). Much in these burgeoning fields remains controversial and debated, and scientists face new challenges as a consequence of the epistemological shift from a dichotomous separation of "genetic" versus "environmental" causes to the "inextricable mixture of social and biological factors typical of the epigenetics and postgenomic landscape" (Meloni 2014: 6). However, this recognition of the difficult methodological and epistemic questions facing molecular genetics is no reason to deny that epigenetic research invalidates the nature/nurture dichotomy or to downplay the great potential of epigenetics in elucidating "the pathways through which the social shapes and is literally inscribed into the body" (Meloni 2014: 6). Indeed, that epigenetics and medical microbiology remain open, debated fields underscores the need for social scientists to engage in this debate from the beginning to shape biosocial research as it "[hesitates] at a crossroads between reductionism and holism" (Morange 2006: 356). As a recent editorial in Nature (2012: 143) opined: "It is time for sociologists and biologists to bury the hatchet and cooperate to study the effects of environmental stress on how people behave."

This matters for medical sociology. The field is at a crossroads where the intersection between biology and the social is ripe for research and theorizing (e.g., Landecker & Panofsky 2013; Meloni 2014). Rejecting heritability studies and the false nature-nurture dichotomy and gene-centric model on which they are grounded

will pave the way for a reconceptualization of the link between the biological and the social in shaping health outcomes, one which is consistent with contemporary bioscientific knowledge. Achieving a comprehensive understanding of these processes will not be easy and will take time. It will also require an integrative approach; social scientists, geneticists, microbiologists, and neuroscientists need to form interdisciplinary alliances to facilitate the holistic study of the development of adverse health outcomes from cell chemistry to whole body physiology, recognizing environmental influences from the cellular to the macrostructual levels.

Heritability studies and estimates of adverse health outcomes have gained a lot of traction and attention in recent years based in part on the belief that they are undergirded by rigorous, state-of-the art science.[19] There is value in showing how this work does not measure up to that billing both methodologically and theoretically (Burt & Simons 2015). Although I, among others, find fault in the heritability model, I hope that those who read this critique take it in the spirit of constructive criticism in which it was offered. I hope this critique advances science.

[19] See Panofsky 2014 for an enlightening discussion of the historical development of the field of behavioral genetics and strategies used to maintain its footing in various scientific fields.

ACKNOWLEDGEMENTS

The author would like to thank Brea Perry and two anonymous reviewers for valuable comments on earlier drafts of the manuscript, and a special thanks to Kara Hannula for extensive feedback on multiple drafts. The arguments presented herein are entirely those of the author and do not reflect the views of those who provided feedback.

REFERENCES

Bagot, R. C., & Meaney, M. J. (2010). Epigenetics and the biological basis of gene x environment interactions. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49, 752–771.

Beckwith, J., & Morris, C. A. (2008). Twin studies of political behavior: Untenable assumptions? *Perspectives on Politics*, 6, 785-791.

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, 135, 885-908.

Belsky, J., & Pluess, M. (2013). Beyond risk, resilience, and dysregulation: Phenotypic plasticity and human development. *Development and Psychopathology*, 25, 1243-1261.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development*, 16, 6-21.

Boardman, J. D., Alexander, K. B., & Stallings, M. C. (2011). Stressful life events and depression among adolescent twin pairs. *Biodemography and Social Biology*, 57,53-66.

Boardman, J. D., Blalock, C. L., Pampel, F. C., Hatemi, P. K., Health, A. C. & Eaves, L. J. (2011). Population composition, public policy, and the genetics of smoking. *Demography*, 48, 1517-1533.

Boardman, J. D., Domingue, B. W., & Daw, J. (2014). What can genes tell us about the relationship between education and health? Social Science & Medicine, Advance online publication doi:10.1016/j.socscimed.2014.08.001

Boardman, J. D., & Fletcher, J. M. (2014). To cause or not to cause? That is the question, but identical twins may not have the answer. *Social Science & Medicine*, Advance online publication. http://dx.doi.org/10.1016/j.socscimed.2014.10.013

Boardman, J. D., Roettger, M. E., Domingue, B. W., McQueen, M. B., Haberstick, B. C., & Harris, K. M. (2012). Gene-environment interactions related to body mass: School policies and social context as environmental moderators. *Journal of Theoretical Politics*, 24, 370-388.

Boardman, J. D., Saint Onge, J. M., Haberstick, B. C., Timberlake, D. S., & Hewitt, J. K. (2008). Do schools moderate the genetic determinants of smoking? *Behavior Genetics*, 38, 234-246.

Braun, L. (2004). Commentary. In C. Garcia Coll, E. L. Bearer, & R. M. Lerner (Eds.), *Nature and Nurture: The Complex Interplay of Genetic and Environmental Influences on Human Behavior and Development* (139-143). New York: Psychology Press.

Breen, F. M., Plomin, R., & Wardle, J. (2006). Heritability of food preferences in young children. *Physiology and Behavior*, 88, 443-447.

Brody, G. H., Lei, M., Chae, D. H., Yu, T., Kogan, S. M., & Beach, S. R. H. (2014). Perceived discrimination among African American adolescents and allostatic load: A longitudinal analysis with buffering effects. *Child Development*, 85, 989-1002.

Browning, S. R., & Browning, B. L. (2011). Letter to the editor: Population structure can inflate SNP-Based heritability estimates. *American Journal of Human Genetics*, 132, 191-193.

Burt, C. H., & Simons, R. L. (2014). Pulling back the curtain on heritability studies: Biosocial criminology in the postgenomic era. *Criminology*, 52, 223-262.

Burt , C. H., & Simons, R. L. (2015). Heritability studies in the postgenomic era: The 'fatal flaw' is conceptual. *Criminology*. Advance online publication..

Burt, C. H., Simons, R. L., & Gibbons, F. X. (2012). Racial discrimination, ethnic-racial socialization, and crime: A micro-sociological model of risk and resilience. *American Sociological Review*, 77, 648-677.

Carlsson, S., Andersson, T., Lichtenstein, P., Michaelsson, K., & Ahlbom, A. (2006). Genetic effects on physical activity: Results from the Swedish Twin Registry. *Medicine and Science in Sports and Exercise*, 38,1396-1401.

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

Charney, E. (2008). Genes and ideologies. *Perspectives on Politics*, 6, 299-319.

Charney, E. (2012a). Behavior genetics and postgenomics. *Behavioral and Brain Sciences*, 35, 331–410.

Charney, E. (2012b). Humans, fruit flies, and automatons. *Behavioral and Brain Sciences*, 35, 381-392.

Charney, E. (2013). Still Chasing Ghosts: A new genetic methodology will not find the "missing heritability". *Independent Science News*. URL to article: http://www.independentsciencenews.org/health/still-chasing-ghosts-a-new-genetic-methodology-will-not-find-the-missing-heritability/

Charney, E., & English, W. (2012). Candidate genes and political behavior. *American Political Science Review*, 106, 1–34.

Charney, E., & English, W. (2013). Genopolitics and the science of genetics. *American Political Science Review*, 107, 382–95.

Chaufan, C. (2008). Unpacking the heritability of diabetes: The problem of attempting to quantify nature and nurture. *International Journal of Critical Statistics*, 2, 23-38.

Cho, I., & Blaser, M. J. (2012). The human microbiome: At the interface of health and disease. *Nature Reviews Genetics*, 13, 260-270.

Colodro-Conde, L., Sanchez-Romera, J. F., & Ordonana, J. R. (2013). Heritability of initiation and duration of breastfeeding behavior. *Twin Research and Human Genetics*, 16, 575-580.

Combe, B. E., Defaye, A., Bozonnet, N., Putheir, D., Royet, J., Leulier, F. (2014). Drosophila microbiota modulates host metabolic gene expression via IMD/NF-κB Signaling. *PLoS ONE*, 9, e104120.

Corby, P. M. A., Biesbrock, A., Gerlach, R., Corby, A. L., Moreira, A., Schork, N. J., & Bretz, W.A. (2014). Treatment responses to tooth whitening in twins. *Twin Research in Human Genetics*, 17, 23-26.

Cronk, N. J., Slutske, W. S., Madden, P. A. F., Bucholz, K. K., Reich, W., & Heath, A. (2002). Emotional and behavioral problems among female twins: An evaluation of the equal environments assumption. *Journal of the American Academy of Child and Adolescent Psychology*, 41, 829–37.

Crusio, W. E. (2012). Heritability estimates in behavior genetics: Wasn't that station passed long ago? *Behavioral and Brain Sciences*, 35, 361–62.

David, L. A., Materna, A. C., Friedman, J., Campos-Baptista, M. I., Blackburn, M. C., Perrotta, A., Erdman, S. E. & Alm, E. J. (2014). Host lifestyle affects human microbiota on daily timescales. *Genome Biology*, 15, R89.

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., Ling, A. V., Devlin, A. S., Varma, Y., Fischbach, M. A., Biddinger, S. B., Dutton, R. J., & Turnbaugh, P. J. (2013). Diet rapidly and reproducibly alters the human gut microbiome. *Nature*, 505, 559-563.

Ding, T., & P.D. Schloss. (2014). Dynamics and associations of microbial community types across the human body. *Nature*, 509, 357-360.

Duster, T. (2003). *Backdoor to Eugenics*. New York: Routledge.

Eaves, L. J. (1988). Dominance alone is not enough. *Behavior Genetics*, 18, 27–33.

Ellis, B. J., Del Guidice, M., Dishion, T.J.. Figueredo, A. J., Gray, P. et al. 2012. The evolutionary basis of risky adolescent behavior: Implications for science, policy, and practice. *Developmental Psychology*, 48,598-623.

Erlich, S. D., & The MetaHIT Consortium. (2011). MetaHIT: The European Union Project on Metagenomics of the Human Intestinal Tract. In K.E. Nelson (Ed.), *Metagenomics of the Human Body*. (Pp.307-316). New York: Springer.

Feil, R., & Fraga, M. F. (2011). Epigenetics and the environment: Emerging patterns and implications. *Nature Reviews Genetics*, 13, 97-109.

Franklin, T. B., Russig, H., Weiss, I. C., Graff, J., Linder, N., Michalon, A., Vizi, S., & Mansuy, I. M. (2010). Epigenetic transmission of early stress across generations. *Biological Psychiatry*, 68, 408-415.

Forsythe, P., Sudo, N., Dinan, T., Taylor, V. H., & Bienenstock, J. (2010). Mood and gut feelings. *Brain, Behavior, and Immunity*, 24, 9-16.

Fowler, J. H., Baker, L. A., & Dawes, C. T. (2008). Genetic variation in political participation. *American Political Science Review* 102: 233-248.

Francis, R. C. (2011). *Epigenetics: The Ultimate Mystery of Inheritance*. New York: Norton.

Gill, S. R., Pop, M., DeBoy, R. T., Eckburg P. B., Turnbaugh, P. J., Samuel, B. S., Gordon, J. I., Relman, D. A., Fraser-Liggett, C. M., & Nelson, K. E. (2006). Metagenomic analysis of the human distal gut microbiome. *Science*, 312, 1355-1359.

Gottlieb, G. (1997). *Synthesizing Nature-Nurture: Prenatal Roots of Instinctive Behavior*. Mahway, NJ. Erlbaum.

Gottlieb, G. (1998). Normally occurring environmental and behavioral influences on gene activity: From central dogma to probabilistic epigenesist. *Psychological Review*, 105, 792-802.

Grice, E. A., & Segre, J. A. (2012). The human microbiome: Our second genome. *Annual Review of Genomics and Human Genetics*, 13, 151-170.

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature*, 389, 349-352.

Harris, J. R. (1998). *The Nurture Assumption: Why Children Turn Out the Way They Do*. New York: Free Press.

Hirsch J. (1997). Some history of heredity-vs-environment, genetic inferiority at Harvard(?), and *The* (incredible) *Bell Curve*. *Genetica*, 99, 207-224.

Hirsch, J. (2004). Uniqueness, diversity, similarity, repeatability, and heritability. In C. Garcia Coll, E. L. Bearer, and R. M. Lerner (Eds.) *Nature and nurture* (127-138). New York: Psychology Press.

Hogben, L. (1933). *Nature and Nurture*. London: Williams and Norgate.

Hoffmann, A., & Spengler, D. (2014). DNA memories of early social life. *Neuroscience,* 264, 64-75.

Horwitz, A. V., Videon, T. V., Schmitz, M. F., & Davis, D. (2003). Rethinking twins and environments: Possible social sources for assumed genetic influences in twin research. *Journal of Health and Social Behavior,* 44, 111-129.

Hubbard, R. (2013). The mismeasure of the gene. In S. Krimsky & J. Gruber (Eds.), *Genetic Explanation: Sense and Nonsense*. Cambridge, MA: Harvard University Press.

Hubbard, R., & Wald, E. (1999). *Exploding the Gene Myth*: Boston, MA: Beacon Press.

Hunley, K. L., Healy, M. E., & Long, J. C. (2009). The global pattern of gene identity variation reveals a history of long-range migrations, bottlenecks, and local mate exchange: Implications for biological race. *American Journal of Physical Anthropology*, 139, 35-46.

Jablonka, E., & Lamb, M. J. (2006). *Evolution in Four Dimensions. Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge, MA: MIT Press.

Jackson, D. D. (1960). A critique of the literature on the genetics of schizophrenia. In D. D. Jackson (Ed.), *The Etiology of Schizophrenia*. New York: Basic Books.

Jakobsson, H. E., Jernberg, C., Andersson, A. F., Sjolund-Karlsson, M., Jansson, J. K., & Engstrand, L. (2010). Short-Term Antibiotic Treatment Has Differing Long-Term Impacts on the Human Throat and Gut Microbiome. *PLoS One*, 5, e9836.

Janss, L., de los Campos, G., Sheehan, N., & Sorensen, D. (2012). Inferences from genomic models in stratified populations. *Genetics,* 192, 693-704.

Johnson, W., Kyvik, K. O., Mortensen, E. L., Skytthe, A., Batty, G. D., & Deary, I. J. (2010). Education reduces the effects of genetic susceptibilities to poor physical health. *International Journal of Epidemiology*, 39, 406-414.

Joseph, J. (2004). *The Gene Illusion: Genetic Research in Psychiatry and Psychology Under the Microscope*. New York: Algora.

Joseph, J. (2006). *The Missing Gene: Psychiatry, Heredity, and the Fruitless Search for Genes*. New York: Algora.

Kaplan, G., & Rodgers, L. J. (2003). *Gene Worship: Moving Beyond the Nature/Nurture Debate over Genes, Brain, and Gender.* New York: Other Press.

Keller, E. F. (2010). *The Mirage of a Space between Nature and Nurture*. Durham, NC: Duke University Press.

Keski-Rahkonen, A., Viken, R. J., Kaprio, J., Rissanen, A., & Rose, R. J. (2004). Genetic and environmental factors in breakfast eating patterns. *Behavior Genetics*, 34, 503–514.

Kessler, R. C., Berglund, P. A., Bruce, M. L., et al. (2001). The prevalence and correlates of untreated serious mental illness. *Health Services Research*, 36, 987-1007.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62, 617-627.

Krieger, N. (2000). *Discrimination and health*. In L. F. Berkman & I. Kawchi (Eds.), *Social Epidemiology* (Pp. 36-75). Oxford University Press.

Kruger, R. F., South, S., Johnson, W., & Iacono. W. (2008). The heritability of personality is not always 50%: Gene-environment interactions and correlations between personality and parenting. *Journal of Personality*, 76, 1485-1522.

Landecker, H., & A. Panofksy. (2013). From social structure to gene regulation and back: A critical introduction to environmental epigenetics and society. *Annual Review of Sociology*, 39, 333-357.

Leahey, T. H., & Leahey, G. E. (1983). *Psychology's Occult Doubles: Psychology and the Problem of Pseudoscience*. Chicago: Nelson-Hall.

Lewontin, R. C. (1974). The analysis of variance and the analysis of causes. *American Journal of Human Genetics,* 26, 400–411.

Lewontin, R. C. (1986). Race and intelligence, In N. J. Block and G. Dworkin (Eds.), *The IQ Controversy* (pp. 107-112). New York: Pantheon Books.

Lewontin, R. C., Rose, S., & Kamin, L. J. (1984). *Not in our Genes: Biology, Ideology, and Human Nature*. New York: Pantheon Books.

Mayer, E. A. (2011). Gut feelings: The emerging biology of gut-brain communication. *Nature Reviews Neuroscience,* 12, 453-466.

Meier, M. H., Slutske, W. S., Heath, A. C., & Martin, N. G. (2010). Sex differences in the genetic and environmental influences on childhood conduct disorder and adult antisocial behavior. *Journal of Abnormal Psychology*, 120, 377-388.

Meloni, M. (2014). The social brain meets the reactive genome: Neuroscience, epigenetics, and the new social biology. *Frontiers in Human Neuroscience*, 8, 309.

Moffitt, T. E., & Beckley, A. (2015). Abandon twin research? Embrace epigenetic research? Premature advice for criminologists. *Criminology*, advance online publication.

Moore, D.S. (2014). Big B, little b: Myth #1 is that Mendelian genes actually exist. In S. Krimsky & J. Gruber (Eds.), *Genetic Explanations: Sense and Nonsense* (pp.43-50). Cambridge MA: Harvard University Press.

Morris-Yates, A., Andrews, G., Howie, P., & Henderson, S. (1990). Twins: A test of the equal environments assumption. *Acta Psychiatrica Scandinavica*, 81, 322–326.

Murgatroyd, C., Patchev, A. V.,Wu, Y., Micale, V., Bockmühl, Y., Fischer, D., Holsboer, F., Wotjak, C. T., Almeida, O. F. X., & Spengler, D. (2009). Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nature Neuroscience,* 12, 1559-1568.

Neale, M. C., & Cardon, L. R. (1992) *Methodology for Genetic Studies of Twins and Families*. Boston, MA: Kluwer Academic.

Plomin, R. (2011). Commentary: Why are children in the same family so different? Non-shared environment three decades later. *International Journal of Epidemiology,* 40, 582–92.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from each other. *Behavioral and Brain Sciences*, 10: 1-16.

Plomin, R., DeFries, J. C., Knopik, V. S.,& Neiderhiser, J. M. (2012). *Behavioral Genetics*, 6th ed. New York: Worth.

Press, N. (2006). Social construction and medicalization: Behavioral genetics in context. In E. Parens, A.R. Chapman, & N. Press (Eds.), *Wrestling with Behavioral Genetics: Science, Ethics, and Public Conversation* (pp.131-149). Baltimore: John Hopkins University Press.

Reitveld, M. J. H., Hudziak, J. J., Bartels, M., van Beijstervelt, C. E. M., & Boomsma, D. I. (2004). Heritability of attention problems in children: Longitudinal results form a study of twins, age 3 to 12. *Journal of Child Psychology and Psychiatry*, 45, 577-588.

Rende, R., Slomkowski, C., Lloyd-Richardson, E., & Niaura, R. (2005). Sibling Effects on Substance Use in Adolescence: Social Contagion and Genetic Relatedness, *Journal of Family Psychology*, 19, 611-618.

Rowe, David C. (1994). *On the Limits of Family Influence: Genes, Experience, and Behavior*. New York: Guilford Press.

Rutter, M. (2002). Nature, nurture, and development: From evangelism through science toward policy and practice. *Child Development*, 73, 1-21.

Rutter, M. (2006). *Genes and Behavior: Nature-Nurture Interplay Explained*. Malden, MA: Blackwell Publishing.

Rutter, M. (2007). Gene-environment interdependence. *Developmental Science,* 10, 12–18.

Rutter, M., & Plomin, R. (1998). Child development, molecular genetics, and what to do with genes once they are found. *Child Development*, 69, 1223-1242.

Saudino, K. J., Ronald, A., & Plomin, R. (2005). The etiology of behavior problems in 7-year-old twins: Substantial genetic influence and negligible shared environmental influence for parent ratings and ratings by same and different teachers. *Journal of Abnormal Child Psychology*, 33,113–130.

Scheilzeth, H., & Husby, A. (2014). Challenges and prospects in genome-wide quantitative trait loci mapping of standing genetic variation in natural populations. *Annals of the New York Academy of Sciences*, 1320, 35-57.

Segal, N. L., & Johnson. W. (2009). Twin studies of general mental ability. In Y. Kim (Ed.) *Handbook of Behavior Genetics* (pp. 81-100). London: Springer.

Skinner, Michael K. (2008). What is an epigenetic transgenerational phenotype: F3 or F2? *Reproductive Toxicology*, 25, 2-6.

Spencer, M. B., & Harpalani, V. (2004). In C. Garcia Coll, E.L. Bearer, & R.M. Lerner (Eds.), *Nature and Nurture* (pp.53-77). New York: Psychology Press.

Stoltenberg, S. F. (1997). Coming to terms with heritability. *Genetica* ,99, 89-96.

Stoolmiller, M. (1999). Implications of the restricted range of family environments for estimates of heritability and nonshared environment in behavior-genetic adoption studies. *Psychological Bulletin* 125:392–409.

Suhay, E., Kalmoe, N. P. (2010). The equal environment assumption in twin studies of political traits: Social confounds and suggested remedies. Unpublished manuscript: *http://sites.lafayette.edu/suhaye/files/2011/01/Violations-of-the-EEA-Suhay-Kalmoe.pdf*

Sweatt, J. D. (2013). The emerging field of neuroepigenetics. *Neuron,* 80, 624-632.

Svedberg, P., Ropponen, A., Alexanderson, K., Lichtenstein, P., & Narusyte, J. (2012). Genetic susceptibility to sickness absence is similar among women and men: Findings from a Swedish birth cohort. *Twin Research and Human Genetics* 5: 642-648.

Turker, N., Toh, Z. Q. Karagiannis, T. C., & Liccidardi, P. V. (2014). Anti-inflammatory effects of probiotics and their metabolites: Possible role for epigenetic effects. In N. Maulik and T. Karagiannis (Eds.), *Molecular Mechanisms and Physiology of Disease: Implications for Epigenetics and Health* (pp. 127-150) New York: Springer.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, 9, 160-169.

Turkheimer, E. (2004). Spinach and ice cream: Why social science is so difficult. In L.F. DiLalla (Ed.) *Behavior Genetics Principles: Perspectives in Development, Personality, and Psychopathology* (pp.161-189). Washington, DC: American Psychological Association.

Turkheimer, E. (2011). Commentary: Variation and causation in the environment and genome. *International Journal of Epidemiology*, 40, 598-601.

Turkheimer, E., Pettersson, E., & Horn, E. E. (2014). A Phenotypic Null Hypothesis for the Genetics of Personality. *Annual Review of Psychology*, 65, 515-540.

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., & Gordon, J. I. (2007). The Human Microbiome Project. *Nature*, 449, 804-810.

Tuvblad, C., Grann, M., & Lichtenstein, P. (2006). Heritability for adolescent antisocial behavior differs with socieoeconomic status: Gene-environment interaction. *Journal of Child Psychology and Psychiatry*, 47, 734-743.

Wahlsten, Douglas 1979. A critique of the concepts of heritability and heredity in behavioral genetics. In eds. J.R. Roye & L.P. Mos (Eds.), *Theoretical Advances in Behavioral Genetics*. Alphen aan den Rijn, the Netherlands: Sijthoff & Noordhoff.

Wahlsten, Douglas. (1990). Insensitivity of the analysis of variance to heredity-environment interaction. *Behavioral and Brain Sciences* 13: 109-120.

Wahlsten, D. (1994). The intelligence of heritability. *Canadian Psychology*, 35, 244-258.

Weaver, I. C.G., Cervoni, N., Champagne, F. A., et al. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, 7, 847-854.

Weaver, I. C. G., Champagne, F. A., Brown, S. E., et al. (2005). Reversal of maternal programming of stress responses in adult offspring through methyl supplementation: Altering epigenetic marking later in life. *The Journal of Neuroscience*, 25, 11045-11054.

Weaver, I. C. G., Meaney, M. J., & Szyf, M. (2006). Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors in the offspring that are reversible in adulthood. *PNAS*, 103, 3480-3485.

Williams, D. R. (1999). Race, socioeconomic status, and health: The added effects of racism and discrimination. *Annals of the New York Academy of Sciences*, 896, 173-188.

Yang, Jian, Beben Benyamin, Brian P. McEvoy, et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42, 565-569.

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., Heath, A. C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J. G., Lozupone, C. A., Lauber, C., Clemente, J. C., Knights, D., Knight, R., & Gordon, J. I.. (2012). Human gut microbiome viewed across age and geography. *Nature,* 486 222-228.