For Published Version and To Cite this Book Chapter see:

Beckwith J, Morris-Singer CA. "Missing Heritability: Hidden Environment in Genetic Studies of Human Behavior". In: Fleming J, Levinson S NOMOS LII: Evolution and Morality. New York: NYU Press; Published 2012. p. 122-136.

"Missing Heritability: Hidden Environment in Genetic Studies of Human Behavior"

Jon Beckwith[1] and Corey A. Morris-Singer[2]

1. Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115
   Telephone 617-432-1920 FAX 617-738-7664 e-mail jbeckwith@hms.harvard.edu

2. Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115
   Telephone 617-432-1248 FAX 617-432-1144 e-mail c.morris.singer@gmail.com

Philip Kitcher argues that the sciences of sociobiology and evolutionary psychology are severely limited in their utility for understanding human nature and contributing to the ethical project. He argues for a much broader set of approaches that would necessarily lessen the importance of biology in this effort (see this volume). Nevertheless, some biologists have made the strong case for the relevance of the biological sciences to this project. For instance, E.O. Wilson proposed in his 1975 catalytic book “Sociobiology” that "scientists and humanists should consider together the possibility that the time has come for ethics to be removed temporarily from the hands of the philosophers and biologicized" (Wilson 1975). Like Philip Kitcher, we as geneticists worry that some in the social sciences and humanities, in order to enrich their analyses, are too eagerly grasping at claims emanating from certain research areas in genetics and biology. In our commentary, we focus, in particular, on a complementary area of biology- the search for biological explanations of variation in human behavior and aptitudes. For the most part, these explanations have been provided from the field of human behavior genetics. We do not say that any relevant biology should be ignored, but rather argue that this field has had a sorry history of failed and biased studies that have continued to this day and have tragically influenced social attitudes and policies. We urge a much more critical look at the science and its underlying assumptions, before it is used to support particular ethical, legal and social policy proposals.

An examination of the history of the field of human behavioral genetics reveals a research area that has repeatedly been forced to re-examine its assumptions (Beckwith 1999). Perhaps this is not surprising; one’s social attitudes about such issues as intelligence, criminality, etc. are difficult to separate from the doing of the research itself. Scientists, who were eminent in their own right, such as Sir Francis Galton (who coined the term eugenics in the 19th century),

and Charles Davenport, along with other leading geneticists in the early 20th century, extended their science well beyond what it could really tell them (Kevles 1985). The consequent impact on social policy of the eugenics movement in the 20th century was the sad outcome. But this overextension of the field of genetics and the attendant controversies have continued, most notably in the reports of strong genetic contributions to intelligence, anti-social behavior and gender differences in aptitudes. It is this history that provides a warning that we should look at contemporary research quite closely to see if, in fact, the conclusions drawn are warranted and worthy of taking into consideration for social and ethical political interpretations.

Today, the two main efforts to correlate genetic variation with variations in human behavior are 1)the classical studies of genetically identical twins and 2)research to directly correlate human behavioral variation with alterations of specific sites on the human genome. These two classes of studies are tied together in that the claims for strong heritability of certain behavioral traits derived from twin studies have given confidence to genome researchers that their search for genes will be successful. However, the assumptions underlying the twin studies have been repeatedly and seriously challenged requiring a more tempered view of their conclusions. Furthermore, studies carried out since the human genome was completely sequenced in the year 2000 reveal a much greater complexity than expected in the relationship between genes, environment and human traits, including behavior. This development has led to talk in the scientific literature about the "missing heritability"- the ostensible gap between the heritabilities for traits calculated by those doing twin studies and the actual contribution of genes detected by genome mapping.

Twin Studies and Heritability

The almost universally misunderstood meaning of the scientific term "heritability" represents a starting point in judging how much we might learn from twin studies. Before actually delving into how heritability numbers are arrived at, we will focus on this word. To make this discussion less abstract, we will use scores on an IQ test, presumed by researchers to reflect intelligence, as an example of a trait that is often measured in twin studies.

The heritability of a trait is the amount (or percentage) of the variance of that trait in a particular population that can be attributed to genetic factors (Falconer 1989). However, a heritability number (say 60%) for intelligence in a population says nothing about the degree to which the intelligence of an individual in that population is influenced by genes. It also says nothing about what the heritability of intelligence would be in a different population. Furthermore, since the variance of a trait is measured in a particular environment and at a particular time, the variance of the trait in that population could be dramatically different in a different environment and different time. In fact, a published study of twins, some of whom came from impoverished environments and some from well-off backgrounds, reports that **"**In the families of low SES [socio-economic status], the heritability of IQ is around 10% while in families of high SES it is 78%" (Turkheimer et al. 2003b). This result strikingly illustrates how a trait such as IQ or any other behavioral trait may vary enormously in its heritability according to the environment people inhabit. (In discussing problems with twin studies below, we will raise questions about the significance of high heritability numbers such as 78% for high SES twins reported in this study.)

With all of these limitations to the application of heritability calculations, one might have thought that researchers in the field and the media would have been more careful in presenting

the results to the public. Unfortunately, despite the fact that twin studies only yield information about the particular population under study in a particular environment at a particular time, a very different impression is often communicated to the public. This miscommunication arises, in part, because the term heritability shares the same root as inherited, and is thus conflated with inheritance. For instance, a typical journalistic representation of studies that reports on an IQ and genetics study, will use phrases such as "IQ is shown to be 60% heritable." These representations of the science, more often than not, lead to a description of the trait as being one that is relatively fixed and unchangeable. This conflation may be further exacerbated by statements of the researchers themselves and/or because of the way in which the media presents the results.

Thus, the misuse or misrepresentation of the word heritability and the limitations to its significance when applied to human studies have been controversial from its inception. Leading population geneticists and others have criticized the use or misuse of this approach and its language ever since it began to be applied to human studies. Some examples follow:

> "Apparently it does little good to warn against oversimplifying the idea of heritability. Perhaps this is a basic argument for coining a new word when an idea is to be presented precisely and in a way in which it cannot be misunderstood." (J.L. Lush, 1936, cited in Bell, his emphasis) (Bell 1977)

> "[heritability] is one of those unfortunate short-cuts which have emerged from biometry for lack of a more thorough analysis of the data." (Fisher 1951)

"When a word [heritability] has both technical and folk meanings, it is the responsibility of the specialist to avoid promoting confusion by either using extremely cautious and precise language when using the term or….abandoning the term in favor of one without a widely understood folk meaning." (Stoltenberg 1997)

"The problem with heritability is that it sounds like a property of the feature itself, when in fact it is merely a description of the population in which the trait appears." (Marks 2002)

Twin studies and environment

It is also instructive to consider how the use of the term heritability came to be applied in genetics initially. Agricultural scientists were the first to calculate heritability coefficients. In their studies, they sought to understand the relationship between genes and environment in contributing to traits valuable for agricultural purposes. To do this, they varied the environment of their animal or plant research subjects and varied the genetics by using different genetically inbred strains. Thus, to apply this approach to humans, behavioral geneticists had to argue that they too could control the environment and the genetics of human research subjects. They pointed out that the genetics could be controlled by studying genetically identical twins. There are two major classes of studies utilizing identical twins. First, researchers have compared human behavioral traits in genetically identical twins who were raised by their biological parents with those who were raised by adoptive parents. In order to assess the strength of genetics vs. environment in this case, it would be necessary that the environments of the adoptive families be different from those of the biological families. If that were the case, this class of experiment

would crudely, at least, mimic agricultural experiments where, for instance, plants with the same genes were placed in different environments

The second class of studies examines also identical twins who have been raised by their biological parents, but, in this case, compares them with non-identical twins also raised with their biological parents. Fraternal twins share on average 50% of their genes while identical twins share all of their genes. The difference of the degree to which the identical and fraternal twins show expression of a particular trait allows the calculation of heritability. It is argued that that since twins grow up together at the same time in the history of their family and the history of their culture, a pair of genetically identical twins (100% shared genes) can be considered to have grown up sharing the same environment to the same degree as do a pair of non-identical twins (50% genetics shared). This would not be true of ordinary (non-twin) siblings born at different times in the history of their family. This assumption, known as the Equal Environment Assumption (EAA), it is argued, allows researchers to control the environment.

To many observers, the simplification of concepts of environment in these studies is problematic, raises questions about the studies' conclusions and, thus, limits the application of those conclusions to questions of social or moral import. (For a more detailed discussion of the issues that follow and of the meaning of heritability, we highly recommend either one of two similar impressively lucid essays by the philosopher Elliott Sober (Sober 2000; Sober 2001).)

The failings in the treatment of environment can be attributed both to the simplifying of the complex nature of human social interactions and to the limitations in our knowledge of the environmental factors that contribute to behavioral traits. In studies comparing adopted identical twins with those twins raised by their biological family, the nature of the adoption process itself can seriously interfere with this apparently well thought-out experiment. Twins who are adopted,

either via adoption agencies or through the efforts of the biological parents themselves, are rarely placed in homes with environments very different from those of the biological parents. This criticism, which suggests that environment and genetic contributions to behaviors cannot be distinguished by these studies, was made most convincingly in the 1970s (Kamin 1974). As a result, behavioral geneticists made an effort to directly assess the "environment" of adoptive homes. When studying the heritability of IQ, for example, they counted the number of books in the homes and noted other features that would indicate or not an intellectually rich environment (Bouchard et al. 1990; Rowe 1994). While certainly the environmental features noted are ones that seem potentially relevant to IQ, simply counting the number of books, etc. does not give a direct sense of that environment. Paradoxically, although the question presumably being asked in such studies is what are the sources of higher IQ scores, the researchers are assuming that they already know the "trait-relevant" environmental factors for which to test.

Strong criticisms have also been raised of the Equal Environment Assumption, a central assumption of the twin studies that compare fraternal and identical twins raised in their biological families (Joseph 1998; Alford et al. 2008; Beckwith and Morris 2008). There are several reasons for suspecting that, in contrast to this assumption, the environments experienced by fraternal twins are different from those experienced by identical twins. First, the physical identity of identical twins can influence the way in which they are treated by their parents and responded to by the outside world. How people respond to an individual's physical features, for example short stature, physical beauty, obesity, skin color, etc, can vary with cultural attitudes. These responses, which vary from positive to negative, can, in turn, influence the behavior of identical twins in similar ways that would not be the case for fraternal twins. In the face of these criticisms, behavior genetic researchers have attempted to assess the validity of the EEA

assumption. If the EEA for the trait under study is incorrect, and identical twins indeed experience more similar environments than fraternal twins, then the calculation of heritability coefficients in such studies would be artificially high, since different concordance of IQ scores in the two types of twins could be due to either genetic or environmental differences.

We have recently reviewed studies that attempt to test the EEA, and concluded that they do not establish the validity of the EEA (Beckwith and Morris 2008). We point out how difficult a task it is to tease out the complex environmental effects that can occur in such studies. The studies published were done with small sample sizes with low statistical power, used problematic retrospective interviews with parents and children in evaluating the childhood environments, and make assumptions about what trait-relevant environmental differences to measure (see the discussion above). Furthermore, several of the studies contain evidence that the trait-relevant assumptions have been violated.

How Heritability Reports Can Mask Environmental Influences

Beyond the potential confounding environmental influences characteristic of these types of twin studies, we want to point out a surprising way in which researchers can turn what some might consider an environmental factor into a genetic one and thus calculate "inflated" heritabilies. We have talked earlier about how identical features of twins can result in identical interactions with their parents and identical experiences with people in society at large. Such effects that may influence their behavior would be thought of by many as a gene-environment interaction. Let's consider an extreme example. In the United States, people with genes that determine dark skin color are liable to meet prejudice to such an extent that their behavior,

education, blood pressure, etc. is significantly influenced. It is true that the genes are responsible for causing the skin color, but it is our culture and prejudices at this time in history that cause discriminatory behavior and its impact. In this case, the genes and the environment are interacting, but, in another less-prejudicial environment this effect could disappear.

However, behavior geneticists often turn this argument around and include such gene-environment interactions in the genetic component of the calculations. Twin researcher, Thomas Bouchard among others has presented this rational, stating that "Identical twins elicit, choose, seek and create environments very similar and, accordingly, the impact of these environments is considered a genetic influence." He and others appear to agree that the environment of identical twins may be more similar than that of fraternal twins, but argue that this similarity is caused by the genes themselves (Bouchard et al. 1990). Their argument may have some validity in a technical sense, but it hides from view the important impact of environmental context on heritability calculations, thus, leaving to the reader a sense that the studies have shown a strong direct and more immutable influence of genes on behavior.

A Shift in Human Behavior Genetics

We have argued that the field of behavior genetics, through much of its history has treated environment in a simplistic manner in its efforts to define heritability of traits. This oversimplification was surely influenced by the success of methodological reductionism in other fields of science, particularly molecular biology, which behavior genetics developed alongside in the 20th century. Within the last ten years, however, there have been indications of a significant shift in the emphasis that behavior geneticists have placed on the role of the environment. We

have already mentioned the recent paper on the enormous shifts in heritability of IQ with changes in environment which has challenged basic assumptions (Turkheimer et al. 2003a). Elsewhere, one of the authors of this latter paper has pointed quite clearly to the difficulties confronting his field in defining trait-relevant environments that influence the development of specific behaviors or aptitudes (Turkheimer 2006). Furthermore, leading behavior geneticist James Flynn, has roiled the field with his studies of changes in average IQ of populations over the several decade period from 1950 to 2000 (Flynn 2007). This was to test the underlying assumption in the field that IQ scores, deemed to have a strong genetic component, should remain fixed and not change with time in a population. What he found was a steady increase over the years and, in addition, that the gap between black and white IQ's in the U.S. had steadily narrowed. Flynn suggests various cultural and educational changes that are responsible for increased IQ scores.

A very different class of behavior genetic studies that indicate a shift in direction for this field has emerged. These studies take as their starting point gene-environment interactions. A requirement for such studies is there be an already identified gene that, because of its function, is a candidate for one that affects human behavior. Caspi et al. reported that people carrying a variant of a gene (MAOA) that should control serotonin levels in the brain, conferred susceptibility to anti-social behavior (Caspi et al. 2002). However, that susceptibility was only observed with those individuals who carried the variant copy of the gene and had been subjected to child abuse. The authors conclude that taking into account potential environmental factors on a behavior can actually help researchers to detect genetic contributions to that behavior. This gene-environment approach may reflect an interesting trend. Nevertheless, as an indication of the difficulty of coming up with meaningful results, subsequent attempts to replicate this study

have yielded mixed results. The mixed results themselves may indicate complex environmental or genetic differences between the populations that were used in each study (Kim-Cohen et al. 2006; Morris et al. 2007).

One last example illustrates the problems in treating environment too lightly when seeking biological contributions to human behavior or aptitudes. In a 1980 study on math performance of boys and girls, researchers reported in Science Magazine that eighth grade boys outnumbered girls 13:1 at the highest levels of scores on the SAT math test. In their article, the authors of the report implied that scientists should pay more attention to biological differences in explaining the gap between boys and girls in math achievement (Benbow and Stanley 1980). Unfortunately for this proposal, the significance of the finding has almost disappeared in the 30 years after the study as the boy/girl ratio of 13:1 had fallen to 2.8:1 in 2005 (Monastersky 2005). Furthermore, the percentage of women who obtained PhD's in mathematics had increased from 6 % in 1970 to 30% in 2003 (Ceci and Williams 2010b). These changes occurred (obviously) without any genetic engineering or time for evolution. These dramatic changes may have been due to the success of the women's movement since the 1970's in instituting numerous educational programs to enhance the interest and performance of girls in math (Beckwith 1983; Ceci and Williams 2010a). While this example does not represent studies that used twins as their subject, it is relevant in that it shows the danger of treating environment in a simplistic way.

The Human Genome and Human Behavior Genetics

With the initiation of the Human Genome Project in the late 1980's, hopes surged for identifying genes involved in the human behaviors and aptitudes studied by behavior geneticists.

These hopes were bolstered by initial rapid identification of genes in which mutations were highly correlated with specific, but rare diseases. For example, mutations were characterized in a single gene that caused Huntington's chorea, in another gene that caused Muscular Dystrophy and in another that caused Gaucher's Disease. However, in recent years it has become clear that for many common human diseases, single mutations in single genes will not provide a causative explanation. Many, perhaps most common human diseases are much more complex in terms of genetic, genetic/environment interaction and environmental sources than was thought. This complexity is illustrated by recent investigations into the genetic contributions to Crohn's disease. Using an approach called Genome Wide Association Studies (GWAS), researchers found 30 different genetic locations in human chromosomes where mutations made small contributions to Crohn's (Barrett et al. 2008). Further, from statistical analyses, they predicted that mutations at approximately 100 such sites total will eventually be found to contribute to Crohn's Disease. That is, perhaps mutations in 100 genes, each mutation having a small effect, contribute to this disease.

It has been significantly more difficult to identify genetic variants associated with behavioral traits or mental illnesses such as schizophrenia. And, for IQ, no replicable finding has been made of a gene that correlates with IQ scores. Researchers in the field have concluded that many human trait variations may be due to small effects of a large number of genes, interaction of those genes with the environment or may be largely environmentally caused.

Missing Heritability

In a recent article, researchers have used the powerful tool of Genome Wide Association Studies to identify genes related to a wide variety of personality traits (Verweij et al. 2010). However, they report that "No genetic variants that significantly contribute to personality traits were identified while our sample provides over 90% power to detect variants that explain only 1% of the trait variance. This indicates that common genetic variants of this size or greater do not contribute to personality trait variation…." The authors refer to the problem of "missing heritability" discussed in our Introduction and predict that "Newer technologies…. and novel statistical approaches combined with larger samples and meta-analyses…." will overcome this problem. The appeal to newer genetic technology and larger sample sizes has become a common refrain in the search for gene-trait correlations; however, what's not commonly acknowledged is that these studies often reveal diminishingly small effects on the trait under study. Furthermore, nowhere do these authors mention the possibility that the heritability calculations may be biased by deficits in the consideration of environment in twin studies. Instead, we would suggest that a major reason for the problem of finding proposed genes involved in human behavior may be a simple one. The researchers who have used the twin study approach with the goal of determining whether there are strong genetic influences on such traits as intelligence have simplified concepts of environment, minimizing its definition and its effects, in order to make their studies feasible. Or, their methodology obscures ways in which environmental influences of human behavior are hidden within their heritability estimates of genetic effects. The result is that they either overestimate the heritabilty for the traits they study or include environmental influences in their calculations of heritability. There may be no missing heritability, but there is certainly missing or hidden environment.

Rather than being at a time where sufficient information has now accumulated from behavior genetic approaches to start employing them for an understanding of human social behavior, we are instead at the beginning of a new era of exploration that recognizes complexity. New approaches are being taken to incorporate environment in a more significant way than in the past in trying to understand the role of the environment in the development of human traits. To those who would like to use the approaches or some of the results of behavior genetics, we cannot emphasize enough how important it is to look with a critical eye at the kinds of assumptions that underly research in this field.

References

1. Alford, J. R., C. L. Funk, and J. R. Hibbing. 2008. "Twin Studies, Molecular Genetics, Politics and Tolerance." *Perspectives on Politics* 6:793-797.
2. Barrett, J. C. et al. 2008. "Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease." *Nat.Genet.* 40(8):955-962.
3. Beckwith, J. 1983. "Gender and math performance: does biology have implications for educational policy?" *Journal of Education (Boston University)* 165:158-174.
4. Beckwith, J. 1999. "Simplicity and complexity: Is IQ ready for genetics?" *Current Psychology of Cognition* 18:161-169.
5. Beckwith, J., and C. A. Morris. 2008. "Twin studies of political behavior: untenable assumptions?" *Perspectives on Politics* 6:785-791.
6. Bell, A. E. 1977. "Heritability in retrospect." *J.Hered.* 68(5):297-300.
7. Benbow, C., and J. Stanley. 1980. "Sex differences in mathematical ability: Fact or artifact?" *Science* 210:1262-1264.
8. Bouchard, T. J. Jr. et al. 1990. "Sources of human psychological differences: the Minnesota study of twins reared apart." *Science* 250:223-228.

9. Caspi, A. et al. 2002. "Role of genotype in the cycle of violence in maltreated children." *Science* 297(5582):851-854.

10. Ceci, S. J., and W. M. Williams. 2010b. "Sex Differences in Math-Intensive Fields." *Curr.Dir.Psychol.Sci.* 19(5):275-279.

11. Ceci, S. J., and W. M. Williams. 2010a. "Sex Differences in Math-Intensive Fields." *Curr.Dir.Psychol.Sci.* 19(5):275-279.

12. Falconer, D. S. 1989. *Introduction to Quantitative Genetics.* 3 ed. London: Longman Scientific and Technical.

13. Fisher, R. A. 1951. "Limits to Intensive Production in Animals." *J.Hered.*(4):217-218.

14. Flynn, J. R. 2007. *What is Intelligence.* Cambridge: Cambridge University Press.

15. Joseph, J. 1998. "The equal environment assumption of the classical twin method: a critical analysis." *The Journal of Mind and Behavior* 19:325-358.

16. Kamin, L. 1974. *The Science and Politics of I.Q.* Potomac, Md.: Earlbaum Associates.

17. Kevles, D. 1985. *In the Name of Eugenics: Genetics and the Uses of Human Heredity.* Berkeley: University of California Press.

18. Kim-Cohen, J. et al. 2006. "MAOA, maltreatment, and gene-environment interaction predicting children's mental health: new evidence and a meta-analysis." *Mol.Psychiatry* 11(10):903-913.

19. Marks, J. 2002. *What it means to be 98% chimpanzee: apes people and their genes.* Berkeley: University of California Press.

20. Monastersky, R. 2005. "Primed for Numbers." *Chronicle of Higher Education* March 4:A1.

21. Morris, C. et al. 2007. "Deconstructing Violence." *GeneWatch* 20(2):3-10.

22. Rowe, D. C. 1994. *The Limits of Family Influence: Genes, Experience and Behavior.* New York: The Guilford Press.

23. Sober, E. 2000. "The meaning of genetic causation." Pp. 347-370 in *From Chance to Choice: Genetics and Justice.* Edited by A. Buchanan, D. W. Brock, N. Daniels, and D. Wikler. Cambridge: Cambridge University Press.

24. Sober, E. 2001. "Separating Nature and Nurture." Pp. 47-78 in *Genetics and Criminal Behavior.* Edited by D. Wasserman and R. S. Wachbroit. Cambridge: Cambridge University Press.

25. Stoltenberg, S. F. 1997. "Coming to terms with heritability." *Genetica* 99:89-96.

26. Turkheimer, E. 2006. "Mobiles: A Gloomy View of Research into Complex Human Traits." Pp. 100-108 in *Wrestling with Behavioral Genetics: Science, Ethics and Public Conversation.* Edited by E. Parens, A. R. Chapman, and N. Press. Baltimore: Johns Hopkins University Press.

27. Turkheimer, E. et al. 2003b. "Socioeconomic status modifies heritability of IQ in young children." *Psychol.Sci.* 14(6):623-628.

28. Turkheimer, E. et al. 2003a. "Socioeconomic status modifies heritability of IQ in young children." *Psychol.Sci.* 14(6):623-628.

29. Verweij, K. J. et al. 2010. "A genome-wide association study of Cloninger's temperament scales: implications for the evolutionary genetics of personality." *Biol.Psychol.* 85(2):306-317.

30. Wilson, E. O. 1975. *Sociobiology:The New Synthesis.* Cambridge: Harvard University Press.