Quantifying and Addressing Parameter Indeterminacy in the Classical Twin Design

Matthew C. Keller[1] and William L. Coventry[2,3]

[1]Center for Society and Genetics, University of California, Los Angeles, United States of America
[2]Queensland Institute of Medical Research, Brisbane, Australia
[3]School of Psychology, University of New England, Armidale, Australia

The *classical twin design* (CTD) is the most common method used to infer genetic and environmental causes of phenotypic variation. As has long been acknowledged, different combinations of the common environment/assortative mating, and additive, dominant, and epistatic genetic effects can lead to the same observed covariation between twin pairs, meaning that there is an inherent indeterminacy in parameter estimates arising from the CTD. The CTD circumvents this indeterminacy by assuming that higher-order epistasis is negligible and that the effects of either dominant genetic variation or the common environment are nonexistent. These assumptions, however, lead to consistent biases in parameter estimation. The current paper quantifies these biases and discusses alternative strategies for dealing with parameter indeterminacy in twin designs. One strategy is to model the similarity among other relatives in addition to twins (extended twin-family designs), which reduces but does not eliminate indeterminacy in parameter estimates. A more general strategy, applicable to all twin designs, is to present the parameter indeterminacy explicitly, as in a graph. Presenting the space of mathematically equally likely parameter values is important, not only because it aids the proper interpretation of twin design findings, but also because it keeps behavioral geneticists themselves mindful of methodological assumptions that can easily go unexamined.

At the heart of all twin designs is the comparison between MZ (monozygotic) and DZ (dizygotic) twin similarities. The most common of these designs, the *classical twin design* (CTD), compares only MZ and DZ twins reared together, and has been the mainstay of efforts to understand the genetic and environmental causes of behavior in humans over the last 30 years. The popularity of the CTD is attributable not only to its ability to inform the degree of genetic and environmental causes of variation in traits, but also to its convenience: MZ and DZ twins reared together are easier to find than other types of genetically informative relatives, such as randomly placed adopted twins. Because of this, very large twin registries in many countries have been launched (Boomsma et al., 2002).

In some cases, follow-up efforts have gathered data from twins' parents, nontwin siblings, spouses, and offspring, allowing for analyses employing the *extended twin-family design* (ETFD). While the CTD and ETFD are generally thought of as tools for discerning environmental and genetic causes of phenotypic variation, the power afforded by the sizes of data sets from modern twin registries have allowed testing increasingly nuanced hypotheses involving sex differences in genetic and environmental influences, sibling effects, gene-by-environment interactions, scale measurement issues, and the environmental and genetic relationships between multiple variables (Neale & Maes, 2004).

Despite its popularity and convenience, it has long been understood that the CTD is limited in its ability to distinguish many potential causes of phenotypic variation (Eaves et al., 1978; Jinks & Fulker, 1970; Martin et al., 1978). The CTD is unable to differentiate effects of the common environment from the effects of assortative mating, gene–environment covariation, or sibling imitation. Similarly, genetic sources of variation are confounded with gene-by-common environment interactions, sibling competition, and with systematic difference in the ways that MZ and DZ twins are treated. Several of these issues have been explored in depth and do not seem to pose problems for interpreting CTD findings. For example, the well-known *equal environments assumption* of twin designs, whereby twins' environments are assumed to affect MZ and DZ twins equivalently with respect to the phenotype in question, has repeatedly been showed to be valid in most cases (e.g., Borkenau et al., 2002; Kendler et al., 1993).

A problem more at the heart of twin designs, which affects both the CTD and, to a lesser degree, the ETFD, is their inability to fully differentiate common environmental effects, additive genetic effects, dominant genetic effects, and epistatic genetic effects. Different combinations of these can lead to

exactly the same observed data, meaning that there is an inherent indeterminacy in parameter estimates. This indeterminacy is the source of the mantra that nonadditive genetic and common environmental variance parameters cannot simultaneously be estimated in the CTD, but the issue goes much deeper than this, as has been discussed with regard to genetic dominance (Eaves et al., 1978; Jinks & Fulker, 1970; Martin et al., 1978) and epistasis (Grayson, 1989).

Despite the attention the subject has received, we argue that fruitful discussion has yet to be exhausted regarding parameter indeterminacy in twin designs. Because it is the most commonly used twin design, the current paper focuses chiefly on the CTD. The accepted method of dealing with parameter indeterminacy in the CTD leads to consistent biases in parameter estimates. We first quantify these biases and rectify an error in the literature regarding parameter bias in the presence of epistasis. We then discuss parameter indeterminacy in the CTD and gauge which of the two causes of this indeterminacy is the most consequential. Finally, we discuss alternative methods of dealing with parameter indeterminacy in the CTD — specifically, the merits of modeling broad- rather than narrow-sense heritability, methods of presenting parameter indeterminacy, and ways to limit parameter indeterminacy.

Twin designs are undoubtedly useful vehicles for scientific discovery, and yet, as with all scientific methods, they are not infallible. Because twin designs are based upon a robust theory — biometric genetics — researchers using them have the advantage of being able to quantify the alternative explanations for their data. This ability, we argue, should be utilized.

The Classical Twin Design

Ignoring for the moment several complications (see below), phenotypic variation can be separated into additive genetic (V_A), nonadditive genetic (V_{NA}), common environmental (V_C), and unique environmental (V_E) variation. Twins inform the magnitude of these parameters because, although both types of twins share a common familial environment, MZ twins share 100% of their genes while DZ twins share (on average) 50% of their genes. DZ twins — as opposed to nontwin siblings — serve as natural controls for MZ twins because both types of twins are matched for age, the familial environment, and aspects of the environment unique to twins (including *in utero* effects). Thus, the degree to which MZ twins' resemblance exceeds DZ twins' resemblance probably reflects genetic variation ($V_G = V_A + V_{NA}$). If DZ twins' resemblance is less than half of the MZ twins' resemblance, an effect of nonadditive genetic variation (V_{NA}, referred to usually as dominance variation, or V_D) is suggested. When DZ twins' resemblance is more than half of the MZ twins' resemblance, some aspect of common environmental variation (V_C) may affect the phenotype.

Several complicating factors, such as gene-by-environment interactions, assortative mating, sibling interaction effects, and gene–environment correlations make the simple decomposition of variance components outlined above difficult. Some of these issues can be overcome using the CTD. For example, sibling interaction effects can be detected as a difference in total variance between MZ and DZ twins. However, these issues have been discussed in detail elsewhere (e.g., Neale & Maes, 2004), and so we limit discussion of them except for the effect of assortative mating ('like marrying like'), which is confounded with the effects of the common environment in the CTD (Eaves et al., 1978; Neale & Maes, 2004). This occurs because assortative mating tends to increase the genetic relationship between DZ twins for the trait in question above the expected .50 level, and thereby mimics the effects of the common environment. Thus, V_C in the CTD (and as used in the present paper) reflects not only the effects of the common environment, but also the genetic effects of assortative mating. Disentangling these two aspects of V_C is not possible using the CTD.

The way that the hypothesized variance components affect twin covariances can be quantified:

$$CV_{MZ} = V_P - V_E \qquad [1]$$

$$CV_{MZ} = V_A + V_{NA} + V_C \qquad [2]$$

$$CV_{DZ} = \tfrac{1}{2}V_A + rV_{NA} + V_C \qquad [3]$$

where CV_{MZ} is the covariance between MZ twins, CV_{DZ} is the covariance between DZ twins, V_P is the overall phenotypic variance, and r is a scalar for V_{NA}, which ranges from ¼ to 0 (see next section).

Since CV_{MZ}, CV_{DZ}, and V_P are estimated from observations, equations 1 to 3 can be used to estimate V_A, V_{NA}, V_C, and V_E. From algebra, the number of independently informative equations must be equal to or greater than the number of unknown parameters for the unknown parameters to be estimable. Parameter indeterminacy arises in the CTD because five unknowns (V_A, V_{NA}, V_C, V_E and r) are estimated by only three informative equations [1 to 3]. Researchers must therefore fix two underlying parameters at assumed values in order to estimate the remaining parameters. In the CTD, these fixed parameters are $\hat{r}$ (fixed to ¼) and either $\hat{V}_C$ or $\hat{V}_{NA}$ (one of which is fixed to zero).[1] Typically, $\hat{V}_{NA}$ is fixed to zero when the ratio $C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq \tfrac{1}{2}$ while $\hat{V}_C$ is fixed to zero when $C\hat{V}_{DZ}/C\hat{V}_{MZ} \leq \tfrac{1}{2}$. This rule of thumb is sensible, if one must fix either $\hat{V}_C$ or $\hat{V}_{NA}$ to zero, because it ensures that none of the estimated variance parameters will be negative (or at a zero boundary if constrained to be positive). $\hat{V}_E$ can be estimated regardless of which parameters are fixed because it is fully determined by two observed values, $\hat{V}_P$ and $C\hat{V}_{MZ}$.

Fixing two parameters in the CTD is a useful trick for circumventing the parameter indeterminacy in what, strictly, is an unsolvable set of equations. In the real world, however, it is reasonable — and even likely — to expect common environmental, nonadditive

genetic, and additive genetic factors to all simultaneously affect the phenotype. Moreover, and contrary to the mantra, it is perfectly valid mathematically to estimate V_C and V_{NA} simultaneously so long as $\hat{A}$ and $\hat{r}$ are fixed. For reasons explained below, the tradition of setting $\hat{r}$ to ¼ and either $\hat{V}_C$ or $\hat{V}_{NA}$ to zero generally leads to overestimates of $\hat{V}_A$ at the expense of estimates of both $\hat{V}_{NA}$ and $\hat{V}_C$ (Eaves et al., 1978; Grayson, 1989; Jinks & Fulker, 1970; Martin et al., 1978).

Decomposition of Genetic Nonadditivity

Although DZ twins and other full siblings share on average half of their alleles, the effects of these alleles are not usually shared by half. This situation, due to genetic nonadditivity, occurs when the phenotypic effects of alleles depend upon other alleles. This is called *dominance* when a given allele's effect depends upon its 'sister' allele at the same locus of the homologous chromosome and *epistasis* when its effect depends upon one or more alleles at separate loci. The number of other alleles modifying a given allele's effect reduces the covariation between DZ twins or other full siblings on some phenotype (Mather, 1974). Specifically,

$$CV_{DZ} = V_C + ½V_1 + ¼V_2 + ⅛V_3 +$$
$$⅟₁₆V_4 + ⅟₃₂V_5 + ... + (½)^n V_n \qquad [4]$$

where $V_1 = V_A$ is the variation in the marginal effect of alleles (depending only upon the effects of alleles taken individually), V_2 is the variation in effects that depend upon combinations of two alleles (i.e., dominance or additive-by-additive epistasis), V_3 is the variation in effects that depend upon combinations of three alleles (e.g., additive-by-dominance epistasis among others), and so forth for each of n alleles affecting the phenotype in question.[2] Thus, from equation [3],

$$rV_{NA} = ¼V_2 + ⅛V_3 + ⅟₁₆V_4 +$$
$$⅟₃₂V_5 + ... + (½)^n V_n \qquad [5]$$

Although the scalar $\hat{r}$ is set to ¼ in the CTD, it can be seen that its true population value is equal to an average of the scalars, ¼, ⅛, ⅟₁₆, ..., (½)^n, weighted by the magnitude (relative to V_{NA}) of each corresponding variance component V_2 to V_n:

$$r = ¼V_2/V_{NA} + ⅛V_3/V_{NA} + ⅟₁₆V_4/V_{NA} +$$
$$⅟₃₂V_5/V_{NA} + ... + (½)^n V_n/V_{NA} \qquad [6]$$

Thus, $r = ¼$ only when V_{NA} is solely due to dominance and/or additive-by-additive epistasis (i.e., when $V_3 = V_4 = V_5 = ... = V_n = 0$), and otherwise ranges from an upper bound of ¼ to a lower bound that theoretically approaches zero.

If all the alleles and relevant loci were known, it would be possible to estimate r precisely. In practice, however, it is generally impossible to assess the value of r from trait-level analyses in humans because there are too many parameters (V_2, V_3, ... V_n) and their intercorrelations would be too high for any data set to possibly

estimate (Eaves, 1988). Observing parent–offspring correlations that are of similar magnitude to sibling–sibling correlations would suggest that epistatic effects that are purely additive in nature (e.g., additive-by-additive epistasis) probably contribute more to V_{NA} than does dominance or epistasis with dominance components (Heath et al., 1984; Mather, 1974). However, comparing parent–offspring and sibling–sibling correlations does not allow r to be estimated: similar parent–offspring and sibling–sibling correlations can result from interactions from any number of additive loci, and thus could result from many different values of r.

Quantitative trait loci analyses in experimentally bred animals suggest that epistasis is quite common (Li et al., 1997; Long et al., 1995; Shook & Johnson, 1999; Yu et al., 1997), with a recent review estimating that up to a third of the genetic variation in carefully studied animal traits is due to epistasis (Carlborg & Haley, 2004). Given that epistasis appears to be nontrivial in those traits where it has been detectable, it seems unlikely that r will be exactly ¼ in most of the complex and polygenic traits of interest to researchers of human variation. Nevertheless, for reasons discussed below, it is probably not biologically plausible for values of r to approach their theoretical lower bound of zero.

Parameter Biases in the Classical Twin Design

It is relatively straightforward to quantify the degree of bias that exists in parameter estimates when the CTD conventions of fixing parameters are followed. Turning first to the situation $C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq 1/2$, using the fixed parameter values of the CTD ($\hat{V}_{NA} = 0$ and $\hat{r} = ¼$) and algebraic manipulation of equations [2] and [3] leads to the familiar equations:

$$\hat{V}_A = 2(C\hat{V}_{MZ} - C\hat{V}_{DZ}) \qquad [7]$$
and
$$\hat{V}_C = 2C\hat{V}_{DZ} - C\hat{V}_{MZ} \qquad [8]$$

Substituting the actual population values CV_{MZ} and CV_{DZ} from equations [3] and [4] for their estimates $C\hat{V}_{MZ}$ and $C\hat{V}_{DZ}$ in equations [7] and [8], we obtain the true population variance compositions of $\hat{V}_A$ and $\hat{V}_C$ in the CTD when $C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq 1/2$:

$$\hat{V}_A = V_A + 2(1 - r)V_{NA} \qquad [9]$$
$$\hat{V}_C = V_C - (1 - 2r)V_{NA} \qquad [10]$$

replicating the findings by Grayson (1989). Thus, $\hat{V}_A$ overestimates V_A by $2(1-r)V_{NA}$ while $\hat{V}_C$ underestimates by V_C by $(1-2r)V_{NA}$ in the CTD. Intuitively, this occurs because nonadditive genetic variation strongly reduces the similarity of DZ twins relative to MZ twins, masking the effects of the common environment and mimicking (when combined with common environmental effects) the effects of additive genes. The lower r is (the more higher-order epistasis exists), the greater these biases become. Obviously, since $\hat{V}_{NA}$ is set to zero, V_{NA} is underestimated by whatever its true value is.

Turning to the situation $C\hat{V}_{DZ}/C\hat{V}_{MZ} \leq \frac{1}{2}$ and using the assumed values $\hat{V}_C = 0$ and $\hat{r} = \frac{1}{4}$, we can once again use equations [2] and [3] to find the estimates of V_A and V_{NA}:

$$\hat{V}_A = 4C\hat{V}_{DZ} - C\hat{V}_{MZ} \qquad [11]$$

and

$$\hat{V}_{NA} = 2C\hat{V}_{MZ} - 4C\hat{V}_{DZ} \qquad [12]$$

Substituting the population parameters into the right sides of equations [11] and [12] leads to the population variance compositions of $\hat{V}_A$ and $\hat{V}_{NA}$ in the CTD when $C\hat{V}_{DZ}/C\hat{V}_{MZ} \leq \frac{1}{2}$:

$$\hat{V}_A = V_A + 3V_C - (1 - 4r)V_{NA} \qquad [13]$$

$$\hat{V}_{NA} = 2(1 - 2r)V_{NA} - 2V_C \qquad [14]$$

These equations are different to those obtained by Grayson (1989), who may have erred by fixing the population parameter, r, to $\frac{1}{4}$ while allowing assumed values of $\hat{r}$ to vary, which is the opposite of what is required.[3] Thus, contra Grayson (1989), who argued that $\hat{V}_A$ always overestimates V_A, V_A can be overestimated or underestimated by $\hat{V}_A$ in the CTD depending on whether $3V_C - (1 - 4r)V_{NA}$ is positive or negative, respectively. When $r = \frac{1}{4}$, $\hat{V}_A$ always overestimates V_A, but when $r < \frac{1}{4}$, $\hat{V}_A$ can underestimate V_A in the presence of little V_C. For example, it is not uncommon for $C\hat{V}_{DZ}/C\hat{V}_{MZ} < \frac{1}{4}$, which may appear as $\hat{V}_A \sim 0$ and very high $\hat{V}_{NA}$ using the CTD, but this more likely represents higher-order epistasis ($r < \frac{1}{4}$) and a concomitant underestimation of V_A. By the same token, V_{NA} can be overestimated when $r < \frac{1}{4}$ and V_C is small, or underestimated when $r \sim \frac{1}{4}$ and V_C is not negligible. It should be noted that when $r \sim \frac{1}{4}$, the effects of assortative mating and/or the common environment mask V_{NA} disproportionately (by a factor of roughly $2V_C$), and thus V_{NA} may often be underestimated in CTD studies.

As noted by Grayson (1989), the normal method of fixing parameters used in the CTD always minimizes estimates of V_C, which seems to lead to the conclusion that CTD studies always report the lowest possible estimates of the effects of the common (familial) environment. However, it must be remembered that the effects of assortative mating are confounded with the effects of the common environment in the CTD, meaning that there is no clear prediction of the direction of bias in parameter estimates of the common environment in CTD studies.
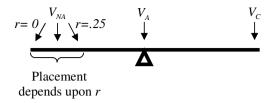
Parameter Indeterminacy and Parameter Space in the Classical Twin Design

Exact parameter values can be estimated using the CTD only because two unknown parameters are fixed at canonical (and somewhat arbitrary) values. This practice leads to the consistent biases quantified above. We now turn to examining the full range of parameter values that are possible when the values normally fixed in the CTD are allowed to vary.
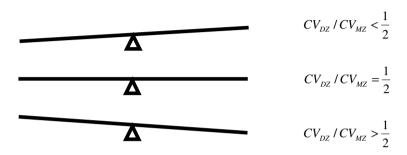
Parameter indeterminacy occurs when different combinations of parameters lead to the same observed statistics, and therefore only a range of possible parameter values, and not their exact values, can be estimated. Parameter indeterminacy in the CTD can be visualized using a *see-saw* analogy (Figure 1). The magnitudes of the three variance parameters are represented by blocks of weight placed upon a see-saw. The V_A weight is placed above the fulcrum, the V_C weight is placed on one end of the see-saw, and assuming that $r = \frac{1}{4}$, the V_{NA} weight is placed midway between the other end of the see-saw and the fulcrum. Lower values of r slide the V_{NA} weight away from the fulcrum toward the left end of the see-saw (Figure 1A). The degree to which the see-saw tilts to one side or the other, corresponding to different levels of $C\hat{V}_{DZ}/C\hat{V}_{MZ}$, provides information about the magnitude of these blocks of weight (Figure 1B). For the analogy to work, we must imagine that (a) unlike normal see-saws, this see-saw tilts only *to the degree* that one side is heavier than the other; and (b) the blocks of weight cannot be seen, so their sizes must be inferred based upon the tilt in the see-saw. Figure 1C shows different combinations of weights and weight placements that all lead to the same tilt in the see-saw — in this case, a see-saw that is evenly balanced. This simple and admittedly imperfect analogy nevertheless gives an intuitive feel for the reason behind parameter indeterminacy in the CTD: many different hypothesized weight sizes and placements can lead to the same observed see-saw tilt.

As illustrated in Figure 1, parameter bias exists in the CTD because the unobservable parameter values probably differ from their assumed values. Systematically varying these assumed values and noting the corresponding estimated parameters reveals the *parameter space* of a given data set, each point within the space corresponding to different combinations of parameters which are, mathematically, equally valid in explaining twin variances and covariances. Describing a data set's parameter space simply involves finding its boundaries. Table 1 shows how fixing different combinations of two parameters, $\hat{r}$ and one of the other variance parameters, leads to different methods of estimating the other two parameters.[4] These methods were derived in the same fashion as were equations [7], [8], [11] and [12], and each method will provide an equivalent fit to any twin-data set. Given that zero is the natural minimum for variance parameters, Table 1 also shows the maximum and minimum for each variance parameter, and consequently the boundaries of the full parameter space. Note that rows 1 to 4 in Table 1 amount to varying $\hat{r}$ from its minimum (zero) to its maximum ($\frac{1}{4}$) and varying $\hat{V}_{NA}$ and from its minimum (zero) to its maximum (which occurs when $\hat{V}_A = 0$). Similarly, rows 5 to 8 in Table 1 amount to varying $\hat{r}$ and $\hat{V}_C$ from their minima to their maxima.

A. Placements of weights on see-saw.

B. Different tilts of see-saw are analogous to different observations of CV_{DZ} / CV_{MZ}.

$$CV_{DZ} / CV_{MZ} < \frac{1}{2}$$

$$CV_{DZ} / CV_{MZ} = \frac{1}{2}$$

$$CV_{DZ} / CV_{MZ} > \frac{1}{2}$$

C. Parameter indeterminacy: possible variance parameters given the same tilt, $CV_{DZ} / CV_{MZ} = \frac{1}{2}$

Classical Twin Design estimates

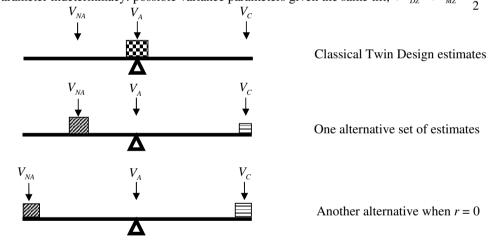One alternative set of estimates

Another alternative when $r = 0$

Figure 1
See-saw analogy of parameter indeterminacy in the Classical Twin Design.

Gauging Which Source of Parameter Indeterminacy is the Most Consequential

There are two sources of the parameter indeterminacy in the CTD: (1) the indeterminacy caused by altering assumed values of $\hat{r}$, and (2) the indeterminacy caused by altering assumed values of $\hat{V}_C$ or $\hat{V}_{NA}$. It is of interest to understand which of these two sources causes the most indeterminacy in the estimated parameters, and hence is the greatest potential source of parameter bias in the CTD. 'Potential' bias refers to how much a parameter can vary; it is potential because the actual bias of a parameter from some data set is, of course, unknowable. We define the potential bias of a given variance parameter that is due to r being unknown as the range of values that variance parameter can take when r is allowed to vary from its minimum to its maximum. For example, one could find the range of values (the potential bias) of $\hat{V}_A$ that occurs when $\hat{r}$ varies from zero to $^1/4$. Similarly, we define the potential bias of a variance parameter that is due to the convention of fixing $\hat{V}_C$ or $\hat{V}_{NA}$ to zero as the range of values that variance parameter can take when $\hat{V}_C$ (if $C\hat{V}_{DZ}/C\hat{V}_{MZ} \leq \ ^1/2$) or $\hat{V}_{NA}$ (if

Table 1

Methods of Obtaining $\hat{V}_A$, $\hat{V}_C$, and $\hat{V}_{NA}$ from $C\hat{V}_{MZ}$ and $C\hat{V}_{DZ}$ and the Boundaries of the Parameter Space Given Eight Different Pairs of Fixed Parameters Possible in Twin-Only Designs

Fixed Parameters	$\hat{V}_A$	$\hat{V}_{NA}$	$\hat{V}_C$
$C\hat{V}_{DZ}/C\hat{V}_{MZ} > \, ^1/_2$			
1. $\hat{r} = 0, \hat{V}_A = 0$	0 *min*	$C\hat{V}_{MZ} - C\hat{V}_{DZ}$ *inter*	$C\hat{V}_{DZ}$ *max*
2. $\hat{r} = 0, \hat{V}_{NA} = 0$	$2(C\hat{V}_{MZ} - C\hat{V}_{DZ})$ *max*	0 *min*	$2C\hat{V}_{DZ} - C\hat{V}_{MZ}$ *min*
3. $\hat{r} = .25, \hat{V}_A = 0$	0 *min*	$^4/_3(C\hat{V}_{MZ} - C\hat{V}_{DZ})$ *max*	$^4/_3 C\hat{V}_{DZ} - ^1/_3 C\hat{V}_{MZ}$ *inter*
4. $\hat{r} = .25, \hat{V}_{NA} = 0$	$2(C\hat{V}_{MZ} - C\hat{V}_{DZ})$ *max*	0 *min*	$2C\hat{V}_{DZ} - C\hat{V}_{MZ}$ *min*
$C\hat{V}_{DZ}/C\hat{V}_{MZ} < \, ^1/_2$			
5. $\hat{r} = 0, \hat{V}_A = 0$	0 *min*	$C\hat{V}_{MZ} - C\hat{V}_{DZ}$ *inter*	$C\hat{V}_{DZ}$ *max*
6. $\hat{r} = 0, \hat{V}_C = 0$	$2C\hat{V}_{DZ}$ *max*	$C\hat{V}_{MZ} - 2C\hat{V}_{DZ}$ *min*	0 *min*
7. $\hat{r} = .25, \hat{V}_A = 0$	0 *min*	$^4/_3(C\hat{V}_{MZ} - C\hat{V}_{DZ})$ *max*	$^4/_3 C\hat{V}_{DZ} - ^1/_3 C\hat{V}_{MZ}$ *inter*
8. $\hat{r} = .25, \hat{V}_C = 0$	$4C\hat{V}_{DZ} - C\hat{V}_{MZ}$ *inter*	$2C\hat{V}_{DZ} - 4C\hat{V}_{MZ}$ *inter*	0 *min*

Note: max = maximum, min = minimum, and inter = intermediate parameter estimates within each group (rows 1–4 and 5–8) of fixed parameters.

$\hat{V}_C$ is not fixed to 0 when $C\hat{V}_{DZ}/C\hat{V}_{MZ} > \, ^1/_2$ (rows 1–4) because this implies that $\hat{V}_{NA} < 0$ and similarly $\hat{V}_{NA}$ is not fixed to 0 in rows 5–8 because this implies that $\hat{V}_C < 0$.

Rows 4 and 8 represent the CTD.

$C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq \, ^1/_2$) are allowed to vary from their minima to their maxima. If the ranges of variance parameters are greatest across different fixed values of $\hat{r}$, then the convention of fixing $\hat{r} = \, ^1/_4$ would be the greatest source of potential bias in the CTD, and vice versa for different fixed values of $\hat{V}_C$ or $\hat{V}_{NA}$. Which of these is the greatest potential source of bias in CTD estimates?

The absolute differences between the parameter estimates of rows 1 and 2 in Table 1 give the ranges of the three variance parameter values that are caused by varying $\hat{V}_{NA}$ from its minimum (zero) to its maximum (which occurs when $\hat{V}_A = 0$) given that $\hat{r}$ is zero and that $C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq \, ^1/_2$. In this case, for example, the range in

$\hat{V}_C$ values across all possible values of $\hat{V}_{NA}$ is $C\hat{V}_{MZ} - C\hat{V}_{DZ}$. This quantity is maximized when $C\hat{V}_{DZ} = \, ^1/_2 C\hat{V}_{MZ}$ (given the constraint that $C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq \, ^1/_2$), and thus $C\hat{V}_{MZ} - \, ^1/_2 C\hat{V}_{MZ} = \, ^1/_2 C\hat{V}_{MZ}$ is the *maximal* potential bias in $\hat{V}_C$ when $\hat{r} = 0$ and $C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq \, ^1/_2$. In a similar way, the maximum potential parameter biases due to altering assumed values of $\hat{V}_C$ or $\hat{V}_{NA}$ are isolated by finding the absolute differences in parameter estimates between rows 1 and 2, rows 3 and 4, rows 5 and 6, and rows 7 and 8 in Table 1. The maximum parameter biases due to altering assumed values of $\hat{r}$ are isolated by finding the absolute differences in parameter estimates between rows 1 and 3, rows 2 and 4, rows 5 and 7, and rows 6 and 8.

We calculated all the differences in parameter estimates as indicated above to gain insight into which source of bias is most influential. We found that the effects of altering assumed values of $\hat{V}_C$ or $\hat{V}_{NA}$ are greatest when $C\hat{V}_{DZ} = \, ^1/_2 C\hat{V}_{MZ}$ and nonexistent when $C\hat{V}_{DZ}/C\hat{V}_{MZ} = 1$ or 0, while the effects of altering assumed values of $\hat{r}$ are greatest when $C\hat{V}_{DZ} = \, ^1/_4 \hat{V}_{MZ}$ and nonexistent when $C\hat{V}_{DZ}/C\hat{V}_{MZ} = 1$ or 0. To further complicate matters, the effects of both sources of indeterminacy depend upon the levels of the other fixed parameter. To simplify these relationships, Table 2 quantifies the *maximal* potential bias (i.e., the maximal range of parameter values) for every variance parameter separately for the two sources of bias. In most situations, the second source of bias, due to altering assumed values of $\hat{V}_C$ or $\hat{V}_{NA}$, has greater potential impact on parameter estimates than does altering assumed values of $\hat{r}$. However, not shown in Table 2, the potential bias due to altering assumed values of $\hat{r}$ increases as the ratio $C\hat{V}_{DZ}/C\hat{V}_{MZ}$ shrinks. When $C\hat{V}_{DZ} \leq \, ^1/_4 C\hat{V}_{MZ}$, the two sources of bias are equally influential (additional details available from the author upon request).

Modeling Broad-Sense Heritability

Of the two sources of parameter indeterminacy in the CTD, we have shown that the inability to simultaneously estimate both $\hat{V}_C$ and $\hat{V}_{NA}$ — or equivalently, the inability to simultaneously estimate all three variance parameters — should generally have the greatest potential to bias parameter estimates. An alternative modeling approach might therefore be to collapse two variance parameters into a single parameter. This can

Table 2

Comparison of the Maximal Biases of $\hat{V}_A$, $\hat{V}_{NA}$, and $\hat{V}_C$, that are due to the Two Sources of Parameter Indeterminacy in Twin-Only Designs

Parameter	Indeterminacy due to varying $\hat{r}$		Indeterminacy due to varying $\hat{V}_{NA}$ or $\hat{V}_C$	
	Maximum bias	Where maximum bias occurs	Maximum bias	Where maximum bias occurs
$\hat{V}_A$	$^1/_2 C\hat{V}_{MZ}$	$C\hat{V}_{DZ} = \, ^1/_4 C\hat{V}_{MZ}$ & $\hat{V}_C = 0$	$C\hat{V}_{MZ}$	$C\hat{V}_{DZ} = \, ^1/_2 C\hat{V}_{MZ}$ & $\hat{r} = \, ^1/_4$
$\hat{V}_{NA}$	$^1/_2 C\hat{V}_{MZ}$	$C\hat{V}_{DZ} = \, ^1/_4 C\hat{V}_{MZ}$ & $\hat{V}_C = 0$	$^2/_3 C\hat{V}_{MZ}$	$C\hat{V}_{DZ} = \, ^1/_2 C\hat{V}_{MZ}$ & $\hat{r} = \, ^1/_4$
$\hat{V}_C$	$^1/_4 C\hat{V}_{MZ}$	$C\hat{V}_{DZ} = \, ^1/_4 C\hat{V}_{MZ}$ & $\hat{V}_A = 0$	$^1/_2 C\hat{V}_{MZ}$	$C\hat{V}_{DZ} = \, ^1/_2 C\hat{V}_{MZ}$ & $\hat{r} = 0$

be done by modeling broad-sense heritability explicitly, estimating $\hat{V}_G = \hat{V}_A + \hat{V}_{NA}$ rather than $\hat{V}_A$ and $\hat{V}_{NA}$ separately. The expected covariances between MZ and DZ twins would simply be:

$$CV_{MZ} = V_G + V_C \qquad [15]$$

$$CV_{DZ} = qV_G + V_C \qquad [16]$$

where

$$q = {}^1\!/_2\, V_A/V_G + r\, V_{NA}/V_G \qquad [17]$$

The value q represents an amalgam of both r and the relative degree of V_A and V_{NA}, and will range from almost zero (when both V_A and $r \sim 0$) to $^1\!/_2$ (when $V_{NA} = 0$). Solving for the estimates of V_G and V_C, we have:

$$\hat{V}_G = \frac{1}{1-\hat{q}}(C\hat{V}_{MZ} - C\hat{V}_{DZ}) \qquad [18]$$

$$\hat{V}_C = (1 - \frac{1}{1-\hat{q}})\, C\hat{V}_{MZ} + \frac{1}{1-\hat{q}}\ C\hat{V}_{DZ} \qquad [19]$$

As is evident from equations [18] and [19], in all cases, $\hat{V}_G$ is maximized when $\hat{q} = {}^1\!/_2$ and minimized when $\hat{q} = 0$, while $\hat{V}_C$ is maximized when $\hat{q} = 0$ and minimized when $\hat{q} = {}^1\!/_2$.

To estimate $\hat{V}_G$ and $\hat{V}_C$, $\hat{q}$ must be fixed to some value. Table 3 shows how fixing different values of $\hat{q}$ alters the parameter estimates for four hypothetical observations of $C\hat{V}_{DZ}$ and $C\hat{V}_{MZ}$. Table 3 also shows that as the value $C\hat{V}_{DZ}/C\hat{V}_{MZ}$ approaches unity, the ranges of possible parameter values shrink, becoming invariant when $C\hat{V}_{DZ} = C\hat{V}_{MZ}$, in which case $\hat{V}_G = \hat{V}_A = \hat{V}_{NA} = 0$ and all covariance between twins would be attributable to $\hat{V}_C$. Unfortunately, the reverse logic also holds: as $C\hat{V}_{DZ}/C\hat{V}_{MZ}$ decreases, the range in parameter values that could account for a given set of data becomes greater.

An advantage of modeling $\hat{V}_G$ (broad-sense heritability) rather than $\hat{V}_A$ and $\hat{V}_{NA}$ separately is that it is simpler and loses no information compared to the traditional approach in that $\hat{r}$, and the relative degree of $\hat{V}_A$ and $\hat{V}_{NA}$ are indeterminate anyway. The parameter $\hat{V}_G$

will also have much tighter confidence intervals than either $\hat{V}_A$ or $\hat{V}_{NA}$, and thus can allow for more powerful tests. Furthermore, it is easier to find and present the ranges of possible parameter values that arise from different values of $\hat{q}$, and interpretation of the quantity $\hat{q}$ brings researchers face-to-face with the very real problem of untangling genetic effects using the CTD. Nevertheless, in an important sense, modeling only broad-sense heritability merely pushes the issue of the makeup of genetic architecture back one level, but substantively it resolves none of the problems of parameter indeterminacy discussed above. The degree of indeterminacy in $\hat{V}_A$, $\hat{V}_{NA}$, and in $\hat{V}_C$, remains the same regardless of whether $\hat{V}_A$ and $\hat{V}_{NA}$ are modeled separately or combined into a single construct, $\hat{V}_G$.

Mathematical Validity Versus Biological Plausibility of Parameter Spaces

Table 1 shows the four boundaries of the parameter space that will arise from any CTD study. Any set of four parameters lying within this space will be mathematically equally valid in explaining a given data set. However, not every set of parameters within the space are biologically equally plausible. Specifically, values of $\hat{r}$ and $\hat{V}_A$ that are close to zero are less biologically plausible than spaces where these values are more positive. Equivalently, values of $\hat{q}$ that are close to zero are less biologically plausible than values of $\hat{q}$ that are closer to $^1\!/_2$. This is because dominance and higher-order epistatic effects tend to contribute variation to lower-order, simpler variance components while the reverse is not true. Two interacting alleles, which one might think should contribute variation only to V_2, nevertheless typically have marginal effects which also contribute variation to V_1 (i.e., to V_A); alleles whose effects are determined by three alleles have marginal effects that contribute to both V_1 and V_2; and so forth.[5] The exact degree of this 'carryover' to simpler variance components depends upon the strength of the interaction and allelic frequencies (see Eaves, 1988; Falconer & Mackay, 1996; Mather, 1974). For example, consid-

Table 3

Effects of Altering Assumed Values of $\hat{q}$ on Parameter Estimates Given Different Ratios of $C\hat{V}_{DZ}/C\hat{V}_{MZ}$

	$C\hat{V}_{DZ} = .60$ $C\hat{V}_{MZ} = .60$		$C\hat{V}_{DZ} = .40$ $C\hat{V}_{MZ} = .60$		$C\hat{V}_{DZ} = .20$ $C\hat{V}_{MZ} = .60$		$C\hat{V}_{DZ} = .00$ $C\hat{V}_{MZ} = .60$	
$\hat{q}$	$\hat{V}_G$	$\hat{V}_C$	$\hat{V}_G$	$\hat{V}_C$	$\hat{V}_G$	$\hat{V}_C$	$\hat{V}_G$	$\hat{V}_C$
.5	.00	.60	.40	.20	.80	−.20	1.20	−.60
.4	.00	.60	.33	.27	.67	−.07	1.00	−.40
.3	.00	.60	.29	.31	.57	.03	.86	−.26
.2	.00	.60	.25	.35	.50	.10	.75	−.15
.1	.00	.60	.22	.38	.44	.16	.67	−.07
.0	.00	.60	.20	.40	.40	.20	.60	.00
Range	.00		.20		.40		.60	

Note: Range is the absolute difference between the upper and lower bounds of both $\hat{V}_G$ and $\hat{V}_C$.

Variance estimates were not constrained to be ≤ 1 or ≥ 0 for symmetry of presentation. Altering these constraints does not qualitatively change the conclusions.

ering only single locus effects (ignoring epistasis), high levels of V_D (V_2) are only observed when either (a) numerous loci with infrequent and close to recessive alleles affect the trait; or (b) loci with frequent and over/underdominant alleles affect the trait. In most other cases, even if the alleles show various levels of dominance, the V_G will be mostly additive in nature.

A similar type of carryover effect exists with epistasis, although quantifying it becomes cumbersome (Mather, 1974) and thus simulation studies are a preferable approach. A simulation conducted by Eaves (1988) indicated that increasing the number of interacting loci had less effect on DZ twins' genetic covariances (q) than might be expected. As the number of interacting loci increased from two to five, q decreased by only .01 to .14 units, depending on (a) allele frequencies; (b) the strength and direction of dominance; and (c) the strength and direction of epistasis. With five loci, for no configuration of these factors was $q < 1/8$, and generally $.30 < q < .50$ (Eaves, 1988).

Using Eaves' findings as a rough guide, we might expect $1/8 < q < 1/2$ in most cases. Using standardized parameters (i.e., $V_G = 1$), from equation (17), this is equivalent to stating that $1/8 < (1/2 V_A + r V_{NA}) < 1/2$, which corresponds to $1/8 < r < 1/4$ and $V_A > 0$ in most cases. However, more work is needed to better understand what areas in the parameter space are more or less likely. Eaves' study surveyed two or five loci, but it is becoming clear that certain complex traits will be affected by many more loci than this. Similarly, his study surveyed many ranges of degrees of dominance, but some degrees of dominance appear to be more common than others (deVicente & Tanksley, 1993). Now that data regarding distributions of degrees of dominance, degrees of epistasis, and minor allele frequencies are increasingly available from QTL analyses in animals, such information can be used in simulation studies to gain a better understanding of what areas of the parameter space are highly unlikely and can therefore be safely ignored, and which areas are biologically more likely and should be given greater *a priori* weight. We leave further development on this topic for future research.

Presenting and Limiting Parameter Indeterminacy in Twin Studies

Current Reporting Practices of Classical Twin Design Findings

To gain an understanding of how parameter indeterminacy is typically reported, we reviewed 40 publications randomly selected from a larger pool of 241 publications (identified using the Medline search engine) of primary research that used the CTD or the *twin plus siblings* design.[6] Of the 40 publications, 38 used the CTD and two used the twin plus siblings design. We found that 33 (82%) of the reviewed publications made no mention of the potential trade-off between $\hat{V}_C$ and $\hat{V}_{NA}$ nor the ramifications of fixing $\hat{r} = 1/4$. The coverage on indeterminacy in the remaining 18% (seven publications) was comprised of the following: two that did not report directly on the indeterminacy but noted limitations of estimating parameters with the CTD; three that conveyed it was possible for both V_C and V_{NA} to coexist, but that only one could be estimated using the CTD; one that reported the potential for $\hat{V}_A$ to overestimate V_A; and finally, one that provided a rather comprehensive account of indeterminacy. In all cases, parameter estimates were derived from the specific assumption that $\hat{r} = 1/4$ and that either $\hat{V}_C = 0$ or $\hat{V}_{NA} = 0$, and in no case was the degree of parameter indeterminacy reported explicitly.[7]

We question whether the typical reporting practice from twin results — reporting only parameter estimates that arise from one particular set of assumptions — is optimal given that these assumptions are predicted to lead to biased estimates that might thereby convey somewhat false impressions of the causes of human variation. Among other reasons, this is potentially troubling because findings from twin studies are increasingly used by researchers outside of mainstream behavioral genetics as a way to understand the causes of variation in traits of interest. Researchers conducting linkage analysis, for instance, might use twin findings to understand not only how 'genetic' a trait is, but also to understand the degree of genetic nonadditivity present, which could have ramifications on the type of statistical model employed. Similarly, developmental psychologists might interpret evidence for common environmental effects without understanding how its estimate depends upon two often hidden model assumptions.

Presenting Parameter Indeterminacy

We forward that it is important to convey to consumers of twin research the range of possible parameter values that could account for a given twin data set. For the CTD, parameter indeterminacy arising in twin studies can be presented graphically by running the analysis (using LISREL, Mx, etc.) four different times (corresponding to the rows 1 to 4 in Table 1 when $C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq 1/2$ and rows 5 to 8 when $C\hat{V}_{DZ}/C\hat{V}_{MZ} \leq 1/2$), with each run fixing $\hat{r}$ and one of the other parameters. The remaining two parameters from each run are freely estimated and their values recorded. As each run leads to a minimum or maximum for at least one of the parameters, the four sets of two fixed and two estimated parameters represent the boundaries of the parameter space which can then be graphed. That any set of four parameters within this area is equally likely (statistically) can be checked by choosing a point within the area, fixing the four parameters to the corresponding values, and checking that the fit function is exactly the same as the one obtained from the CTD.

Because indeterminacy exists in two dimensions — not four — the parameter space can be fully presented graphically in two dimensions, although some information must be condensed.[8] Figure 2 shows the two-dimensional standardized parameter space from a

twin plus sibling study (Keller et al., 2005) on extraversion where $C\hat{V}_{DZ}/C\hat{V}_{MZ} < 1/2$, and Figure 3 shows the standardized parameters space from data on perceived social support using a CTD (Coventry et al., 2005) where $C\hat{V}_{DZ}/C\hat{V}_{MZ} > 1/2$. The y-axes are $\hat{V}_C$, the x-axes are the degree to which $\hat{V}_A$ is greater than $\hat{V}_{NA}$, and $\hat{r}$ is represented by shading. Parameter estimates from the CTD are single points on the extreme boundaries of the spaces, showing graphically that the assumptions of the CTD lead to estimates that are biased in the predictable directions discussed above. The parameter space in Figure 3 is triangular because there is no difference between $\hat{r} = 0$ and $\hat{r} = 1/4$ when $\hat{V}_{NA} = 0$.

In both figures, the lower limits for $\hat{r}$ and $\hat{V}_A$ were set to zero, which is probably unrealistic, as described above. Thus, it must be stressed that the parameter space presented in these figures shows the full range of mathematically equally valid parameters. Darker areas of the figure (where $1/8 < r < 1/4$) and areas where $\hat{V}_A$ is more positive (perhaps where $\hat{V}_A > 1/2\hat{V}_{NA}$) should be considered more biologically likely. In future research, we plan on more precisely identifying the areas that are more likely, and introducing a presentation technique that limits the parameter space to those regions that are biologically plausible.

The graphical approach employed here can be extended to multivariate phenotypes by fixing parameters in the same way as with univariate phenotypes.

Furthermore, if a parameter interacts with sex or some environmental variable, the parameter space can be separately analyzed at different levels of the interacting variable (e.g., different graphs for males and females if there is evidence for a gene-by-sex interaction).

The ranges of $\hat{V}_A$ and $\hat{V}_{NA}$ in Figures 2 and 3 demonstrate that the CTD is poorly suited to divining the genetic architectures of traits. Nevertheless, it should also be noted that the range of the sum $\hat{V}_A + \hat{V}_{NA}$ ($\hat{V}_G$) will always be much less than the range of the difference $\hat{V}_A - \hat{V}_{NA}$. While $\hat{V}_A - \hat{V}_{NA}$ has a range of .62 (–.39 to .23) in Figure 2 and .47 (–.19 to .28) in Figure 3, $\hat{V}_G$ has a range of .18 (.29 to .47) in Figure 2 and .09 (.19 to .28) in Figure 3. This graphically demonstrates the fact that broad-sense heritability estimates will always be less variable than narrow-sense heritability estimates from the CTD (Martin et al., 1978).

Limiting Parameter Indeterminacy

We forward three nonmutually exclusive strategies for limiting the degree of parameter indeterminacy in twin designs. The first is to extend, whenever possible, the CTD to include parents, siblings, spouses and children (Truett et al., 1994). Though more complicated and computationally intensive, the ETFD has two principal advantages over the CTD. First, by adding more genetically informative relatives, $\hat{V}_C$ can be decomposed into four components: variance due to parental influence (vertical transmission), variance specific to the
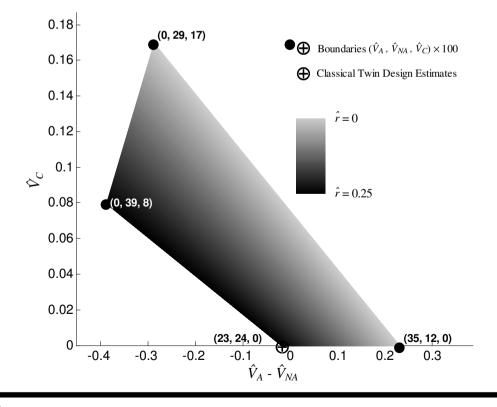

Figure 2

Space of mathematically equally valid parameter values given $C\hat{V}_{MZ} = .47$ and $C\hat{V}_{DZ} = .17$. Areas where $1/8 < r < 1/4$ and $V_A > 0$ may be more biologically plausible (see text).
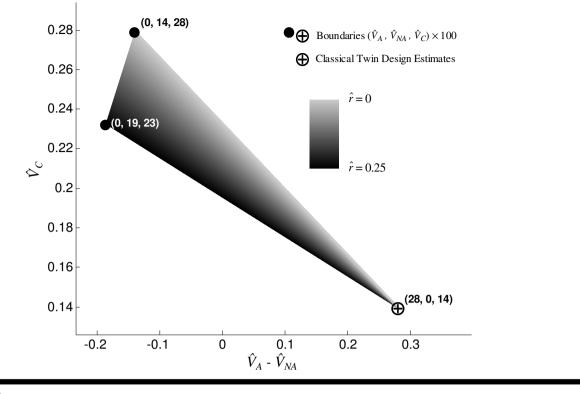
Figure 3

Space of mathematically equally valid parameter values given $C\hat{V}_{MZ} = .42$ and $C\hat{V}_{DZ} = .28$.

twin environment, variance specific to the sibling environment, and (genetic) variance due to assortative mating. Second, the ETFD allows V_C and V_{NA} to be estimated simultaneously, precluding the need to fix one of these parameters to zero. Nevertheless, parameter indeterminacy still persists in the ETFD, albeit at lower levels, because the value of r remains inestimable. When $r = 1/4$, the variance components in the ETFD are unbiased. However, to the degree that $r < 1/4$, V_{NA} will be overestimated and V_C will be underestimated. Because of this, it seems preferable to explicitly report the parameter indeterminacy caused by varying $\hat{r}$ in ETFD studies.

A second possible way to limit parameter indeterminacy in twin designs is to report parameters from the *full* rather than *reduced* statistical models. While it is traditional to drop parameters (i.e., set them to zero) that statistically cannot be shown to differ from zero, such an approach probably maximizes the biases of the estimated variance parameters. As demonstrated above and shown graphically in Figures 2 and 3, setting parameters to their lower bound of zero pushes the estimated parameters to extreme boundaries of the space of their possible values. In a Bayesian sense, this should maximize on average the distance from the estimated parameters to the true parameter values. Nevertheless, reduced models may sometimes be preferable simply because they are more powerful and make more easily falsified predictions. It seems preferable therefore to report parameters from

both the full and reduced statistical models, leaving interpretation to the reader (M. C. Neale, personal communication, February, 2005).

A third possible way to narrow the parameter indeterminacy in twin designs is to use prior research or theory to fix $\hat{V}_C$ (or other parameters) at more likely values. For example, if adoption data exists for a phenotype of interest that allows for reliable estimates of V_C, $\hat{V}_C$ can be fixed at that value in future classical twin studies of the same phenotype. Stoolmiller (1999) and others have argued that adoption data may often underestimate V_C because the range of family environments are restricted due to screening by adoption agencies. However, V_C might also be overestimated due to selective placement, and in either case, statistical methods can be used to control for these sources of bias (Neale, 1998). While we agree with those that would caution that this strategy should be used only when there is good reason to believe that the fixed values are fairly accurate and relevant to the study in question, it must also be acknowledged that the traditional method of fixing parameters (e.g., fixing $\hat{V}_C = 0$ simply because $C\hat{V}_{DZ}/C\hat{V}_{MZ} \leq 1/2$) is not only more arbitrary than the proposed method, it also leads to extreme parameter estimates.

Conclusion

How substantial are the biases in parameter estimates from CTDs? The maximal *potential* bias can be quite

large, depending upon the ratio $C\hat{V}_{DZ}/C\hat{V}_{MZ}$ and the absolute level of $C\hat{V}_{MZ}$ (Table 2). In a companion report (Coventry & Keller, 2005), we compared all ETFD parameter estimates available in the literature to those that would be obtained from the CTD as a way to gain some idea of the *actual* bias in CTD estimates that is due to the inability to simultaneously estimate $\hat{V}_C$ and $\hat{V}_{NA}$, as well as several other parameters. As predicted, we found that $\hat{V}_A$ and $\hat{V}_{NA}$ parameters from the CTD could be very misleading, although broad-sense heritability ($\hat{V}_A + \hat{V}_{NA}$) appeared to be much less biased. Furthermore, for certain phenotypes, what would appear to be $\hat{V}_C$ in the CTD turned out to be due to assortative mating in the ETFD, demonstrating that CTD studies do not always underestimate the effects of the common environment. For other phenotypes, the effects of the common environment were masked in CTD studies. These results demonstrate the problems described in this paper with interpreting CTD parameter estimates, but they also suggest that CTD estimates are useful if interpreted in light of their limitations.

The degree to which epistasis contributes variation to traits and the consequent bias it introduces (due to the indeterminacy in *r*) poses a more difficult problem for the interpretation of parameter estimates from twin designs. Epistasis disproportionately masks the effects of the common environment and assortative mating, and drives DZ similarity to much lower levels than MZ similarity, which expands the amount of indeterminacy in estimated parameters. While its potential to bias parameter estimates tends to be lower than the trade-off between $\hat{V}_C$ versus $\hat{V}_{NA}$, the true bias it causes cannot be gauged by comparing CTD and ETFD estimates because is inestimable in the ETFD. It has become clear from both genetic mapping and animal studies that epistasis cannot be ignored as an important factor in the genetic architecture of traits (Carlborg & Haley, 2004), and the common observations of $C\hat{V}_{DZ}/C\hat{V}_{MZ} < 1/2$ and even $C\hat{V}_{DZ}/C\hat{V}_{MZ} < 1/4$ are consistent with this possibility (Lykken et al., 1992).

We have identified several ways that parameter indeterminacy from twin data can be minimized and/or communicated. Most centrally, we argue that reporting the full space of possible parameters, as in Figures 1 and 2, is easily accomplished, simple to understand, and aids in the proper interpretation of both CTD and ETFD results. This approach is preferable, we argue, to the current practice of reporting a single set of parameters based on only one particular set of assumptions. A clear description of parameter indeterminacy is not only important for accuracy and scientific integrity, but also because it keeps behavioral geneticists themselves mindful of assumptions that can easily go unquestioned.

This paper is not a critique of the CTD nor of twin designs more generally, which have proven over the years to be extraordinarily useful. Indeed, the CTD represents a large improvement over epidemiological studies that include, for instance, parents and offspring

only. Moreover, the issue of parameter indeterminacy is not a problem unique to twin studies; any model in science is open to alternative explanations, and it is a strength of the twin design that these alternatives can be quantified (compare this to many models in the social sciences). However, as Eaves et al. (1978) noted (referring specifically to the CTD), 'If the twin study is to have a future in human quantitative genetics, it will no longer be as the mainstay of an investigation of the "heritability" of a heterogeneous collection of behavioral measures … twins are not an end in themselves since the twin study cannot constitute a definitive basis for genetic analysis' (p. 277). Twenty-five years after this was written, countless studies have employed the CTD. While the researchers conducting these studies have not asserted that their investigations represent the definitive solution, they quite often do not acknowledge the potential biases caused by the model's assumptions, nor the ranges of possible parameter values that could account for their data. Presumably this is assumed knowledge, but such an assumption, like the assumptions in the twin method itself, should be critically examined.

Endnotes

1   We follow the convention that $\hat{V}_O$ and $\hat{r}$ are the estimates (via observation of a sample or deduction) of the unknown population parameters V_O and r.

2   See Mather (1974) for definitions of the various types of epistasis and for alternative methods of variance decomposition due to genetic nonadditivity. The simplified technique presented here applies only to siblings; covariation between parents and offspring depend not only upon the number of modifying alleles, but also upon whether the alleles are at the same or different loci.

3   Grayson (1989) concluded that $\hat{V}_A$ always overestimates V_A by a factor of $(2^{(1-r)}/1-2r)\hat{V}_C$ and V_{NA} always underestimates $\hat{V}_{NA}$ by a factor of $(1/1-2r)V_C$. The differences between Grayson's conclusions and our own appear to be due to incorrect substitution, not simply to alternative assumptions or perspectives. A fuller explanation of this discrepancy and worked-through examples are available from the authors upon request.

4   Each parameter has a natural lower bound (zero), but because *r* also has an upper bound ($1/4$), it makes sense to always fix $\hat{r}$ at its upper and lower limits when exploring the boundaries of the other parameters. Note that if $C\hat{V}_{DZ}/C\hat{V}_{MZ} = \hat{m} < 1/4$, the upper bound of $\hat{r}$ should be $\hat{m}$ rather than $1/4$ to insure that no variance parameters are negative. Observing that $C\hat{V}_{DZ}/C\hat{V}_{MZ} < 1/4$ would also imply that at least some higher-order epistasis affects the trait and that $V_3 … V_n \neq 0$.

5   This occurs for the same reason that some marginal effect generally exists anytime two or more variables interact.

6   Designs that include only twins plus siblings are also sometimes referred to as *extended twin designs*. Although twin plus sibling designs have substantially more power to detect genetic nonadditivity and common

environmental effects (Posthuma & Boomsma, 2000), they suffer from the same parameter indeterminacy issues as does the CTD.

7   Note that parameter indeterminacy is a separate issue from variation in parameter estimates arising from sampling error (i.e., reporting confidence intervals for variance parameters). The latter were commonly reported while the former were not.

8   MATLAB scripts that graph the parameters space for any CTD findings (such as Figures 2 and 3) are available for download at www.matthewckeller.com. The scripts require parameter boundaries as input and output a plot.

Acknowledgment

References

Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics, 3*, 872–882.

Borkenau, P., Riemann, R., Angleitner, A., & Spinath, F. M. (2002). Similarity of childhood experiences and personality resemblances in monozygotic and dizygotic twins: A test of the equal environment assumption. *Personality and Individual Differences, 33*, 261–269.

Carlborg, O., & Haley, C. S. (2004). Epistasis: Too often neglected in complex trait studies? *Nature Reviews Genetics, 5*, 618–625.

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

Coventry, W. L., Medland, S. E., Gillespie, N. A., Heath, A. C., & Martin, N. G. (2005). *Genetic and environmental influences on perceived social support: Differences across sex and the member of the support network.* Unpublished manuscript.

DeVicente, M. C., & Tanksley, S. D. (1993). QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics, 134*, 585–596.

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

Eaves, L. J., Last, K. A., Young, P. A., & Martin, N. G. (1978). Model-fitting approaches to the analysis of human behavior. *Heredity, 41*, 249–320.

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics* (4th ed.). Harlow, UK: Addison Wesley Longman.

Grayson, D. A. (1989). Twins reared together: Minimizing shared environmental effects. *Behavior Genetics, 19*, 593–604.

Heath, A. C., Martin, N. G., Eaves, L. J., & Loesch, D. (1984). Evidence for polygenic epistatic interactions in man? *Genetics, 106*, 719–727.

Jinks, J. L., & Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA and classical approaches to the analysis of human behavior. *Psychological Bulletin, 73*, 311–349.

Keller, M. C., Coventry, W. L., Heath, A. C., & Martin, G. M. (in press). Ubiquitous evidence for genetic non-additivity in Cloninger's and Eysenck's Personality Dimensions in twin sibships. *Behaviour Genetics.*

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

Li, Z., Pinson, S. R., Park, W. D., Paterson, A. H., & Stansel, J. W. (1997). Epistasis for three grain yield components in rice (*Oryza sativa*). *Genetics, 145*, 453–465.

Long, A. D., Mullaney, S. L., Reid, L. A., Fry, J. D., Langley, C. H., & Mackay, T. F. C. (1995). High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. *Genetics, 139*, 1273–1291.

Lykken, D. T., McGue, M., Tellegen, A., & Bouchard, T. J. (1992). Emergenesis: Genetic traits that may not run in families. *American Psychologist, 47*, 1565–1577.

Martin, N. G., Eaves, L. J., Kearsey, M. J., & Davies, P. (1978). The power of the classical twin study. *Heredity, 28*, 79–95.

Mather, K. (1974). Non-allelic interaction in continuous variation of randomly breeding populations. *Heredity, 32*, 414–419.

Neale, M. C. (1998). Adoption studies. In P. Armitage & T. Colton (Eds.), *Encyclopedia of Biostatistics* (pp. 77–81). New York: John Wiley.

Neale, M. C., & Maes, H. H. (2004). *Methodology for genetic studies of twins and families*. Dordrecht, the Netherlands: Kluwer Academic.

Posthuma, D., & Boomsma, D. I. (2000). A note on the statistical power in extended twin designs. *Behavior Genetics, 30*, 147–158.

Shook, D. R., & Johnson, T. E. (1999). Quantitative trait loci affecting survival and fertility-related traits in *Caenorhabditis elegans* show genotype-environment interactions, pleiotropy and epistasis. *Genetics, 153*, 1233–1243.

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

Truett, K. R., Eaves, L. J., Walters, E. E., Heath, A. C., Hewitt, J. K., Meyer, J. M., Silberg, J., Neale, M. C., Martin, N. G., & Kendler, K. S. (1994). A model system for analysis of family resemblance in extended kinships of twins. *Behavior Genetics, 24,* 35–49.

Yu, S. B., Li, J. X., Tan, Y. F., Gao, Y. J., Li, X. H., Zhang, Q., & Maroof, M. A. (1997). Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences USA, 94,* 9226–9231.