University of Massachusetts Boston
ScholarWorks at UMass Boston

Working Papers on Science in a Changing World | Critical and Creative Thinking Program

Spring 3-22-2014

Five Fundamental Gaps In Nature-Nurture Science

Peter J. Taylor
University of Massachusetts Boston, peter.taylor@umb.edu

Follow this and additional works at: https://scholarworks.umb.edu/cct_sicw

 Part of the Biostatistics Commons, and the Genetics Commons

Recommended Citation

Taylor, Peter J., "Five Fundamental Gaps In Nature-Nurture Science" (2014). *Working Papers on Science in a Changing World*. 3.
https://scholarworks.umb.edu/cct_sicw/3

This Article is brought to you for free and open access by the Critical and Creative Thinking Program at ScholarWorks at UMass Boston. It has been accepted for inclusion in Working Papers on Science in a Changing World by an authorized administrator of ScholarWorks at UMass Boston. For more information, please contact scholarworks@umb.edu.

WORKING PAPER on Science in a Changing World

Paper # 3-2014

Five Fundamental Gaps in Nature-Nurture Science

PETER J. TAYLOR

GRADUATE PROGRAM IN CRITICAL AND CREATIVE THINKING
University of Massachusetts, Boston, MA 02125, USA
http://scholarworks.umb.edu/cct_sicw/3

Five Fundamental Gaps In Nature-Nurture Science

Peter J. Taylor
Science in a Changing World graduate track
University of Massachusetts, Boston, MA 02125, USA
peter.taylor@umb.edu

Abstract

Difficulties identifying causally relevant genetic variants underlying patterns of human variation have been given competing interpretations. The debate is illuminated in this article by drawing attention to the issue of *underlying heterogeneity*—the possibility that genetic and environmental factors or entities underlying a trait are heterogeneous—as well as four other fundamental gaps in the methods and interpretation of classical quantitative genetics: "Genetic" and "environmental" fractions of variation in traits are distinct from measurable genetic and environmental factors underlying the traits' development; Standard formulas for partitioning variation in human traits are unreliable; Methods for translation from fractions of variation to measurable factors are limited; and Variation within groups is different from variation between averages for separate groups. Given these five gaps in the estimation and interpretation of components of variance, high heritability values for traits are not a reliable basis for choosing which traits to investigate by molecular techniques; this helps explain why identification of causally relevant genetic variants has not produced the results and insights hoped for.

Genome-Wide Association studies have identified variants at large numbers of genetic loci that confer statistically significant changes in traits, including increases in risk for diseases such as diabetes, heart disease, and cancers in defined populations (Khoury et al. 2007). A consensus has emerged that most medically significant traits are associated with many genes of quite small effect (McCarthy et al. 2008). The detection and identification of variants is further complicated by genetic heterogeneity in its various forms (e.g., mutations in a gene may occur at a variety of points in the gene, the clinical expression of such mutations can vary significantly, and different genetic variants may be expressed as the same clinical entity). The implications to be drawn from difficulties identifying causally relevant genetic variants have been the subject of active debate (Couzin-Frankel 2010). In particular, can variants associated with a significant but very small effect still lead researchers to biologically revealing pathways? Or, is it the case that, taking genetic heterogeneity into account, future advances will come from finding rare alleles having a strong effect (McClellan and King 2010)?

The debate is illuminated in this article by returning to the classical quantitative genetic partitioning of variation in a given trait in some defined population. The conventional wisdom is that "[r]esearch into the genetics of complex traits has moved from the estimation of genetic variance in populations [i.e., classical quantitative genetics] to the detection and identification [made possible by new tools of molecular biology] of variants that are associated with or directly cause variation" (Visscher et al. 2007). This move, however, rests on taking high values of a classical measure, heritability, to indicate a strong genetic contribution for a trait, such as incidence of heart disease, which makes the trait "a potentially worthwhile candidate for molecular research" that might identify the specific genetic factors involved (Nuffield Council on Bioethics 2002, chapter 11). In light of this continuing—and foundational—role for classical quantitative genetics, five fundamental *gaps* in the field's methods and interpretations (Taylor 2010) are discussed. Each gap is presented in a capsule summary that is then elaborated. Researchers and commentators concerned with the difficulties identifying causally relevant variants, or with nature-nurture issues more generally, as well as teachers of the next generation of researchers would benefit from acknowledging and consistently sustaining appropriate responses to *all* of these gaps.

Underlying heterogeneity

When a trait is observed to be similar within a group of individual and different among groups, there may be similar conjunctions of genetic and environmental factors (or, in epidemiology, risk or protective factors) involved in producing the trait, but this need not be the case. That is the first gap. The appropriate response is to allow for the possibility of heterogeneity of factors underlying any given trait.

Consider claims that some human trait, say, IQ test score at age 18, show high heritability (Neisser et al. 1996). These claims can be derived from analysis of data from relatives. For example, the similarity of pairs of monozygotic twins (which share all their genes) can be compared with the similarity of pairs of dizygotic twins (which do not share all their genes). The more that the former quantity exceeds the latter, the higher is the trait's heritability (assuming for purposes of discussion that monozygotic twins are not treated more similarly than are dizygotic twins). Researchers and commentators often describe such comparisons as showing how much a trait is "heritable" or "genetic." However, no genes or measurable *genetic factors* (a generic term used in this article to denote entities such as alleles, tandem repeats, chromosomal inversions, etc.) are examined in deriving heritability estimates (or estimates of other fractions of trait variation in classical quantitative genetics). Nor, as some prominent geneticists have noted (e.g., Rutter 2002, 4), does the method of analysis suggest where to look for them. Moreover, even if the similarity among twins or a set of close relatives is associated with similarity of (yet-to-be-identified) genetic factors, *the factors may not be the same from one set of relatives to the next, or from one environment to the next*. In other words, the underlying factors may be *heterogeneous*. It could be that pairs of alleles, say, AAbbcbDDee, subject to a sequence of environmental factors, say, FghiJ, during the development of the organism are associated, all other things being equal, with the same outcomes as alleles aabbCCDDEE subject to a sequence of environmental factors FgHiJ (Fig. 1). The gap between homogeneous and heterogeneous genetic and environmental factors influencing the development of a trait has yet to be recognized as a significant methodological concern by quantitative geneticists or by critical commentators on heritability research (e.g., Downes 2004 and references therein).

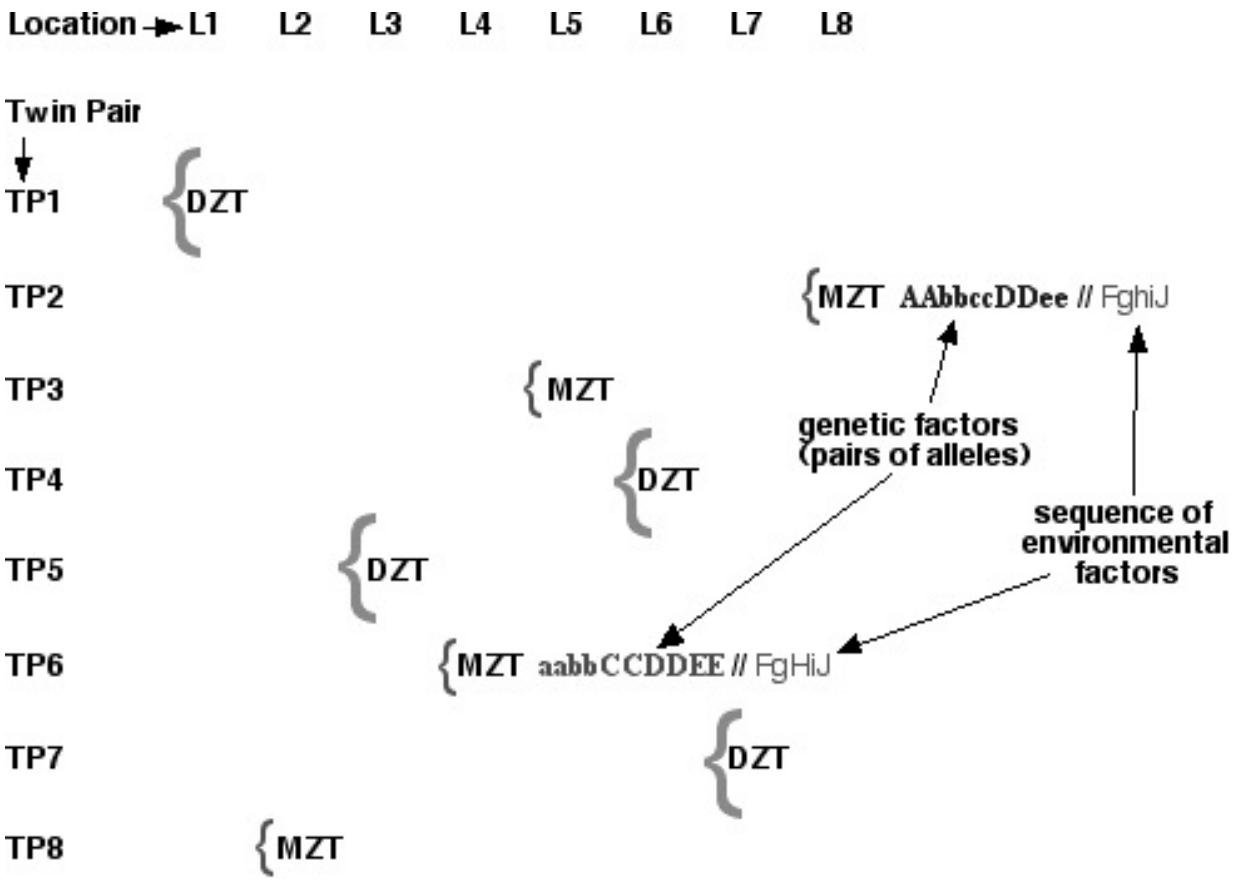

Figure 1. Factors underlying a trait may be heterogeneous even when identical or monozygotic twins (MZT) are more similar than fraternal or dizygotic twins (DZT). The greater similarity is indicated by the *smaller* size of the curly brackets. The underlying factors for two MZ pairs are indicated by upper and lower case letters for pairs of alleles (A-E) and environmental factors to which they are subject (F-J).

Of course, it is not the case that underlying factors are always heterogeneous. Some traits are largely determined by the genes at a single locus more or less independently of the individuals' upbringing—so called high-penetrance major genes (e.g., presence of extra digits or polydactyly). The detection of such traits can, however, be made through examination of family trees; quantitative genetics and heritability estimation need not be involved. If such traits are put aside, there are no obvious grounds to rule out the possibility of heterogeneity in the measurable genetic and environmental factors that underlie patterns in quantitative and other complex traits, such as crop yield, height, human IQ test scores, susceptibility to heart disease,

personality type, and so on. Moreover, because underlying heterogeneity encompasses both environmental and genetic factors, researchers face an even greater challenge than indicated when genomics researchers have responded to difficulties in identifying causally relevant factors by emphasizing *genetic* heterogeneity (McClellan and King 2010).

The appropriate response to the first gap is to acknowledge the possibility of underlying heterogeneity and it implications for quantitative genetics (Taylor 2010). Doing so could, for example, lead researchers to seek to identify the specific genetic and environmental factors without reference to the trait's heritability or the other fractions of the total variance. It could prepare them to expect fruitless molecular investigations on route to finding the special high heritability traits for which the underlying factors are not heterogeneous. It could lead them to restrict attention to variation within a set of relatives. This last path makes sense because, even if the underlying factors are not known, high heritability still means that if one twin develops a trait (e.g., type 1 diabetes) the other twin is more likely to as well. The second twin might be advised to take measures to reduce the health impact if and when the disease started to appear for that twin. However, notice that this path assumes that the timing of getting the condition differs from the first twin to the second; the factors influencing the timing could also be heterogeneous. Further implications of the gap between assuming heterogeneity versus homogeneity of underlying factors will be discussed after presentation of the other four gaps.

Statistical patterns in traits are distinct from measurable factors

The second gap lies between, on one hand, quantitative genetics that deals with the statistical analysis of measurements on a trait for a sample of related and unrelated individuals in a range of situations and, on the other hand, the investigation of measurable genetic and environmental factors influencing the processes through which the trait develops in different individuals. These inquiries are conceptually distinct. This gap needs to be highlighted, not downplayed or obscured.

Conceptual clarity and terminological adjustments can help highlight this gap. As a starting point, the potential for confusion in the varying uses of the term "genetic"

diminishes if genetic is reserved as an adjective in reference to factors that are transmitted from parents to offspring and whose presence can, in principle, be observed. In a similar spirit, “environmental” can be taken to refer only to measurable factors, which can range widely, say, from average daily intake of calories to degree of maltreatment that a person experienced as a child. Potential for confusion associated with the commonly used nouns “genotype” and “environment" can also be reduced. These terms obscure the second gap by suggesting, without warrant, that the quantities estimated through analysis of data about observed traits have a relationship with measurable genetic and environmental factors influencing the development of the trait. Suitable substitutes are provided by the agricultural terms *variety* and *location*. A variety can be thought of simply as a group of individuals whose relatedness by genealogy can be characterized, such as offspring of a given pair of parents, or a group of individuals whose mix of genetic factors can be replicated, as in an open pollinated plant variety. A location is the situation or place in which the variety is raised, such as a family of humans or a plot at an agricultural research station. The use of the terms variety and location does not assume that researchers can specify the genetic or environmental factors that influence the trait in the various variety-location combinations.

Clarity about the second gap allows sound interpretation of the classical quantitative genetic analysis of variation among related and unrelated individuals for a given trait, which centers on partitioning the variation into fractions according to simple additive models (i.e., in statistical terms, undertaking an Analysis of Variance). In these models the value of the trait for a given individual is a sum of separate elements, including ones associated with the individual’s variety and location as well as variety-location combinations or *interaction* and noise or unsystematic influences (e.g., measurement error). (“Element” will be used here in place of the technical term “effect,” whose causal connotations are unwarranted and can confuse the discussion.) The overall variation in the trait becomes a sum of the variances of the elements in the additive model. Figure 2 depicts the partitioning of variation for an agricultural evaluation trial, where it is possible to raise or grow a set of animal or plant varieties in each of a set of locations and to raise replicates for each variety-location combination.

(As will become evident, it is often helpful to consider agricultural studies and to contrast what can be known through those studies with what can be known through analyses of data from humans.)

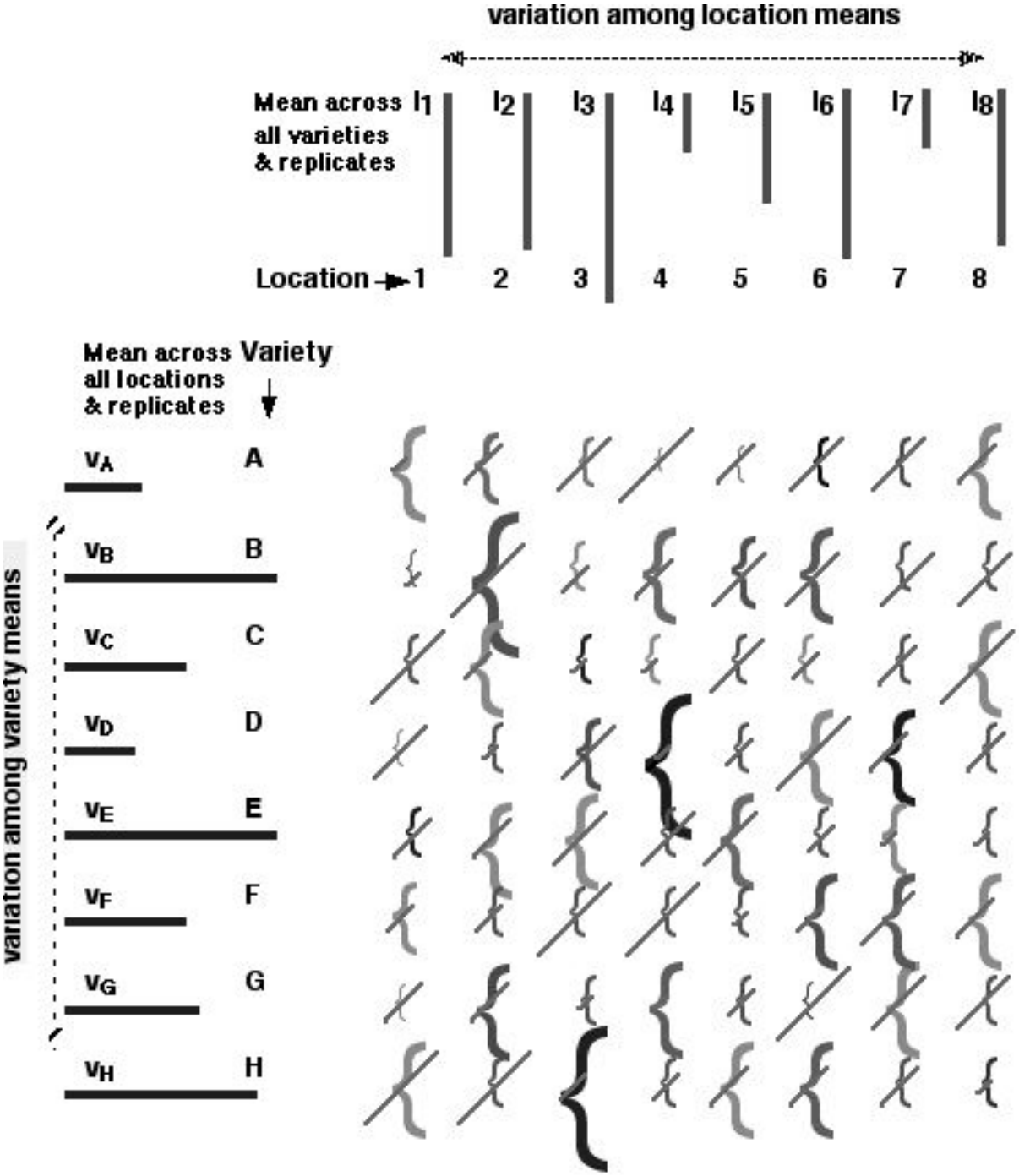

Figure 2. Partitioning of variation in the ideal agricultural evaluation trial where each of a set of varieties is raised in each of a set of locations, and there are two or more replicates in each variety-location combination. The bars next to the varieties and

locations and the diagonal bars for each variety-location combination indicate the average value of the trait for, respectively, each variety, each location, and each variety-location combination (after allowing for the variety and location averages). These are the *elements* referred to in the text. The variation between replicates within variety-location combinations is indicated by the size of the curly brackets. The brackets are given a non-systematic shading to denote that the variation between replicates is not correlated from one variety-location combination to another. Note the contrast between the agricultural evaluation trial and Figure 1, in which the replicates of any variety—twin pairs—are raised in only one location—household—per variety.

Heritability is often described as the fraction of variation in a trait associated with "genetic differences" or "genetic variance," but the original quantitative genetic concept does not concern variation among the *genes* possessed by the individuals. The descriptions in quotes are loose expressions for the variance of the variety elements, where each variety's element is related to the average or mean of the *trait* for the variety across all locations and replicates minus the overall mean. Similarly, the "environmental" variance, sometimes labeled "*shared* environmental" variance, is the variance of the location elements, not variation in environmental factors experienced by the individuals.

The variance of the "variety-location interaction" elements is distinct from the use of the term "gene-environment interaction" for situations in which "gene" denotes a value of a measured genetic factor, the "environment" denotes a value of a measured environmental factor, and an interaction means that the quantitative relation between the trait and one of the factors varies according to the measured value of the other factor (e.g., Moffitt et al. 2005). For the partitioning of trait variance, in contrast, a high degree of variety-location interaction means that the responses of the observed varieties across the range of the observed locations do not parallel one another. That is, one variety may be highest for the trait in one location, but another variety may be highest in another location-or, at least, the difference between any two varieties may change location to location. Variety-location interaction also means that a location that is best for one variety will not be best for all. For agricultural breeders, a high degree of

interaction means that recommendations to farmers have to be made for a delimited set of locations or, for animals, for defined conditions of husbandry.

The residual variance is what remains after the preceding systematic variation for the trait has been taken into account; all that remains is the unsystematic variation between replicates within variety-location combinations. In analyses of human variation, this fraction is often labeled "non-shared environmental" variance and interpreted in terms of non-shared environmental factors (Plomin 1999; see critical review by Turkheimer 2000). Residual variance, however, is equally well non-shared "genetic" (variety) and interaction variance. In any case, to interpret the unsystematic variation in terms of differences in the underlying factors is to forget this second gap, namely, statistical patterns in traits are distinct from measurable factors.

Unreliable partitioning of human variation

The third gap lies between the values generated by the methods commonly used in analysis of human studies and the actual fractions of the overall variance corresponding to heritability, the "environmental" fractions, and so on. The standard methods do not reliably estimate the actual values. The methods need to be repaired and shorn of unsupported or unnecessary assumptions.

To understand how it could be the case that abundant and sophisticated published research that partitions data for human variation has resulted in unreliable estimates and interpretations, three points are key.

1. Similarity of relatives. When it is not possible to observe every variety raised in every location (i.e., cases like that in Figure 1, not the situation depicted in Figure 2), quantitative genetics analyzes variation in ways that take into account the genealogical relatedness of the individuals whose traits are observed. The standard methods assume that, all other things being equal, similarity in traits for relatives is proportional to the fraction shared by the relatives of all the genes that vary in the population (e.g., fraternal or dizygotic twins share half of the variable genes that identical or monozygotic twins share fully). However, plausible models of the contributions of multiple genes to a trait can be shown to result in, all other things being equal, ratios of dizygotic similarity to monozygotic similarity that are not .5 and that vary considerably around their average

(Taylor 2007). An example is given in the Appendix, but this point does not depend on the validity of any particular hypothetical model of multiple genes contributing to the trait. The assumption is unreliable because in quantitative genetics, the name of the field notwithstanding, the relevant correlations need to be based on observed *traits* and, as such, cannot be directly given by the proportion of shared *genes* involved in the development of those traits. For the same reason, heuristic values of the similarity of relatives of other degrees, which are ubiquitous, are also unreliable. (Identifying the exact fraction of genes shared by relatives [e.g., Visscher 2006] does not address this issue.) The methods can be adjusted to allow empirical estimation of parameters to take degree of relatedness into account provided the appropriate classes of data are available (Taylor 2012, Appendix 1). In this regard, data about unrelated individuals or varieties raised in the same location are especially valuable.

2. *Interaction variance*. The standard methods almost always omit reference to a variety-location-interaction variance fraction (Jacquard 1983). This means that fraction is subsumed in the among-variety-means variance fraction, that is, in the heritability estimate. Empirical estimation of the variety-location-interaction variance fraction is possible if the appropriate classes of data are available (e.g., data about monozygotic twins raised in separate locations in which there is no correlation from one twin's location to the other). Although the concerns that agricultural breeders have about high interaction variance (mentioned under the second gap) are not relevant for human research, the size of the interaction variance remains important to anyone wanting to claim that the "shared environmental" (location) fraction is of small importance or smaller importance than had been believed. To support such a claim requires showing not only that the location variance is a small fraction of the total variation, but the variety-location-interaction variance is as well.

3. *Non-essential genetic models*. Partitioning of trait variation into components rests traditionally on models of theoretical, idealized genes with simple Mendelian inheritance and direct contributions to the trait (Falconer and Mackay 1996; Lynch and Walsh 1998). Given that the data are about *traits*, it must also be possible to partition trait variation without referring to theoretical, idealized genes while making use of defined degrees of relatedness among kinds of individuals or varieties (Taylor 2012).

The virtues of such *gene-free* analysis include: making it difficult to forget the second gap or to omit variety-location interaction fraction; highlighting the classes of data needed for empirical estimation of parameters to take degree of relatedness into account; and providing a means to assess the implications of the three points raised under this third gap.

On the implications of the three points, suppose, as in the agricultural evaluation trial (Figure 2), there is as much data on a trait as could be needed. The correct fractions of the variation can be estimated making it is possible to examine how well the standard formulas (e.g., Rijsdijk and Sham 2002) recover those values. Taylor (2012) shows that: the residual fraction is recovered correctly; the so-called heritability estimates subsume the interaction fraction and are inflated or deflated according to whether the relatedness parameter is lower or higher than under the standard assumption; and the shared environmental fraction is correspondingly deflated or inflated. In short, as illustrated numerically in the Appendix (Figures 4 and 5), the standard formulas are unreliable estimates of the fractions of the variation in the trait.

Translation from fractions of variation to measurable factors

The fourth gap lies between fractions of variation and hypotheses about the underlying measurable genetic and environmental factors, or between the methods available and the methods needed for translation from variance fractions to hypotheses. This gap needs to be reckoned with. If it cannot be bridged, the third gap becomes moot and methods of analysis of variation among relatives need to find a basis quite different from that of classical quantitative genetics.

In conventional interpretations, a high heritability value indicates a strong genetic contribution to the trait. The finding that the variance of location elements ("shared environmental" variance) is a small fraction of the variation in human traits relative to the residual ("non-shared environmental") variance—a finding called into question by the third gap—is typically interpreted as the shared environment (e.g., socioeconomic status of the family) being less important (strictly: being associated with less variation in the trait) than social or environmental influences that vary for siblings within a family.

Such interpretations either overlook the second gap (in which fractions of variation in a trait are seen as distinct from measurable factors underlying the trait's development) or presume the existence of some method to expose the measurable genetic and environmental factors. The method might not be explicit, but an obvious initial step would be to assume that the variety elements in the additive models used for partitioning (see Figure 2) are related to the level of some genetic factor (or composite of genetic factors) that remain to be exposed. Similarly, it could be assumed that the location elements are related to the level of some composite of environmental factors, and that the residuals are related to some factors not captured by either of these relations. These assumptions are, however, questionable. In principle, twin studies could be conducted *even if the varieties were drawn from different species*, in which case one would not expect such a genetic-factor gradient to exist. Even if all varieties are from the same species, the genetic factors that influence the trait need not be the same for all varieties. Indeed, recalling the first gap, the combinations of underlying genetic and environmental factors may be heterogeneous. Notice, also, that the calculation of the variety elements involves averaging over a particular set of locations, which means that the variety elements, and thus the variance of these elements, are not properties of the varieties alone. (Similarly for location elements and their variance.)

Agricultural trials can allow generation of hypotheses about the genetic and environmental factors. By describing one way this happens, the difficulty in human research of bridging the fourth gap is accentuated. Whenever a number of varieties of animals or plants are raised or grown in multiple replicates over many locations, techniques of cluster analysis can eb used to group varieties by similarity in responses across all locations (Byth et al. 1976). Varieties in any resulting group tend to be above average for a location in the same locations and below average in the same locations. The wider the range of locations in the measurements on which the grouping is based, the more likely it is that the ups and downs shared by varieties in a group are produced by the same conjunctions of underlying genetic and environmental factors. This feature gives researchers some license to discount the possibility of underlying heterogeneity within a group, allowing them to hypothesize about the group averages—about what factors in the locations elicited basically the same response from varieties in a particular

variety group, a response that distinguishes the group from others. (It should be noted that data analysis is never self-sufficient; knowledge from other sources is always needed to help researchers generate their hypotheses about genetic and environmental factors.) However, clustering becomes infeasible when analyzing measurements from studies of human twins because such studies have only two replicates (twins) in one or at most two locations (families). In short, in agricultural research there is a path to bridge the fourth gap between fractions of variation and hypotheses about the underlying measurable genetic and environmental factors. However, the path is not one that research on human variation can follow.

If the fourth gap is considered together with the second and third, then it is not at all clear, *contra* the so-called laws of behavioral genetics captured by Turkheimer (2000), that heritability is substantial for all human behavioral traits, the influence of being raised in the same family (location) is smaller than the effect of genes, and a substantial portion of the variation in complex human behavioral traits is not accounted for by the effects of genes or families. *Contra* Plomin (1999 C26), it is not clear whether there is a trend for heritability to increase over people's lifetimes and, even if there were such a trend, this is not evidence that "genetic" differences come to eclipse "environmental" differences (see interaction variance under the third gap).

Fortunately, it is now possible to undertake research to identify the specific, measurable genetic and environmental factors without reference to the trait's heritability or the other fractions of the total variance (e.g., Moffitt et al. 2005, Davey Smith and Ebrahim 2007, Khoury et al. 2007). Yet, as indicated in the introduction, difficulties have become apparent in identifying causally relevant genetic variants for humans.

Differences within groups and among averages for separate groups

The fifth gap lies between within-group variation and between-group differences (i.e., variation among the means of the groups). The two kinds of variation have no logical or methodological relationship. This gap is widely acknowledged, but then sometimes hedged in the contentious debates about differences among the averages for racial and other groups when writers propose that high heritability confers plausibility on hypothesizing a role for genetic factors in explaining those differences (e.g., Jensen in

Miele 2002, 111ff). The within-group/between-group gap is, however, firm and its deep implications need to be kept always in mind.

If the fourth gap is not bridged, partitioning of variation among traits and heritability estimates provide little or no guidance in hypothesizing about measurable factors underlying observations of traits *within* one group of varieties. It follows logically that such analysis can provide little or no guidance about measurable factors associated with differences *between* the means of two groups.

Even if the fourth gap were bridged, hypothesizing about the difference between the average or *mean* values for varieties replicated within, but not across, locations is subject to the limitations of any nested analysis of variation. A textbook example following Lindman (1992) illustrates these limitations. Consider high school students' test scores in algebra viewed in relation to their teacher and school. A significant difference among the mean scores for the schools might, at first sight, be interpreted in terms of differences among the schools' facilities or organization. However, in practice, the students within a school could be randomly assigned to a teacher in their usual school, but neither the students or the teachers would have been assigned randomly across schools. The influences of the teachers in the different schools and the capacity of the students are also, therefore, involved in the differences among the schools' mean scores. The observed differences between schools could be due to some characteristic of the school as a whole, or to the fact that some schools have better teachers or their students are more avid learners, or to combinations of factors, such as students responding worse to teachers whose class-preparation time has been reduced because their school's administrators insist more on detailed documentation of student performance, and so on. In short, analysis of variation cannot help researchers hypothesize about the difference in the mean scores from one school to the next because the teachers are replicated in their students' test scores only within schools, not across schools. To translate this into the concerns here, nested analysis of variation cannot help researchers hypothesize about the difference in the mean scores from one location to the next when each variety is replicated only within some location. Researchers might just as well conduct a separate analysis for each subset of varieties and location. In the context of racial differences for human traits, this would mean a

separate analysis for each combination of group of individuals and experience of membership in different racial groups. (Note that to respect that nested analysis has this methodological limitation is not to make the claim that disjunct kinds of causes *must* be operating in the different racial groups.)

Implications for past and future analysis of human variation

The five fundamental gaps in quantitative genetics pose challenges to the common interpretations or key results of classical quantitative genetic analyses of variation, especially analyses of human variation. Recognition of the gaps might lead researchers, as well as historians and philosophers of science who have commented extensively on nature-nurture science (Downes 2004), to revisit studies that have interpreted heritability and "genetic variance" as a quantity that measures the contribution of the genetic factors in influencing variation in outcomes of the process through which the trait develops. Such a review might well, given the five gaps, find that key results and interpretations from many decades of human quantitative genetics are not justified or, at best, are unreliable.

Considering the gaps together and highlighting the possible heterogeneity of measurable genetic and environmental factors that underlie patterns for traits, further implications for understanding the analysis of human variation follow.

1. The use of ambiguous terms, such as "variation associated with genetic differences," obscures the gap between fractions of variation in a trait and measurable factors underlying the trait's development, thus making it harder to visualize the possibility of heterogeneity in the underlying measurable factors (first gap).

2. The translation from patterns in variation to hypotheses about measurable factors is possible in agricultural trials when, through clustering (as discussed under the fourth gap), groups can be defined within which underlying heterogeneity is minimized. This cannot be done in defining human groups.

3. Consider what happens if researchers put aside the search for measurable factors and, as is common in agricultural and laboratory breeding, focus on deriving reliable estimates of heritability as a fraction of the variation for the trait. If the actual advance under selective breeding is less than predicted, one source of the discrepancy

could be the underlying heterogeneity of genetic factors and their re-assortment through mating. However, whether this is the case matters little for breeders, because they can compensate for discrepancies: They discard the undesired offspring, breed the desired ones, and continue. This kind of selective breeding and compensating for discrepancies is not, of course, an acceptable option for humans.

4. If measurable factors underlying variation are identified for one group (presumably, for humans, by some means other than classical quantitative genetics), the possibility of underlying heterogeneity tempers any impulse to hypothesize that the same factors apply within other groups as well as to the difference between their means.

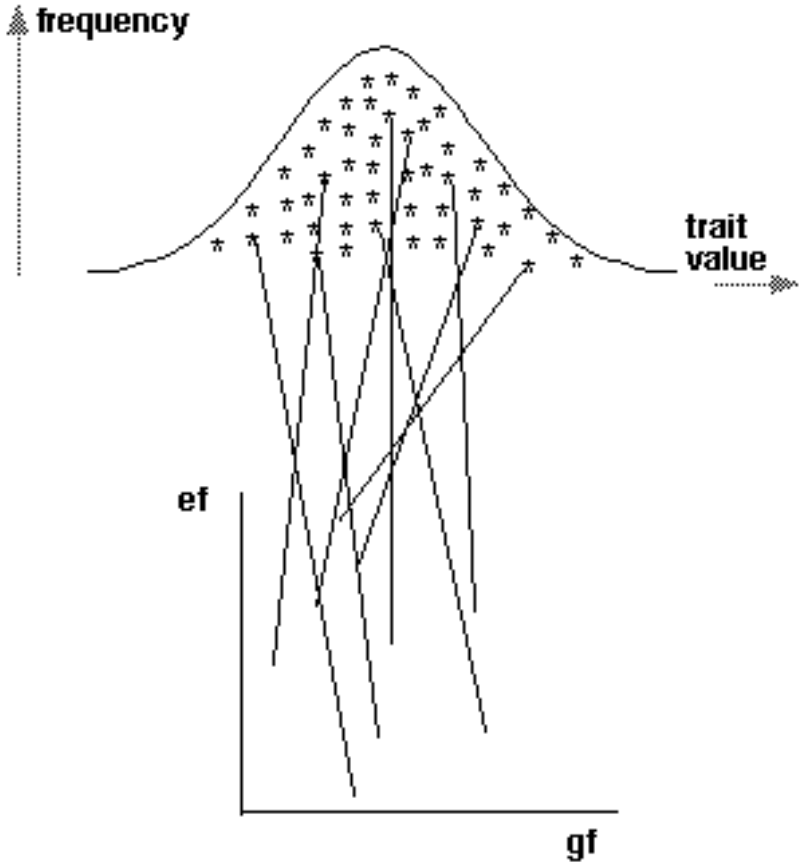

Figure 3. Values of a hypothetical trait for a population, showing one underlying genetic factor and one environmental factor (ef and gf) connected to the values for a sample of individuals. The crossing of the connector lines indicates schematically the first gap, i.e., the underlying factors are heterogeneous for this trait. The separation of the distribution

of the trait at the top and the plot of factors at the bottom reflects the second gap; the difficulty of inferring the factors from the distribution of the trait reflects the fourth gap.

This last observation makes this account of gaps in classical quantitative genetics relevant to other realms of biomedical and social science. Consider, as a pertinent illustration, the phenomenon of the large mean differences on IQ test scores between generations, which still lacks a satisfactory explanation. Dickens and Flynn (2001) propose "reciprocal causation" models, which involve two key features: a *matching* of environments to differences that may initially be small (e.g., children who show an earlier interest in reading will be more likely to be given books and receive encouragement for their reading and book-learning); and a *social multiplier* through which society's average level for the attribute in question influences the environment of the individual (e.g., if people grow up and are educated with others who, on average, have higher IQ test scores, this will stimulate their own development). Such models open up further challenges. Once it is recognized that the potency of social multipliers depends on the capacities of different groups to capitalize on historical changes in society, there is no reason to assume that the multipliers apply uniformly across individuals despite their differences in age, gender, geographical location, culture, and so on, or even that the multipliers move different individuals in the same direction but at different speeds. To adapt a basketball analogy that Dickens and Flynn employ, the onset of TV coverage of basketball acted as a social multiplier by eliciting greater participation in basketball but, at the same time, it elicited more couch-potato spectatorship. Now, once researchers envisage developmental pathways whose heterogeneous components differ among individuals at any given point of time, they have opened up the challenge of developing methods to collect and analyze the data so as to discriminate among many possible models of those pathways. The same challenge applies to explaining persisting gaps between mean IQ or achievement test scores for racially or ethically defined groups in the United States (Rampey et al. 2009).

As noted in the introduction, research on complex human traits now applies tools such as Genome-Wide Association that are more powerful than the formulas of classical quantitative genetics. However, because heritability estimates for human traits

are unreliable, a high heritability value is not a reliable guide for choosing which traits to explore at the molecular genetic level. Even when estimates are reliable in agricultural or laboratory breeding, it is always possible that the genetic and environmental factors underlying patterns in quantitative and other complex traits are heterogeneous. The possibility of underlying heterogeneity puts an exclamation point on the consensus that most medically significant traits are associated with many genes of quite small effect (McCarthy et al. 2008). It diminishes the utility for medical research and potential treatment not only of the results of quantitative genetics but also of Genome-Wide Association studies. The five fundamental gaps and appropriate responses to them should be understood by all researchers who want to build on "the estimation of genetic variance in populations [and move] to the detection and identification of variants that are associated with or directly cause variation" (Visscher et al. 2007) as well as teachers of the next generation of researchers.

Acknowledgements

This article is based on research supported by the National Science Foundation under grant SES–0634744. Some sentences and diagrams have been drawn from Taylor (2010, 2012). A visiting fellowship at the Konrad Lorenz Institute for Evolution and Cognition Research provided conditions conducive to the completion of the manuscript. I acknowledge thought-provoking comments from Richard Lewontin and Lindsay Morgia on an earlier version of this article.

References Cited

Byth DE, Eisemann RL, DeLacy IH. 1976. Two-way pattern analysis of a large data set to evaluate genotypic adaptation. Heredity 37: 215-230.

Couzin-Frankel J. 2010. Major Heart Disease Genes Prove Elusive. Science 328: 1220-1221.

Davey-Smith G, Ebrahim S. 2007. Mendelian randomization: Genetic variants as instruments for strengthening causal influences in observational studies. Pages 336-366 in Weinstein M, Vaupel JW, Wachter KW, eds. Biosocial Surveys. Washington, DC: National Academies Press.

Dickens WT, Flynn JR. 2001. Heritability estimates versus large environmental effects: The IQ

paradox resolved. Psychological Review 108: 346-369.

Downes SM. 2004. Heredity and Heritability in Zalta EN, ed. The Stanford Encyclopedia of Philosophy.

Falconer DS. 1960. Introduction to Quantitative Genetics. New York: The Ronald Press Company.

Jacquard A. 1983. Heritability: One word, Three concepts. Biometrics 39: 465-477.

Kendler KS, Prescott CA. 2006. Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric and Substance Abuse Disorders. New York: The Guilford Press.

Khoury MJ, Little J, Gwinn M, Ioannidis JP. 2007. On the synthesis and interpretation of consistent but weak gene-disease associations in the era of genome-wide association studies. International Journal of Epidemiology 36: 439-445.

Lindman HR. 1992. Analysis of Variance in Experimental Design. New York: Springer-Verlag.

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, Hirschhorn JN. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nature Reviews Genetics 9: 356-369.

McClellan J, King M-C. 2010. Genetic heterogeneity in human disease. Cell 141: 210-217.

Miele F. 2002. Intelligence, Race, and Genetics: Conversations with Arthur Jensen. Boulder, CO: Westview Press.

Moffitt TE, Caspi A, Rutter M. 2005. Strategy for investigating interactions between measured genes and measured environments. Archives of General Psychiatry 62: 473-481.

Neisser U, Boodoo G, et al. 1996. Intelligence: Knowns and unknowns. American Psychologist 51: 77-101.

Nuffield Council on Bioethics. 2002. Genetics and Human Behavior: The Ethical Context. http://www.nuffieldbioethics.org viewed 22 June 2007.

Plomin R, Asbury K. 2006. Nature and nurture: genetic and environmental influences on behavior. The Annals of the American Academy of Political and Social Science 600: 86-98.

Rampey B, Dion G, Donahue P. 2009. The Nation's Report Card: Long-Term Trend 2008. Washington, DC: National Center for Education Statistics.

Richardson K, Norgate S. 2005. The equal environments assumption of classical twin studies

may not hold. British Journal Educational Psychology 75: 339-350.

Rijsdijk FV, Sham PC. 2002. Analytic approaches to twin data using structural equation models. Briefings In Bioinformatics 3: 119–133.

Rutter M. 2002. Nature, Nurture, and Development: From Evangelism through Science toward Policy and Practice. Child Development 73: 1-21.

Taylor PJ. 2007. The unreliability of high human heritability estimates and small shared effects of growing up in the same family. Biological Theory: Integrating Development, Evolution and Cognition 2: 387-397.

—. 2010. Three puzzles and eight gaps: What heritability studies and critical commentaries have not paid enough attention to. Biology & Philosophy 25: 1-31.

—. 2012. A gene-free formulation of classical quantitative genetics used to examine results and interpretations under three standard assumptions. Acta Biotheoretica 60: 357-378.

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

Visscher PM, Medland SE, Ferreira MAR, Morley KI, Zhu G, Cornes BK, Montgomery GW, Martin NG. 2006. Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings. PLoS Genetics 2.

Visscher PM, Macgregor S, Benyamin B, Zhu G, Gordon S, Medland S, Hill WG, Hottenga J-J, Willemsen G, Boomsma DI, Liu Y-Z, Deng H-W, Montgomery GW, Martin NG. 2007. Genome Partitioning of Genetic Variation for Height from 11,214 Sibling Pairs. American Journal of Human Genetics 81: 1104-1110.

Appendix: Numerical illustrations of conceptual points from the article

1. Ratios of DZ similarity to MZ similarity under an illustrative model.

Consider a model in which the trait occurs when the combined "dosage" from 10 loci exceeds a threshold of 5, where each pair of alleles contributes a full, zero, or half dose according to whether the alleles are, respectively, both the same for one variant, same for the other, or one of each. Dominance is zero for the subset of the full results in the table below; location (environment) and noise (error) are not included in the model. The frequency of the first variant for a given locus is randomly chosen from the range in the first two columns. The third column gives the intraclass correlation for DZ twins. Given that the intraclass correlation for MZ twins is 1, the third column also gives the ratio of DZ similarity to MZ similarity. The average of these values is .60.

Frequency of first kind of allele chosen from range:		**Intraclass correlation for DZ twins**
Lower limit	**Upper limit**	
0.00	0.50	-0.01
0.49	0.51	0.78
0.38	0.63	0.61
0.25	0.75	0.53
0.13	0.88	0.73
0.00	1.00	0.60
0.75	0.75	0.81
0.69	0.81	0.56
0.63	0.88	0.53
0.56	0.94	0.41
0.50	1.00	1.00

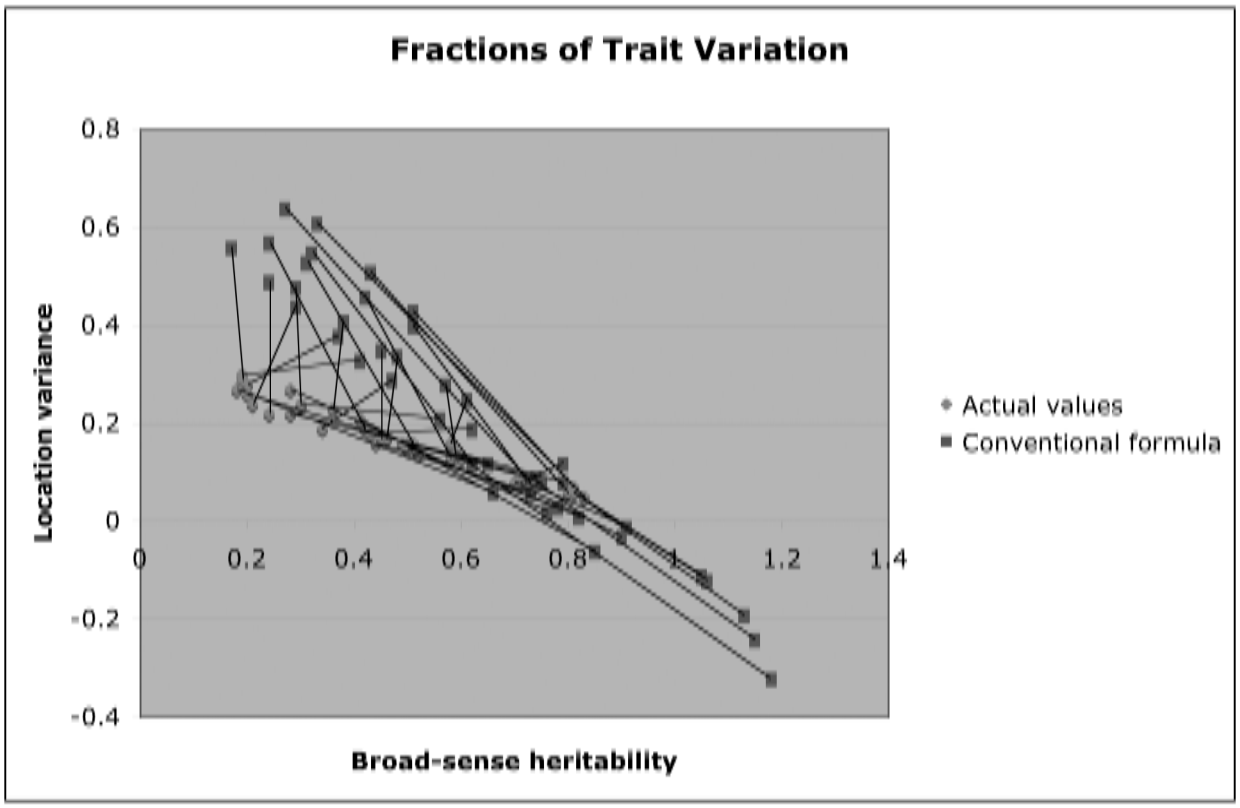

Figure 4. Contrast between the actual values and the values estimated using the conventional quantitative genetic formulas for a range of combinations of broad-sense heritability and the location-variance fraction (so-called shared environmental effect) (values drawn from Taylor 2012, Table 4).

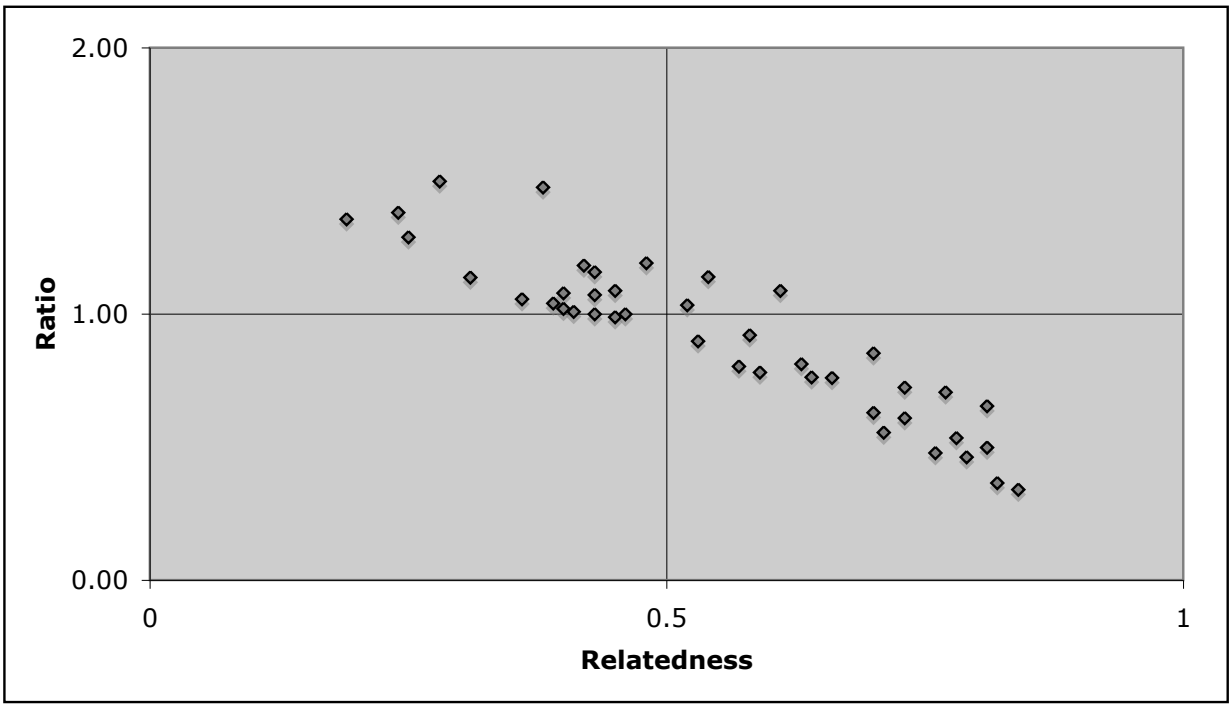

Figure 5. Ratio of the value estimated using the conventional quantitative genetic formula for broad-sense heritability and the sum of the actual values for heritability and interaction fraction of the variance plotted against the empirically estimated parameter for relatedness for dizygotic twins, a parameter that is conventionally assumed to be 0.5 (values drawn from Taylor 2012, Table 4).