European Journal of Personality, *Eur. J. Pers.* **25**: 254–266 (2011)
Published online (wileyonlinelibrary.com) **DOI**: 10.1002/per.836

Heritability in the Era of Molecular Genetics: Some Thoughts for Understanding Genetic Influences on Behavioural Traits

WENDY JOHNSON[1,2]*, LARS PENKE[1] and FRANK M. SPINATH[3]

[1]*Centre for Cognitive Ageing and Cognitive Epidemiology and Department of Psychology, University of Edinburgh, Edinburgh, UK*
[2]*Department of Psychology, University of Minnesota, Minneapolis, Minnesota, USA*
[3]*Department of Psychology, Saarland University, Saarbruecken, Germany*

Abstract: Genetic influences on behavioural traits are ubiquitous. When behaviourism was the dominant paradigm in psychology, demonstrations of heritability of behavioural and psychological constructs provided important evidence of its limitations. Now that genetic influences on behavioural traits are generally accepted, we need to recognise the limitations of heritability as an indicator of both the aetiology and likelihood of discovering molecular genetic associations with behavioural traits. We review those limitations and conclude that quantitative genetics and genetically informative research designs are still critical to understanding the roles of gene-environment interplay in developmental processes, though not necessarily in the ways commonly discussed. Copyright © 2011 John Wiley & Sons, Ltd.

Key words: genetic influences; twin study; heritability

Much is often made of new findings of the presence of genetic influences on measures of behavioural traits or psychological constructs. We live in an age of fascination with genetic mechanisms. Psychological and medical researchers seek the genetic determinants of illnesses and undesirable behaviours, with the aim of eventually developing appropriate biochemical interventions. Sociologists, historians and even economists seek evolutionary explanations for current social and behavioural patterns, with the aim of eventually developing policies that accurately reflect human nature. Thus, findings of genetic influences on new measures of behaviours or psychological constructs are often used to justify claims of the validity of the measures, their bases in biology and, by extension, the reality of the constructs, their evolutionary significance, the appropriateness of policies accepting present social circumstances as inevitable and the expense of molecular genetic search for the specific genes involved.

For example, in developing their Five Factor Theory of personality, McCrae et al. (2000) took the strong position that personality traits are 'endogenous dispositions, influenced not at all by the environment' (page 175). Though acknowledging it to be an oversimplification, they justified this position by pointing to the heritability of personality—not only of factors such as extraversion and neuroticism which are generally recognised to be temperamental but also of the other major dimensions in their Five Factor Theory, and even of more specific aspects of personality such as their underlying facets. They noted that the same studies that show that personality is heritable also show that it is subject to environmental influences but emphasised that these environmental influences are generally nonshared, thus confounded with measurement error and likely to be much smaller than estimated. They even pointed out that some of the legitimately nonshared environmental influences could lead to biological determination as well, because of 'prenatal hormonal environment, minor brain damage or infection, or the imperfect operation of genetic mechanisms' (page 176), thus rendering the proportion of personality variation potentially subject to manipulation by the postnatal environment even smaller.

More recently, in a journal as high profile as the Proceedings of the National Academy of Sciences USA, Wallace, Cesarini, Lichtenstein, and Johannesson (2007) reported that the heritability of responder behaviour in a game-theoretical experiment was 42% in the Swedish Twin Registry. This behaviour was willingness to deviate from maximally personally advantageous game strategies in order to punish opponent behaviour perceived to be unfair. The authors pointed to the consistency of their findings that this behaviour was heritable with those from other studies of economically relevant social attitudes. They also pointed to other researchers' observations of increased testosterone levels and activation in the dorsolateral prefrontal cortex during game performance as evidence of the biological significance of their findings. They noted that their findings had important implications for economic theory and policy as well as for understanding behaviour in experimental conditions and how observed patterns of game responding behaviour may have evolved.

Rushton, Bons, and Hur (2008) carried such inferences even further. They demonstrated that they could obtain general factors of personality in several samples and that these general factors of personality showed substantial

*Correspondence to: Wendy Johnson, Centre for Cognitive Ageing and Cognitive Epidemiology and Department of Psychology, 7 George Square, Edinburgh, UK EH8 9JZ. E-mail: wendy.johnson@ed.ac.uk

Received 13 April 2010
Revised 15 December 2010, Accepted 15 December 2010

heritability. Noting that the genetic influences appeared to be nonadditive (subject to interactions among gene loci), they concluded that they had obtained evidence for evolutionary selection effects on the general factor of personality, with those higher on this factor likely evidencing higher reproductive fitness.

There is no doubt about the findings of heritability of behavioural traits, whether they be personality or ability in specific or in general, or particular patterns of activities such as game response, or TV viewing. In the last 30 years, the field of behaviour genetics has accomplished a paradigm shift in psychology, biology and sociology, primarily through the estimation of heritability in trait after trait after trait. It is now generally accepted that behavioural traits are subject to genetic influences and genetic influences on behavioural traits are ubiquitous, so much so that Turkheimer (2000) enshrined this as the First Law of Behaviour Genetics. This general prevalence of genetic influences is of critical importance for our understanding that major sources of behavioural differences lie in differences intrinsic to each individual, and therefore, it is a great accomplishment of behaviour genetics that they are now so firmly established and widely accepted. But what exactly does it mean for the underlying biology, what does it have to do with evolution, and what does it tell us about the inevitability of life outcomes and social structures if all behavioural traits are moderately heritable? This is a more subtle question. We perceive that fascination with the presence of genetic influences on behavioural traits has outstripped awareness of the subtleties involved in heritability statistics among many, and the time has come to revisit such details, particularly in the light of recent developments in molecular genetics. That is the purpose of this target article. It is in no sense intended to question the validity of heritability statistics as indicators of the presence of genetic influences on behavioural and psychological traits but to sharpen the ability of psychologists to make effective interpretive use of the information they provide.

EVERYTHING IS HERITABLE

Genetic influences underlie all behavioural traits. They should thus no longer be surprising or even individually noteworthy, as they form the foundation of individual differences in responses to all forms of stimuli. In fact, they are so ubiquitous that even nonsensical arbitrary collections of items reveal them. For example, we randomly grouped collections of 20 items from well-validated personality scales, the NEO-PI-R (NEO Personality Inventory; Costa & McCrae, 1992, 240 items assessing five broad domains with eight facets each) and the Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008, 300 items assessing 11 traits that form three broad domains). They showed significant mean heritability estimates of .31 and .28 in two sizable samples of twins, not much lower than the values typical of the carefully designed intended scales. This empirical demonstration makes clear that innate characteristics are always involved in any systematically manifested behaviour: the specific situation and the actor's environmental history for both matter but so does the genotype of the actor. Moreover, outside the lab (and even inside it, as participation in psychology experiments is not random in the population), the actor's genotype is involved in the likelihood that any specific situation will even be encountered, let alone how the actor will behave in that situation once it is encountered. Anyone could be mugged and responses to mugging will vary, but people who have a tendency to walk alone at night in dangerous areas are more likely to end up displaying their individual responses to mugging. As psychologists, we want to understand why people behave as they do, which means understanding not only the behaviour typically displayed but how people come to be in positions in which the behaviour is relevant. The important question is no longer whether genetic influences are involved in this—by now, we know that they are—but how they are involved. Ignoring or disregarding genetic influences is impossible if any true understanding is to be achieved. Estimating heritability provides a rough and ready way of cutting the processes involved into genetic and environmental pieces, and the persistent observation that the genetic piece is not small has real meaning. But heritability alone does not help us identify and understand the processes that generate patterns of behaviour nor what can or cannot be done about them.

UNDERSTANDING THE LIMITATIONS OF HERITABILITY ESTIMATES

The limitations of heritability estimates for understanding underlying biology have long been known to behaviour geneticists but not necessarily to the many social scientists who are becoming newly interested in the presence of genetic influences on behavioural traits. Fortunately for psychologists, many of the limitations of heritability estimates for understanding biology are related to psychometric principles, so the concepts involved should have a familiar ring. Simply put, better measurement in and of itself generates higher heritability, but that alone tells us nothing about underlying biology. We begin our discussion by exploring the relations between psychometric principles and heritability estimates and then move to a discussion of heritability's genetic and broader biological implications.

These two sections have a common theme: the ubiquity of the presence of substantial genetic influences on psychological traits is important for understanding their emergence and development, but the magnitudes of the heritability estimates for specific traits are not. As the discussion will make clear, both measurement properties and the biological mechanisms involved in traits influenced by many genes act to make heritability estimates of psychological traits congregate in the middle of their possible distributions: perhaps 30–60%. Thus, many are very similar to each other. At the same time, those that differ from this very general pattern are as likely outliers for psychometric or sampling reasons as for biological/genetic or environmental reasons, and there is generally no straightforward way of distinguishing among the possible reasons. For example, the heritabilities of our scales based on

DOI: 10.1002/per

randomly grouped personality items were somewhat lower than those of the carefully designed intended scales. In this case, because we know how the scales were compiled and have a direct basis of comparison in the form of the intended scales, we can be reasonably certain that it was because of the unusually large amount of error variance in these incoherent scales, but if we did not have a reason to suspect this explanation, we would not be able to distinguish error variance as a reason for lower heritability from some biological property such as lower genetic penetrance. Moreover, it is very easy to demonstrate that specific environmental circumstances often exert considerable moderating effects so that heritability estimates differ even within samples, and studies providing such demonstrations are becoming increasingly common (e.g., Agrawal et al., 2009; Johnson, Deary, & Iacono, 2009; McCaffery, Papandonatos, Bond, Lyons, & Wing, 2009). Traits that show high heritabilities (even those well above 60% such as height) also can demonstrate pronounced secular trends. For these reasons, particular values of heritability estimates, whether within individual samples or in comparisons between traits (and the attendant estimates of shared and nonshared environmental influences that are interdependent of them) matter little in psychology once we have accepted that all behavioural traits are heritable.

COMMON PRINCIPLES OF PSYCHOMETRICS AND HERITABILITY

Of course, the random personality item 'scales' we created were not nonsense. The items we randomly grouped were all carefully selected to be markers of personality traits for which we have substantial evidence from many sources. Truly nonsensical collections of items such as random number assignments of Likert scores would not generate perceivable heritability. There are three processes intimately bound up with the basic psychometric principles of scale construction; however, that will, generate heritability, and they will do so without regard to biological relevance.

The first is aggregation. One basic psychometric principle of scale construction is that individual items are always noisy indicators of the construct of interest. Scales require a number of items sufficient to allow the independent sources of error in individual items to cancel each other out through aggregation. This way, the items more clearly reveal the construct of interest they share (Rushton, Brainerd, & Pressley, 1983). In the 'scales' we created, described earlier, that shared construct of interest probably was the general personality factor that has recently been described by Musek (2007) and Rushton et al. (2008). The heritability of our scales, however, does not help us to evaluate whether the general factor of personality is anything other than a statistical artefact, and in particular, whether it should be characterised as a reflection of general life history strategy (Rushton, 1985), evolutionary fitness (Miller, 2007), social desirability (Bäckström, Björklund, & Larsson, 2008), an informant-specific response bias (Riemann & Kandler, 2010), personal resilience, or anything else. Note, though, that relative absence of heritability, where we have some basis of comparison, could cast doubt on some of these possible explanations (Riemann & Kandler, 2010).

Even individual items tend to show heritability (Neale, Rushton, & Fulker, 1986), but they tend to be less heritable than scales, in part because of the principle of aggregation (Johnson, Gangestad, Segal, & Bouchard, 2008). The aggregation principle is also involved in heritability at the molecular genetic level. Polygenic traits tend to show heritability in part because the traits themselves are the aggregations of the expression of many genetic polymorphisms. The more broadly we consider a trait, the more likely it is to display a good solid heritability that is rather stable across populations. Thus, longer and more general scales tend to show higher heritability than do shorter and more narrowly focused scales, all else being equal. The 'all else being equal' condition is important in evaluating this statement. All else rarely is equal, and differences in internal consistency in scales of equal length would be an example of the kinds of differences that would impact the magnitudes and stabilities of heritability estimates. Thus, for example, the general factor of cognitive ability tends to show higher heritability than do tests of specific cognitive abilities (e.g., Johnson et al., 2007) because of the pervasively large proportion of variance for which it accounts in most datasets, whereas the much weaker general factor of personality does not (e.g., Riemann & Kandler, 2010).

The second process that contributes to heritability is test–retest reliability, and its place in the psychometric principles of scale construction is also clear. Scales can only be valid if people's responses are consistent over relevant time periods. The items we used to generate our arbitrary 'scales' all came from personality inventories with good test–retest reliability. Though item responses still are less reliable over time than the scales to which they are intended to contribute, the fact that they do have test–retest reliability alone contributed something to the specific heritability estimates for our arbitrary 'scales'. One way to see this is to consider a 'scale' formed by collecting easily measurable, reliable, highly heritable but likely biologically independent traits such as height, body fat percentage, skin colour on a gradient from light to dark, ability to taste bitterness, ability to process the chemicals in cilantro and quality of tooth enamel. Scores on this 'scale' would be heritable because of the reliable heritability of each of the items, and it would be easy to estimate the reduction in heritability that would result from introduction of a non-heritable item such as primary native language. Thus, temporally more reliable scales tend to show higher heritability, again all else being equal (cf. McCrae, Kurtz, Yamagata, & Terracciano, 2010).

The third process that contributes to heritability is frequency of response. Heritability is a property of variance, so there will be no heritability when everyone responds in the same way. For example, textbooks commonly use the presence of two eyes across most of the animal kingdom as an illustration of the fact that the heritability of a clearly genetically influenced trait is 0 when there is no variability. Understanding of the contribution of frequency of response (often known as item difficulty) to psychometric scale construction is relatively recent. Its role in heritability estimates

DOI: 10.1002/per

of psychological traits is probably even less well appreciated. Classical test theory suggests that the most reliable and valid scales consist of items of moderate endorsement frequency, but development of item response theory has made it clear that we improve accuracy of measurement and thus validity when scales include items that reliably tap the full distribution of the underlying construct. Items at the extremes of the distribution tend to have rather skewed response patterns, however, as most people endorse the items on one end and few people endorse the items on the other end. Response skew tends to reduce the heritabilities of these items because there is little difference among participants' responses, let alone between the responses of people with different degrees of biological relationship on which most heritability estimates are based (Wicherts & Johnson, 2009). To the extent they have been explored using the item response theory; however, both the NEO-PI-R and MPQ contain relatively large numbers of items of medium endorsement frequency, so the specific magnitudes of the heritability estimates for our arbitrary 'scales' could in part be attributed to the medium endorsement frequency of the items.

Like aggregation, this also has an analogue in molecular genetics. Genetic polymorphisms differ in frequency, and individual polymorphisms that are very infrequent in a population will contribute little to the manifestation of heritability, though they could contribute dramatically to the phenotype involved. The single gene that is the major cause of Huntington's Disease is an example of this. Given our current lack of medical interventions to prevent disease onset, presence of the deleterious allele completely determines development of the disease, but this allele is so rare that heritability is effectively 0 in any population sample. There are techniques that can be used to estimate the heritabilities of traits with very skewed population distributions, but these are not the techniques most commonly used to estimate the heritabilities of behavioural traits (Plomin, DeFries, McClearn, & McGuffin, 2007). Response frequencies contribute to the magnitudes of heritability estimates for practical rather than theoretical reasons. That is, most measures of psychological traits tap mostly the moderate range of response frequency. This is an issue separate from the techniques that have been used to estimate the heritabilities of scales, whether they tap primarily this range or not.

HERITABILITY AS A BIOLOGICAL STATISTIC

The concept of heritability was developed to measure response to selection by agriculturalists interested in developing strains of plant and animal species with higher production yields (Hartl & Clark, 1997). Thus, for example, dairy farmers were interested in selecting cows with higher milk production, and corn farmers were interested in selecting corn seeds that would grow into plants with higher oil production. These are situations far removed from even the most controlled psychological laboratory conditions involving human subjects, let alone the community-dwelling twin samples on which most heritability estimates of psychological traits are based. The practical application of heritability in controlled breeding explains many of the concepts used in defining heritability and manipulating its estimates, the limitations of heritability with respect to understanding human psychological traits and the history of resistance to the use of the heritability concept on human psychological traits, as the following sections will explain.

Heritability is defined as the ratio of genetic variance to total trait variance. Immediately, however, things get complicated—especially when breeding and environments cannot be controlled, but even when they can. Not all genetic variance is transmissible from one generation to the next because the genetic material at individual genetic loci (alleles) can have interactive effects on other alleles, either at the same locus on the corresponding chromosome (dominance), or on other alleles at other loci without regard to chromosome (epistasis).[1] Only so-called additive genetic variance not involved in such epistatic or nonadditive effects is transmissible directly to the next generation, as each individual receives one of each chromosome from each parent, and the chromosomes recombine during meiosis. Because agriculturists were interested in response to breeding selection, it was necessary to distinguish the transmissible from the nontransmissible genetic variance. This led to the development of the concepts of narrow and broad heritability, reflecting the transmissible genetic variance and total (nontransmissible as well as transmissible) genetic variance, respectively. When applied to human behaviour, heritability estimates are usually but not always assumed to reflect only transmissible genetic variance. The calculations actually made, however, more accurately reflect at least some of the nontransmissible genetic variance as well. This is yet another reason that the estimates of heritability that are produced in most studies are of much greater importance for their indication of genetic influence in general than for the precise magnitudes.

When (narrow) heritability is, say, 60%, we expect that 60% of the variance will 'breed true'. That is, if we use corn seeds from only those plants that produced corn oil in excess of 1 standard deviation above the mean in the current population, we expect to get plants with a mean at least .6 standard deviation above the original population mean in the next generation, assuming that environmental conditions are the same across the generations. Of course, heritability estimates are never used in this manner in humans because we do not do this kind of selective breeding in humans, despite the frequent estimation of heritability for human traits. This focuses on two important points about heritability estimates.

First, it is not 60% of the phenotype that is passed on to the next generation, or even 60% of genes related to deviation from the original population average in any way that has any meaning for the individual in that next generation. Each selected member of the original generation passes basically one-half of its genes to the next generation, but which half it passes is different for each offspring and typically completely unrelated to the trait of interest. Even individuals with the same parents can receive very different 'packets' of genes related to the trait of interest (it is even

[1]This is the nonadditive genetic variance that was mentioned in the Rushton et al. (2008) paper discussed above.

DOI: 10.1002/per

theoretically possible, though extremely unlikely, that full siblings share none of the genes on which humans can vary), and throughout the population, we should expect to see even greater variability in 'packets' from one generation to the next. In a randomly assorted population, about half the variation is between families and half the variation is within families. This is because each parent and offspring shares basically half their genes by descent and full siblings born of the same parents also share on average half their segregating genes. This implies that, for a trait with 60% heritability and selection at 1 standard deviation above the mean in the original generation, the standard deviation in the next generation will be about .84, or almost as great as the original standard deviation, indicating that even quite substantial heritability does little to create actual similarity among family members. Thus, despite personality's substantial heritability, even monozygotic (MZ) twins are not particularly similar in personality. Their mean personality differences as measured by self-report questionnaires are about 80% of those of randomly paired individuals. That is, on average, randomly paired individuals will differ in personality by a little over 1 standard deviation when the personality measure is at least approximately normally distributed. With typical heritability on the order of 40%, MZ twins will differ on average on the same scale by just less than .9 standard deviation. Dizygotic (DZ) twins, of course, will fall in between (Plomin & DeFries, 1980). Many people find this relative lack of personality similarity surprising in the face of its substantial heritability and the attendant correlation between self-reports and co-twin reports.

Second, whether the second generation mean is actually .6 standard deviations higher when a trait with .60 heritability is selected depends on whether the relevant environment is constant across the generational time span. Many genetic effects are only elicited by specific environmental conditions (for example, genetic influences on smoking and drinking can only be manifested by exposure to cigarettes and alcohol), and changes in the environment between generations can have major effects on the response to selection. Staying with the corn selection example that was previously mentioned, if rainfall was sufficient for the original generation of corn but inadequate for the offspring generation, mean corn oil production could be lower in the second generation than in the original, despite the rather stringent selection process and substantial heritability. Together, these two points make clear that heritability does not indicate the extent to which phenotype, or observed trait level, is genetically determined in any meaningful way. As stated earlier, for traits completely genetically determined, such as the presence of two eyes in humans, heritability is nonexistent because there is no variance. Given the large variations in economic and social conditions across generations, understanding this makes clear that it would be easily possible for the heritability of a psychological trait to be consistent across several generations, yet mean levels could differ substantially between generations. The so-called 'Flynn effect' on intelligence test scores, the well-replicated observation that intelligence test scores have increased continuously over the past hundred years or so, is a demonstration of this.

As a ratio of genetic to total (genetic plus environmental) variance, heritability is dependent on the magnitudes of both the genetic and the environmental variances. With the same magnitude of genetic variance, heritability can be high because environmental variance is relatively low, or low because environmental variance is relatively high. This again points to the importance of the environment in what is measured in heritability. Where relevant environmental circumstances are tightly constrained, heritability will be high even if there is little genetic variance. And where relevant environmental circumstances vary widely, heritability will be low even with substantial genetic variance. Because heritability fluctuates with environmental variation, differences in environmental circumstances that affect its variability, such as the introduction of experimental variation in livestock feed or differences among nations in access to education can have dramatic effects on heritability estimates. For the same reason, heritability can only be informative about reasons for between-group differences when we can be reasonably sure that the environments experienced by the groups are similar with regard to all factors that affect the development of the trait.

A consequence of the many factors that influence the sizes of heritability estimates is that the fact that the probability of detecting genes of measurable effect increases with heritability in simulations of gene-mapping scenarios does not necessarily imply that higher trait heritability increases the probability that there are genes of measurable effect (Visscher, Hill, & Wray, 2008). This helps to place the current absence of findings from genome-wide association studies that can account for the genetic variance of such well-established heritable traits as human height (Weedon & Frayling, 2008) and intelligence (Butcher, Davis, Craig, & Plomin, 2008) and even oil production in corn (Laurie et al., 2004) in perspective. It also suggests greater circumspection on the part of researchers who use estimates of heritability to justify molecular genetic searches for specific genes involved in polygenic traits.

Heritability estimates are based on the assumption that genetic and environmental influences are independent. For agricultural purposes, this assumption is probably pretty reasonable: even when specific individuals do receive different environmental manipulations, crop and animal breeders do not generally allow the individuals to select their own circumstances with respect to the manipulated variables. But humans do select and manipulate their own environmental circumstances, violating the independence assumption with profound implications for heritability estimates. The ability to select environmental circumstances creates correlations and interactions between genetic background and environmental circumstances. Gene-environment correlations exist when there is genetic control of exposure to the environment, or environmental control of genetic expression (Purcell, 2002). The environment that is correlated with the genes can be either shared (when it acts to make family members similar to each other) or nonshared (when it acts to make family members different). Sometimes, gene-environment correlations are also classed as passive, active, or evocative (Scarr & Weinberg, 1983), depending on whether

DOI: 10.1002/per

the individual merely inherits both genes and environmental circumstances that reinforce each other (as when children of antisocial parents both inherit genes for antisocial behaviour and are maltreated), actively selects situations that reinforce genetic inclinations (as when bright children seek out books), or acts in a way that evokes particular kinds of environmental responses (as when people genetically inclined to be social end up with many friends). These two ways of looking at gene-environment correlation do not map directly one onto the other, though some class gene-shared environmental correlation as passive and gene-nonshared environmental correlation as active and/or evocative. This mapping is problematic, as the following example illustrates: one child in a family may have inherited Dad's interest in woodworking. In the process of creating an active and evocative gene-nonshared environmental correlation by teaching this child to build furniture in the lonely garden shed, Dad may also provide a passive nonshared environmental correlation that reinforces the lack of extraversion this child also inherited from Dad. In contrast, this child's sister may have the opposite experience with respect to extraversion through pursuit of a common interest in sports with their much more social Mom.

Gene-environment interactions exist when there is genetically controlled differential sensitivity to the environment or environmental control of genetic response, as when males with the allele producing relatively low levels of monoamine oxidase A are more likely to become antisocial upon experiencing childhood maltreatment than those with the allele producing relatively high levels (Caspi et al., 2002). Human ability to move toward environments that seem 'comfortable' and away from environments that are 'uncomfortable' means that gene-environment correlations and interactions will tend to co-exist (Johnson, 2007).

The presence of gene-environment correlations and interactions introduces systematic biases in heritability estimates made under the independence assumption in twin studies. The nature of these biases can be demonstrated algebraically (Purcell, 2002). As twin designs are most commonly used to make heritability estimates, understanding the biases in these studies is important. The biases occur because the covariance caught up in the correlations and interactions cannot be separately estimated in the classical twin design that only involves MZ and DZ twin pairs. Instead, gene-environment correlations and interactions must be absorbed by one of the three variance components that are estimated: genetic and shared and nonshared environmental. If genetic and shared environmental influences that make family members similar are correlated in ways that affect trait development, estimates of genetic influences are understated. In contrast, if genetic and nonshared environmental influences that make family members different are correlated, estimates of genetic influences are overstated. If genetic and shared environmental influences interact, estimates of genetic influences are also overstated, but if genetic and nonshared environmental influences interact, estimates of genetic influences are understated. Of course, if some combination of correlations and interactions exists, there are many different possibilities for the direction of bias in estimates of genetic influences.

The existence of these kinds of systematic biases in our heritability estimates can have important implications for understanding developmental processes in traits such as intelligence and antisocial behaviours that show changes in heritability with age. For example, the observation that the heritability of IQ increases with age is commonly attributed to the idea that, over time, the genes will 'out' (Bouchard, 2009). The interpretation would be different if it were clear that the increasing heritability estimates with age were due to distortions in the heritability estimates resulting from the presence of correlations between genes and environments that shift over time from gene-shared environmental correlations to gene-nonshared environmental correlations. More specifically, in early childhood, estimates of 35% shared environmental (or even considerably higher, as in Spinath, Ronald, Harlaar, Price, & Plomin, 2003) and 30% genetic influences for IQ are common (Plomin et al., 2007). The shared environmental influences decrease to effectively zero in adult samples, whereas the genetic influences increase to as much as 80% (Deary, Penke, & Johnson, 2010). Based on this, it is not unreasonable to suspect that 30% of the variance in IQ could be caught up in gene-environment correlations that shift from shared to nonshared environmental with age. If this interpretation is correct, it would be more appropriate to view the early shared environment as providing familial support (or lack thereof) for early brain, interest, confidence, motivation and skill development that place the individual increasingly (or decreasingly) in a position to seek out and make use of further opportunities for such development on his/her own.

These kinds of systematic biases in heritability estimates also complicate interpretations of their relative magnitudes. If, for example, one trait shows an estimated heritability of 50% but its development actually involves substantial but unmeasured nonshared environmental correlation, whereas another trait shows an estimated heritability of 30% but involves substantial but unmeasured nonshared environmental interaction without correlation, what are we to infer about the relative importance of genetic variance in the two traits?

Moreover, many constructs also show rather disparate heritability levels depending on precisely how they are measured. For example, Plomin and Foch (1980) used pedometer readings over a 1 week period to measure children's activity levels, generating an estimate of heritability of about 10%, with shared environmental influences of about 89%, likely largely tapping family activities. Wood, Rijsdijk, Saudino, Asherson, and Kuntsi (2008), however, used a composite of parent and teacher reports of hyperactivity and actigraph measures to generate a heritability estimate of 92% for activity level in similarly aged children, likely tapping fidgeting and restlessness during class and other supposedly quiet activities to a much greater degree. Others (e.g., Spinath, Wolf, Angleitner, Borkenau, & Riemann, 2002b [adults]) have obtained more usual moderate estimates of heritability in the .40 range. Such wide discrepancies in heritability estimates may of course arise in part because of differences in the psychometric characteristics of the chosen measurements, but they also may arise because they involve different ways of conceptualising the trait involved or because different measurements involve different kinds of rater bias that are

DOI: 10.1002/per

correlated or interact with the relevant environment in different ways. For example, the heritability of the general factor of personality (Rushton, Bons, & Hur, 2008) may be inflated in many studies based on self-report questionnaires because of the conflation of self-report bias with actual trait level.

The existence of epistatic or nonadditive genetic processes also introduces complications in interpreting the magnitudes of heritability estimates. Completely epistatic processes can generate heritability estimates that appear to reflect largely additive genetic variance (Hill, Goddard, & Visscher, 2008). At the same time, many very specific forms of behaviour such as style of dress and choice of home furnishings show evidence of substantial epistatic or nonadditive genetic variance. Lykken (1982) referred to this as 'emergenesis' or the effects of specific constellations of genes that are identical in MZ twins but are dispersed in any other pairs of relatives. Theoretical models, supported by some empirical evidence, predict that traits that have been under evolutionary selection show increased levels of nonadditive genetic variance (Crnokrak & Roff, 1995; Merilä & Sheldon, 1999; Stirling, Réale, & Roff, 2002). A history of evolutionary selection is not, however, the only possible source of nonadditive genetic variance in a trait. Therefore, the converse, like all converses, that traits that show nonadditive genetic variance have been under directional evolutionary selection, does not necessarily hold. Indeed, the amount of nonadditive genetic variance in a trait might be among the weaker genetic indicators of evolutionary selection (Penke, Denissen, & Miller, 2007b) and needs to be combined with other information on the trait's genetic structure, like the average number, frequency and effect size of associated genetic polymorphisms (Penke, Denissen, & Miller, 2007a). Thus, Rushton et al.'s (2008) observation that there is apparently nonadditive genetic variance in the general factor of personality is consistent with a history of evolutionary selection on that trait, but inferring selection based on this observation alone was certainly premature. Similarly, it is premature to conclude, as did Hill et al. (2008) and many, many others, that estimates of high additive genetic variance for a trait indicate higher chances of success for genome wide and other genetic association studies, which rely on the assumption that genes have primarily additive effects, in revealing the specific polymorphisms underlying the trait.

The presence of measurable heritability in any trait is strong evidence that genes are involved. From a biological perspective, gene involvement is the most important conclusion that can be drawn from heritability estimates. Beyond that, however, heritability alone provides little information about the trait's evolutionary importance, the kinds of genetic and environmental transactions involved in its development, the degree to which we can expect the heritability to be stable across different populations and environmental circumstances, or the degree to which the trait may be responsive to environmental manipulation.

GENETIC CORRELATION

Genetic correlation, or the extent to which the genetic influences on two traits overlap, is often held up as evidence that the same genes directly influence both traits. For example, Kendler and Myers (2010) referred to genetic correlations between measures of personality and incidence of depression as reflecting shared genetic risk factors. This is of course possible, but the inference is premature. Genetic correlations are like phenotypic correlations: Some common factor C (i.e., the same genes) may contribute to both Trait A and Trait B as people often assume, or heritable Trait A may contribute to trait B (whether heritable or not), or heritable B may contribute to A, whether heritable or not. If the unidirectional contributions of one trait to another were to be completely environmental but not universal and to take place, say, in infancy, the genetic influences on one trait would still bleed, over time, into the other and any examination of the two traits after infancy would show genetic correlation. The fact that such situations do exist is, for example, the basis for the research technique of Mendelian randomization (Davey-Smith, 2010).

For a practical example, consider the three most likely explanations for the correlation between physical and cognitive function in old age: (i) Many chronic physical illnesses common in old age are cognitively as well as physically debilitating; (ii) some kind of constitutional integrity may protect both cognitive and physical function in old age; and (iii) good cognitive function throughout life may promote better health habits that protect physical function in old age. Because both lifelong cognitive function and many chronic physical illnesses common in old age are genetically influenced, finding a genetically mediated correlation between them is essentially guaranteed. The presence of genetic influences on the covariance between cognitive and physical functions in old age cannot help to distinguish among these explanations, which are of course not mutually exclusive either (Johnson, Deary, McGue, & Christensen, 2009). This is because, if (i) is the primary explanation for the correlation, the effects of genes involved in any genetically influenced physical illness such as vascular disease that has cognitive symptomatology will 'bleed' into the genetic influences on cognitive function in anyone in the sample who has this disease, contributing to any overall estimate of genetic correlation because they contribute to the overall estimate of heritability of cognitive function in the sample. At the same time, if (ii) is the primary explanation, the genes involved in constitutional integrity will also generate a genetic correlation, this time through actual 'common cause'. But if (iii) is the primary explanation, the genetic influences on cognitive function will 'bleed' into the genetic influences on physical function, contributing to the genetic correlation because they contribute to the overall estimate of heritability of physical function. This kind of confounding is typical of any situation in which there may be reciprocal effects between two traits that develop over time.

Moreover, genetic covariances cannot be used to infer the presence of common molecular polymorphisms (Carey, 1988). Genetic correlations will be high when genes lay down a biological structure and that structure is responsible for individual differences in two variables. But when the same genes are involved in the development of different

DOI: 10.1002/per

biological structures, the degree of genetic correlation that will result is not clear because the genes in common may not be the only genes involved and because the common genes may be very important to one trait but far less so to the other. Similarly, the same genetic loci could be involved in the development of a trait throughout its developmental period, but some could be of major importance during initial stages, whereas others are important at the end of development, creating a negative correlation over time among these loci that leads to a low or even no genetic correlation overall. On the other hand, as noted earlier, if one genetically influenced trait contributes even indirectly and merely phenotypically to the development of another trait, the genetic influences on the first trait will both create genetic correlation between the two traits and increase the apparent genetic variance in the second trait. Finally, nonadditive genetic influences and linkage disequilibrium (the tendency for closely spaced genetic loci to be inherited together) may create genetic correlation between traits even when there are no loci in common at all.

WHERE ARE THE GENES WHEN WE ESTIMATE HERITABILITY?

To date, molecular genetic linkage, association and even genome-wide association studies have yielded much less substantive and replicable results than initially expected, despite large-scale application to many traits of high scientific and commercial interest (McClellan & King, 2010). Height is a good example of the kinds of problems that appear to be involved. Three genome-wide association studies (Gudbjartsson et al., 2008; Lettre et al., 2008; Weedon et al., 2008) covering about 63 000 people revealed some 54 genetic polymorphisms involved in height, but there was almost no overlap across the studies in the polymorphisms that appeared to be involved, and in total, the 54 polymorphisms accounted for less than 5% of the variance in height, and the most recent study (Genetic Investigation of Anthropocentric Traits Consortium, 2010) reported at least 180 common variant associations, explaining only about 10% of the variance. This has left geneticists mystified. Though many remain optimistic that better statistical procedures and more closely spaced genetic markers will reveal the specific polymorphisms that contribute to the heritability we can so readily calculate, some that are suggesting that the currently applied modelling approach does not fully cover the existing theoretical frameworks (e.g., Maher, 2008; Manolio et al., 2009), whereas others are suggesting that we may need to rethink fundamental aspects of our understanding of genetic and evolutionary mechanisms (e.g., Gibson & Dworkin, 2004; Gerhart & Kirschner, 2007; Le Rouzic & Carlborg, 2007; Jablonka & Lamb, 2008; Mill et al., 2008; Eyre-Walker, 2010; Penke, 2010; Johnson, in press; Mitchell, in press). To us, the apparently extremely polymorphic nature of the genetic influences on height in combination with its substantial cohort effects suggest that more favourable environmental conditions (e.g., absence of infection and parasite load, good nutrition) may make possible the expression of many different genes involved in growth (and probably other traits) that would otherwise be silent (unexpressed and/or without phenotypic effects). As we will show, this would be consistent with the new evolutionary and genetic mechanisms under consideration.

Genetic association studies, whether of specific candidate genes or genome wide, rely on the assumption that common traits are associated with common genetic variants, or the correctness of the so-called common disease-common variant hypothesis (Chakravarti, 1999; Iles, 2008). They also rely on the traditional understandings that heredity occurs through the transmission of discrete units of DNA located on chromosomes, that genetic variance is equivalent to variance in DNA sequence and is the consequence solely of many random combinations of pre-existing alleles and new mutations (Jablonka & Lamb, 2008) and that each unit of DNA has a specific function in coding for specific proteins. Thus, by implication, they rely on the assumption that genetic action is not affected by the developmental history of the individual. Linkage studies rely on these assumptions as well, but also on the assumption that the genetic polymorphisms that are relevant to a trait in the family groupings studied are relevant to that trait in more general populations. That is, linkage studies rely by implication on the additional assumption that traits arise through the same genetic mechanism(s) in all families who display them. To date, progress in using these studies to identify genetic polymorphisms has been slow. Increasingly, tests are revealing the inaccuracy of the assumption that gene expression is unaffected by developmental history (Robinson, 2004; Robinson, Grozinger, & Whitfield, 2005; for a review oriented toward psychologists see Roberts & Jackson, 2008) and geneticists are suggesting that it may be more accurate to think in terms of a rare variant-common disease (or trait) hypothesis (Keller & Miller, 2006; Iyengar & Elston, 2007; Penke et al., 2007a; Mitchell, in press), in which many rare genetic variants and mutations contribute to the interruption of networks of genetic systems that control networks of biological systems whose disruption contributes to common diseases (Barabasi, 2007; Goh et al., 2007).

This tangle of underlying assumptions is complicated by our rapidly developing understanding of the presence of epigenetic mechanisms, copy number variants and genetic polymorphisms that appear to exert primarily regulatory effects on other polymorphisms. That is, these gene variants do not code directly for the transcription of amino acids that would otherwise be missing and that in combination build traits, but rather, they are involved in the synthesis of molecules that regulate the rates and magnitudes of transcription of amino acids. Epigenetic mechanisms refer to differences in patterns of gene expression that are not caused by differences in DNA sequence. Some of these differences can be passed from generation to generation despite the fact that, for years, standard biology textbooks have taught us that such Lamarkian transmission is false. For example, feeding a mouse a certain diet can change the coat colour of its offspring and grand-offspring due to DNA methylation (effectively blocking) of expression of the coat

DOI: 10.1002/per

colour gene (Waterland & Jirtle, 2003). Copy number variants refer to often rather long stretches of DNA that can be inserted, deleted, or duplicated in some but not all individuals. Their involvement is increasingly suspected in major behavioural disorders such as schizophrenia and autism (Sebat et al., 2007; International Schizophrenia Consortium, 2008; Kumar et al., 2008; Morrow et al., 2008; Stefansson et al., 2008; Walsh et al., 2008; Xu et al., 2008). Overall, we are increasingly aware that our understanding that each genetic polymorphism is related to a given amino acid and thus trait in a fixed one-to-one pattern is simply inaccurate, and we will need to develop new molecular genetic tools to investigate the molecular genetic processes that we are beginning to understand only now. In other words, heritability estimates help us to recognise that molecular genetic processes are involved, but they do not help us to identify which processes or which genetic polymorphisms.

Some examples of epigenetic phenomena can help to make the complexities of their implications for heritability estimates clear. First, in a much cited study, Weaver et al. (2004) reported that increased licking, grooming and arched-back nursing of pups by rat mothers methylated (blocked genetic expression of) a glucocorticoid receptor in the hippocampus involved in stress response. These effects were reversed with cross-fostering in infancy and external treatment in adulthood but persisted into adulthood without intervention. The study is important for many reasons that go beyond this article, but the issue of relevance to heritability estimates is that the difference in genetic expression did not involve differences in gene sequence. This means that to whatever extent analogous effects on stress response exist in humans, these genetic effects will show up as shared environmental influences in twin and adoption studies as long as individual mothers treat infants consistently in the relevant ways. This is complicated by the fact that genetic expression itself appears to be heritable (York et al., 2005). Thus, the heritability estimates we make may reflect both the effects of genes that contribute directly to the trait of interest and genes that contribute to the regulation of genetic expression, either generally or specifically with respect to the trait of interest. Moreover, if a varying environmental influence directly elicits genetic expression of physiological response, the trait will probably show heritability even if the genes directly involved do not differ among humans because gene actions tend to vary with the rest of the genetic background (Flint & Mackay, 2009).

Environmental stresses may also elicit expression of genetic variation that typically lies dormant. This has been demonstrated in *Drosophila melanogaster* by delivering particular stresses during specific developmental periods. For example, at 21–23 hours of pupal development, 4 hours of heat treatment disrupts the posterior cross-veins in a small percentage of flies (Waddington, 1957). Waddington (1953) demonstrated the heritability of this effect in a series of well-known experiments by crossing the specific animals affected by the heat treatment and subjecting their progeny to it again. The proportion of animals affected by the treatment increased each generation in response to the selection until nearly all the animals showed the effect. But some of the control flies from the same selection lines also showed the effect even without receiving the treatment. Thus, previously silent genetic variation revealed by environmental conditions can be concentrated by selection to the degree that the environmental conditions are no longer necessary for expression. In twin studies and adoption studies, the effects of such changes in gene expression on heritability estimates would depend on the degree to which the environmental conditions creating the changes in genetic expression were uniform in the population.

Though MZ twins share a common genetic background, which includes genetic influence on gene expression, significant variation in gene expression remains. The extent of this variation increases with age (Fraga et al., 2005), suggesting environmental influences. Comparing the similarity of MZ twins reared together and apart across multiple traits suggests, however, that post-natal environmental experiences are not the only sources of these epigenetic differences. This is because MZ twins tend to be similar to the same degree regardless of whether they are reared together or apart (Wong, Gottesman, & Petronis, 2005). This 'similarity of similarity' may be due to active gene-environment correlation. But it may also be due to epigenetic similarities at the time of separation of MZ twin blastomeres that do not affect DZ twin blastomeres (Gärtner & Baunack, 1981; Kaminsky et al., 2009). Such effects would act in general to increase heritability estimates from twin studies and would also be sources of nonadditive genetic variance.

But changes in expression even of relevant genes do not always mean changes in the phenotype of interest, depending on the roles played by the products of those genes in the biochemical processes involved. For example, the renin-angiotensin system is generally acknowledged to be one of the most important means through which blood pressure is genetically controlled. In this system, renin (produced in the kidney) acts on angiotensinogen (AGT, produced in the liver) to generate angiotensin I. This is converted by the enzyme angiotensin-converting-enzyme (ACE) to angiotensin II, which acts through several different receptors to increase blood pressure. Kim et al. (1995) artificially increased the numbers of copies of the genes for AGT and ACE in mice. For the AGT gene, this increased the concentration of AGT in the blood and also increased the blood pressure of the mice. For the ACE gene, however, this increased the concentration of ACE in the blood, but it had no effect on the blood pressure of the mice (Krege et al., 1997). This was because ACE effectively acted only as a gatekeeper in the conversion process from angiotensin I to angiotensin II. Because we have little understanding of how processes such as this may be involved in psychological traits, there is no way to assess their effects on heritability estimates.

Our growing awareness of the existence of these genetic phenomena and new understanding of the kinds of roles they play in some specific situations are very exciting developments. We have to assume that the situations of which we do have some understanding are not unique, and similar processes will be involved in other situations. Some humility is no doubt also warranted: As recently as 10 years ago, we

DOI: 10.1002/per

did not consider these phenomena to be of general importance, and there may be many other phenomena and processes we are overlooking today. These phenomena and processes can have many different kinds of effects on heritability estimates. Given the state of our current knowledge of the sources and natures of epigenetic actions, genetic expression mechanisms and copy number variants, it is not possible to generalise about these effects at present.

CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

Despite a few infamous exceptions, the vast majority of scientists who have estimated heritabilities of human behaviour have never been remotely interested in using them for their original purpose of breeding values. In fact, the now well-established result that all behavioural traits are heritable has been as surprising to these behaviour geneticists as to the behaviourists who believed that people were completely products of their environmental circumstances. Decades ago, when the Zeitgeist locked the door marked 'Genetic Influences on Behaviour,' heritability estimates provided the key that opened that door. Having unlocked this door through ambitious and important behaviour genetic studies, we are moving into a room affording a much better (more accurate) view of the developmental processes contributing to behavioural-related outcomes. Now that we are in that room, however, neither the magnitude nor new reports of the existence of heritability in previously unmeasured psychological or behavioural measures alone tells us much of anything. Most importantly, it is not useful as a criterion to judge the biological importance or even construct validity of a psychological measure. Traits like height, which is on the order of 90% heritable and easily and accurately measurable, are clearly malleable by the environment, or North Koreans would not currently be on average 6 inches shorter than South Koreans (Pak, 2004; Schwekendiek, 2008). Environmental malleability no doubt contributes to the difficulties we are experiencing in identifying the molecular polymorphisms involved in its population variation. Yet, we understand the molecular genetics of phenylketonuria, a disease that shows no heritability at the population level because of its rarity and for which we have an efficient environmental preventive intervention.

Heritability estimates do not open the door to the next room we believe we need to enter, which is marked 'Gene-Environment Transactions' (Johnson, 2007; Johnson, Turkheimer, Gottesman, & Bouchard, 2009). As psychologists, however, we do have keys that will open the closets and cabinets in the new room we have entered that contain the explanations of interplay between genetic and environmental influences. Behaviour genetic studies are optimally positioned to make use of these keys because of their unique ability to control for and chart the presence of genetic influences within environmental circumstances and life outcomes that develop over time. These keys are clearer specification of phenotypes and their developmental trajectories, genetically informative mediator and moderator models that rely primarily on twin samples (particularly in longitudinal studies that allow assessment of developmental trajectories), Mendelian randomization (Davey-Smith, 2010) and the identification and exploration of endophenotypes (Gottesman & Gould, 2003), especially in samples of twins. Embracing more complex behaviour, genetic designs including pairs with different degrees of genetic overlap (such as MZ twins, DZ twins, biological full-siblings, biological half-siblings, adoptive children and parent-child dyads; Coventry & Keller, 2005) is also crucial, because this allows more fine-grained and less biassed descriptions of genetic and environmental phenotypic variation (Keller et al., 2009) and a more accurate assessment of the impact of cultural transmission, nonrandom mating and genome-environment covariation than the classical twin design (Medland & Keller, 2009). Such studies must be very carefully designed, however, as the potential for increased ability to reveal developmental processes is heavily dependent on having similar numbers of pairs of participants with various relationships, good ways to address the lack of independence among relationship pairs when large family groupings participate, and measures of the phenotypes at periods in the lifespan that at least overlap.

Increasingly sophisticated behaviour genetic techniques that can be applied only in genetically informative samples such as twin and adoption studies and that also go well beyond heritability continue to hold some of the most important of the keys we need. There are many current studies making use of them. For example, Finkel, Reynolds, McArdle, Hamagami, and Pedersen (2007) separately examined changes with age in various aspects of cognitive function to investigate whether genetically influenced decreases in processing speed contribute to declines with age among older adults. They concluded that this was the case for memory and spatial ability but not verbal ability. From a different perspective, Irons, McGue, Iacono, and Oetting (2007) used Mendelian randomization to test the gateway hypothesis, or the idea that early alcohol use predisposes adolescents to nonalcohol substance use and antisocial behaviour. In a sample of Asian adoptees, they identified those (30% of the sample) who had inherited a deficiency in the aldehyde dehydrogenase 2 enzyme that is important in alcohol metabolism. Presence of the deficiency makes alcohol consumption unpleasant and causes facial flushing, so most people who have the deficiency avoid alcohol. Despite their much lower rates of use of alcohol, the adolescents with the genetic deficiency had levels of nonalcohol substance use and antisocial behaviour that did not differ from those of the adolescents without the alcohol metabolism enzyme deficiency, thus contradicting the idea that adolescent alcohol use typically serves as a gateway to other undesirable behaviours.

In the same area of research, Perlman, Johnson, and Iacono (2009) examined the potential robustness of P300 amplitude reduction as an endophenotype (Gottesman & Gould, 2003) for alcoholism by testing whether its genetic variance and thus expression was affected by adolescent alcohol use. The idea was that, if adolescent alcohol use was associated with changes in genetic variance in P300, it would be an indication that P300 reduction might be an effect of early alcohol use rather than a marker of genetic

DOI: 10.1002/per

vulnerability to alcoholism. Genetic variance was not affected by adolescent alcohol use. This was the case not only because of lack of statistical significance of effects, an issue of statistical power, but because the parameters that would have been associated with moderating effects were estimated to be effectively 0. The study thus provided further evidence for the viability of P300 reduction as an endophenotype. As a final example, Hicks, South, DiRago, Iacono, and McGue (2009) found that across a broad range of different specific environmental circumstances, including stressful life events and association with delinquent peers, genetic influences on externalising problems among adolescents such as antisocial behaviour and substance abuse tended to be more strongly expressed in poor environmental circumstances, and gene-environment correlations were also lower in these situations. This study is particularly powerful because it revealed similarity in the indicated systems of gene-environment transactions across measures of environmental circumstances and personal characteristics. This similarity suggests that poor environments do something quite fundamental to trigger systemic expression of genetic vulnerabilities to externalising problems (broadly construed), and individuals in better circumstances use those circumstances quite powerfully to minimise expression of the same kinds of genetic vulnerabilities.

Such studies form the bases for the behaviour genetic research of the future that goes way beyond simple heritability estimates, and the tools on which they relied are available for use today. They can help us to progress in the fascinating process of uncovering the genetic underpinnings of human behavioural traits—from merely showing that genes are involved to understanding what they actually do.

ACKNOWLEDGEMENTS

We thank Marion Spengler, Juliana Gottschling and Ralph Dirksen from Saarland University for their work in preparing the scale simulations. Wendy Johnson holds a Research Council of the United Kingdom fellowship. Lars Penke is supported by a grant from the UK Medical Research Council Grant (No. 82800), which is part of the Age UK-funded Disconnected Mind project. The work was partly undertaken within The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative. Funding from the BBSRC, EPSRC, ESRC and MRC is gratefully acknowledged. The work itself received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Agrawal, A., Sartor, C. A., Lynskey, M. T., Grant, J. D., Pergadia, M. L., Grucza R., et al. (2009). Evidence for an interaction between age at first drink and genetic influences on DSM-IV alcohol dependence symptoms. *Alcoholism, Clinical and Experimental Research*, *33*, 2047–2056.

Bäckström, M., Björklund, F., & Larsson, M. R. (2008). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality*, *43*, 335–344.

Barabasi, A.-L. (2007). Network medicine: from obesity to the "diseaseome". *The New England Journal of Medicine*, *357*, 404–407.

Bouchard, T. J., Jr. (2009). Genetic influences on human intelligence (Spearman's *g*): how much? *Annals of Human Biology*, *36*, 1–17.

Butcher, L. M., Davis, O. S., Craig, I. W., & Plomin, R. (2008). Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500 K single nucleotide polymorphism microarrays. *Genes, Brains and Behavior 7*, 435–446.

Carey, G. (1988). Inference about genetic correlations. *Behavior Genetics*, *18*, 329–345.

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

Chakravarti, A. (1999). Population genetics – making sense out of sequence. *Nature Genetics*, *21* (supplement), 56–60.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI). Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Coventry, W. L., & Keller, M. C. (2005). Estimating the extent of parameter bias in the classical twin design: a comparison of parameter estimates from extended twin-family and classical twin designs. *Twin Research and Human Genetics*, *8*, 214 – 223.

Crnokrak, P., & Roff, D. A. (1995). Dominance variance: associations with selection and fitness. *Heredity*, *75*, 530–540.

Davey-Smith, G. (2010). Mendelian randomization for strengthening causal inference in observational studies: Application to gene x environment interactions. *Perspectives on Psychological Science*, *5*, 527–546.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11*, 201–211.

Eyre-Walker, A. (2010). Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proceedings of the National Academy of Sciences*, *107*, 1752–1756.

Flint, F., & Mackay, T. F. C. (2009). Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Research*, *19*, 723–733.

Finkel D., Reynolds, C. A., McArdle, J. J., Hamagami, F. & Pedersen, N. L. (2007). Genetic variance in processing speed drives variation in aging of spatial and memory abilities. *Developmental Psychology*, *45*, 820 – 834.

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences*, *102*, 10604–10609.

Gärtner, K., & Baunack, E. (1981). Is the similarity of monozygotic twins due to genetic factors alone? *Nature*, *292*, 646–647.

Genetic Investigation of Anthropocentric Traits Consortium (2010). Hundreds of variants clustered in genetic loci and biological pathways affect human height. *Nature*, doi:10.1038/nature09410.

Gerhart, J. & Kirschner, M. (2007). The theory of facilitated variation. *Proceedings of the National Academy of Sciences*, *104*, 8582–8589.

Gibson, G. & Dworkin, I. (2004). Uncovering cryptic genetic variation. *Nature Reviews Genetics*, *5*, 681–690.

Goh, K.-I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., & Barabasi, A.-L. (2007). The human disease network. *Proceedings of the National Academy of Sciences*, *104*, 8685–8690.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *The American Journal of Psychiatry*, *160*, 636–645.

Gudbjartsson, D. F., Walters, G. B., Thorleifsson, G., Stefansson, B. V., Halldorsson, P., Zusmanovich, S. P., et al. (2008). Many sequence variants affecting diversity of adult human height. *Nature Genetics*, *40*, 609 – 615.

Hartl, D. L., & Clark, A. G. (1997). *Principles of population genetics*. Sunderland, MA: Sinauer Associates, Inc.

Hicks, B. M., South, S. C., DiRago, A. C., Iacono, W. G., & McGue, M. (2009). Environmental adversity and increasing genetic risk for externalizing disorders. *Archives of General Psychiatry*, *66*, 640–648.

Hill, W. G., Goddard, M. E., & Visscher, P. M. (2008). Data and theory point mainly to additive genetic variance for complex traits. *PLoS Genetics*, *4*, 1–10.

Iles, M. M. (2008). What can genome-wide association studies tell us about the genetics of common disease? *PLoS Genetics*, *4*, e33, 1–9.

International Schizophrenia Consortium (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature*, *455*, 237–241.

Irons, D. E., McGue, M., Iacono, W. G., & Oetting, W. S. (2007). Mendelian randomization: a novel test of the gateway hypothesis and models of gene-environment interplay. *Development and Psychopathology*, *19*, 1181–1195.

Iyengar, S. K., & Elston, R. C. (2007). The genetic basis of complex traits: rare variants of "common gene, common disease"? *Methods in Molecular Biology*, *376*, 71–84.

Jablonka, E., & Lamb, M. J. (2008). Soft inheritance: challenging the modern synthesis. *Genetics and Molecular Biology*, *31*, 389–395.

Johnson, W. (2007). Genetic and environmental influences on behavior: capturing all the interplay. *Psychological Review*, *114*, 423–440.

Johnson, W. (in press). What do genes have to do with cognition? In Kreitler, S. (Ed.), *Cognition and Motivation*. New York: Cambridge University Press.

Johnson, W., Bouchard, T. J., Jr., McGue, M., Segal, N. L., Tellegen, A., Keyes, M., & Gottesman, I. I. (2007). Genetic and environmental influences on the Verbal-Perceptual-Image Rotation (VPR) model of the structure of mental abilities in the Minnesota Study of Twins Reared Apart. *Intelligence*, *35*, 542–562.

Johnson, W., Deary, I. J., & Iacono, W. G. (2009a). Genetic and environmental influences underlying educational attainment. *Intelligence*, *37*, 466–478.

Johnson, W., Deary, I. J., McGue, M., & Christensen, K. (2009b). Genetic and environmental links between cognitive and physical functions in old age. *Journal of Gerontology – Series B*, *64*, 65–72.

Johnson, W., Gangestad, S. W., Segal, N. L., & Bouchard, T. J., Jr. (2008). Heritability of fluctuating asymmetry in a human twin sample. *American Journal of Human Biology*, *20*, 651–658.

Johnson, W., Turkheimer, E., Gottesman, I. I., & Bouchard, T. J., Jr. (2009c). Beyond heritability: twin studies in behavioral research. *Current Directions in Psychological Science*, *18*, 217–220.

Kaminsky, Z. A., Tang, T., Wang, S. C., Plak, C., Oh, G. H. T., Wong, G. H. C., et al. (2009). DNA methylation profiles in monzygotic and dizygotic twins. *Nature Genetics*, *41*, 241–245.

Keller, M. C., Medland, S. E., Duncan, L. E., Hatemi, P. K., Neale, M. C., Maes, H. H. M., Eaves, L. J. (2009). Modeling extended twin family data I: Description of the Cascade Model. *Twin Research and Human Genetics*, *12*, 8–18.

Keller, M. C., & Miller, G. F. (2006). Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *The Behavioral and Brain Sciences*, *29*, 385–452.

Kendler, K. S., & Myers, J. (2010). The genetic and environmental relationship between major depression and the Five-Factor Model of personality. *Psychological Medicine*, *40*, 801–806.

Kim, H., Krege, J. H., Kluckman, K. D., Hagaman, J. R, Hodgin, J. B., Best, C. F., et al. (1995). Genetic control of blood pressure and the angiotensinogen locus. *Proceedings of the National Academy of Sciences*, *92*, 2735–2739.

Krege, J. H., Kim, H. S., Moyer, J. S., Jennette, J. C., Peng, L., Hiller, S. K., & Smithies, O. (1997). Angiotensin-converting-enzyme gene mutations, blood pressures, and cardiovascular homeostasis. *Hypertension*, *29*, 150–157.

Kumar, R. A., KaraMohamed, S., Sudi, J., Conrad, D. F., Brune, C., Badner, J. A., et al. (2008). Recurrent 16p11.2 microdeletions in autism. *Human Molecular Genetics*, *17*, 628–638.

Laurie, C. C., Chasalow, S. D., LeDeaux, J. R., McCarroll, R., Bush, D., Hauge, B., et al. (2004). The genetic architecture of long-term response to selection for oil concentration in the maize kernel. *Genetics*, *168*, 2141–2155.

Le Rouzic, A., & Carlborg, O. (2007). Evolutionary potential of hidden genetic variation. *Trends in Ecology & Evolution*, *23*, 33–37.

Lettre, G., Jackson, A. U., Gieger, C., Schumacher, F. R., Berndt, S. I., Sanna, S., et al. (2008). Identification of ten loci associated with height highlights new biological pathways in human growth. *Nature Genetics*, *40*, 584–591.

Lykken, D. T. (1982). Research with twins – the concept of emergenesis. *Psychophysiology*, *19*, 361–373.

Maher, B. (2008). The case of the missing heritability. *Nature*, *456*, 18–21.

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

McCaffery, J. M., Papandonatos, G. D., Bond, D., Lyons, M. J., & Wing, R. R. (2009). Gene x environment interaction of vigorous exercise and body mass index among male Vietnam-era twins. *American Journal of Clinical Nutrition*, *89*, 1011–1018.

McClellan, J., & King, M. C. (2010). Why it is time to resequence. *Cell*, *142*, 353–355.

McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., et al. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, *78*, 173–186.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A, (2010). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, doi:10.1177/10888868310366253.

Medland, S. E., & Keller, M. C. (2009). Modeling extended twin family data II. Power associated with different family structures. *Twin Research and Human Genetics*, *12*, 19–25.

Merilä, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and non-fitness traits: empirical patterns and development of ideas. *Heredity*, *83*, 103–109.

Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., et al. (2008). Epigenetic profiling reveals DNA-methylation changes associated with major psychosis. *American Journal of Human Genetics*, *82*, 696–711.

Miller, G. F. (2007). Sexual selection for moral virtues. *The Quarterly Review of Biology*, *82*, 97–125.

Mitchell, K. J. (in press). The genetics of neurodevelopmental disease. *Current Opinion in Neurobiology*.

Morrow, E. M., Yoo, S. Y., Flavell, S. W., Kim, T. K., Lin, Y. X., Hill, R. S., et al. (2008). Identifying autism loci and genes by tracing recent shared ancestry. *Science*, *321*, 218–223.

Musek, J. (2007). A general factor of personality: evidence for the big one in the five-factor model. *Journal of Research in Personality*, *41*, 1213–1233.

Neale, M. C., Rushton, J. P., & Fulker, D. W. (1986). Heritability of item responses on the Eysenck Personality Questionnaire. *Personality and Individual Differences*, *7*, 771–779.

Pak, S. (2004). The biological standard of living in the two Koreas. *Economics and Human Biology*, *2*, 511–521.

Penke, L. (2010). Bridging the gap between modern evolutionary psychology and the study of individual differences. In D. M. Buss & P. H. Hawley (Eds.): *The evolution of personality and individual differences* (pp. 243–279). New York: Oxford University Press.

Penke, L., Denissen, J. J. A., & Miller, G. F. (2007a). Evolution, genes, and inter-disciplinary personality research - Response. *European Journal of Personality*, *21*, 639–665.

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

Perlman, G., Johnson, W., & Iacono, W. G. (2009). The heritability of P300 amplitude in 18-year-olds is robust to adolescent alcohol use. *Psychophysiology*, *46*, 962–969.

DOI: 10.1002/per

Plomin, R., & DeFries, J. C. (1980). Genetics and intelligence: recent data. *Intelligence*, *4*, 15–24.
Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2007). *Behavioral Genetics*, 5th Edition. New York: Worth.
Plomin, R., & Foch, T. T. (1980). A twin study of objectively assessed personality in childhood. *Journal of Personality and Social Psychology*, *39*, 680–688.
Purcell, S. (2002). Variance component models for gene-environment interaction in twin analysis. *Twin Research*, *5*, 554–571.
Riemann, R., & Kandler, C. (2010). Construct validation using multitrait-multimethod twin data: the case of a general factor of personality. *European Journal of Personality*, *24*, 258–277.
Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic personality psychology. *Journal of Personality*, *76*, 1523–1544.
Robinson, G. E. (2004). Beyond nature and nurture. *Science*, *304*, 397–399.
Robinson, G. E., Grozinger, C. M., & Whitfield, C. W. (2005). Sociogenomics: Social life in molecular terms. *Nature Reviews Genetics*, *6*, 257–266.
Rushton, J. P. (1985). Differential K theory: The sociobiology of individual and group differences. *Personality and Individual Differences*, *6*, 441–452.
Rushton, J. P., Bons, T. A., & Hur, Y.-M. (2008). The genetics and evolution of the general factor of personality. *Journal of Research in Personality*, *42*, 1173–1185.
Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: the principle of aggregation. *Psychological Bulletin*, *94*, 18–38.
Scarr, S. & Weinberg, R. A. (1983). The Minnesota Adoption Studies – genetic-differences and malleability. *Child Development*, *54*, 260–267.
Schwekendiek, D. (2008). Height and weight differences between North and South Korea. *Journal of Biosocial Sciences*, doi:10.1017/S002193200800299X.
Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., et al. (2007). Strong association of de novo copy number mutations with autism. *Science*, *316*, 445–449.
Spinath, F. M., Angleitner, A., Borkenau, P., Riemann, R., & Wolf, H. (2002a). German Observational Study of Adult Twins: a multimodal investigation of personality, temperament, and cognitive ability. *Twin Research*, *5*, 372–375.
Spinath, F. M., Ronald, A., Harlaar, N., Price, T. & Plomin, R. (2003). Phenotypic *g* early in life: on the etiology of general cognitive ability in a large population sample of twin children aged 2 to 4 years. *Intelligence*, *31*, 195–210.
Spinath, F. M., Wolf, H., Angleitner, A., Borkenau, P., & Riemann, R. (2002b). Genetic and environmental influences on objectively assessed activity level in adults. *Personality and Individual Differences*, *33*, 633–635.
Stefansson, H., Rujescu, D., Cichon, S., Pietilainen, O. P., Ingason, A., Steinberg, S., et al. (2008). Large recurrent microdeletions associated with schizophrenia. *Nature*, *455*, 232–236,
Stirling, D. G., Réale, D., & Roff, D. A. (2002). Selection, structure and heritability of behavior. *Journal of Evolutionary Biology*, *15*, 277–289.
Tellegen, A., & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In G. J. Boyle, G. Matthews, & D. H. Saklofske, *Handbook of Personality Theory and Testing, Vol. II, Personality Measurement and Assessment* (pp. 261–292). Thousand Oaks, CA: Sage.
Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, *9*, 160–164.
Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era – concepts and misconceptions. *Nature Reviews Genetics*, *9*, 255–266.
Waddington, C. H. (1953). Genetic assimilation of an acquired character. *Evolution*, *7*, 118–126.
Waddington, C. H. (1957). *The strategy of the genes*. New York: Macmillan.
Wallace, B., Cesarini, D., Lichtenstein, P., & Johannesson, M. (2007). Heritability of ultimate game responder behavior. *Proceedings of the National Academy of Sciences*, *104*, 15631–15634.
Walsh, T., McClellan, J. M., McCarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M., et al. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science*, *320*, 539–543.
Waterland, R. A. & Jirtle, R. L. (2003). Transposable elements: Targets for early nutrition effects on epigenetic gene regulation. *Molecular and Cellular Biology*, *23*, 5293–5300.
Weaver, I. C. G., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., et al. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, *7*, 847–854.
Weedon, M. N., & Frayling, T. M. (2008). Reaching new heights: insights into the genetics of human stature. *Trends in Genetics*, *24*, 595–603.
Weedon, M. N., Lango, H., Lindgren, C. M., Wallace, C., Evans, D. M., Mangino, M., et al. (2008). Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genetics*, *40*, 575–583.
Wicherts, J. M., & Johnson, W. (2009). Group differences in the heritability of items and test scores. *Proceedings of the Royal Society of London – Series B*, *276*, 2675–2683.
Wong, A. H. C., Gottesman, I. I., & Petronis, A. (2005). Phenotypic differences in genetically identical organisms: the epigenetic perspective. *Human Molecular Genetics*, *14*, R11–R18.
Wood, A. C., Rijsdijk, F., Saudino, K. L., Asherson, P., & Kuntsi, J. (2008). High heritability for a composite index of children's activity level measures. *Behavior Genetics*, *38*, 266–276.
Xu, B., Roos, J. L., Levy, S., Van Rensburg, E. J., Gogos, J. A., & Karayiorgou, M. (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature Genetics*, *40*, 880–885.
York, T. P. F., Miles M., Kendler, K. S., Jackson-Cook, C., Bowman, M. L., & Eaves, L. (2005). Epistatic and environmental control of genome-wide gene expression. *Twin Research*, *8*, 5–15.

DOI: 10.1002/per