Estimating the Extent of Parameter Bias in the Classical Twin Design: A Comparison of Parameter Estimates From Extended Twin-Family and Classical Twin Designs

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

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

The *classical twin design* (CTD) circumvents parameter indeterminacy by assuming (1) negligible higher-order epistasis; and (2) either nonadditive genetic or common environmental effects are nonexistent, creating two potential sources of bias (Eaves et al., 1978; Grayson, 1989). Because the *extended twin-family design* (ETFD) uses many more unique covariance observations to estimate parameters, common environmental and nonadditive genetic parameters can be simultaneously estimated. The ETFD thereby corrects for what is likely to be the largest of the two sources of bias in CTD parameter estimates (Keller & Coventry, 2005). In the current paper, we assess the extent of this and other potential sources of bias in the CTD by comparing all published ETFD parameter estimates to CTD parameter estimates derived from the same data. CTD estimates of the common environment were lower than ETFD estimates of the common environment for some phenotypes, but for other phenotypes (e.g., stature in females and certain social attitudes), what appeared as the common environment was resolved to be assortative mating in the ETFD. On average, CTD estimates of nonadditive genetic factors were 43% lower, and additive genetic factors 63% higher, than ETFD estimates. However, broad-sense heritability estimates from the CTD were only 18% higher than ETFD estimates, highlighting that the CTD is useful for estimating broad-sense but not narrow-sense heritability. These results suggest that CTD estimates can be misleading when interpreted literally, but useful, albeit coarse, when interpreted properly.

The *classical twin design* (CTD) is the most commonly used technique to infer genetic and environmental causes of phenotypic variance. It compares the similarity of MZ (monozygotic) twins to DZ (dizygotic) twins in order to estimate parameters of additive genetic (V_A), nonadditive genetic (V_{NA}), common environmental (V_C), and unique environmental (V_E) variation. In addition to these four parameters, the correlation between DZ twins due to genetic nonadditivity, r, is also unknown, although it must range between ¼ (when all genetic nonadditivity is due to additive-by-additive epistasis or dominance) to a theoretical lower limit of 0 (when genetic nonadditivity is due to extreme epistasis). However, there are possibly good reasons, based on the principles of biometrical genetics, to believe that r is typically closer to ¼ than to 0 (Eaves et al., 1978; Keller & Coventry, 2005).

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). A primary limitation of the CTD stems from the fact that it offers only three observations, the total phenotypic variance ($\hat{V}_P$), the covariance between MZ twins ($C\hat{V}_{MZ}$), and the covariance between DZ twins ($C\hat{V}_{DZ}$), from which to estimate five unknown parameters[1]. Algebraically, five unknowns cannot be estimated from only three knowns. This limitation is circumvented in the CTD by fixing two of the unknown parameters at assumed values ($\hat{r}$ to ¼ and either $\hat{V}_C$ or $\hat{V}_{NA}$ to 0), enabling the other three parameters to be estimated.

In a companion report (Keller & Coventry, 2005), we derive the biases in the three estimated parameters that result from fixing (a) $\hat{r}$ to ¼ and (b) either $\hat{V}_C$ or $\hat{V}_{NA}$ to 0 (see also Grayson, 1989). When $\hat{V}_{NA}$ is set to 0 (often called 'ACE' models), V_A tends to be overestimated while V_{NA} (because it is set to 0) and V_C tend to be underestimated. While when $\hat{V}_C$ is set to 0 (often called 'ADE' models), depending on the value of $\hat{r}$, V_A is usually (but not always) overestimated, V_{NA} is usually (but not always) underestimated, and V_C (because it is set to 0) tends to be underestimated (cf. Grayson, 1989). Of the two sources of bias, we found

that the inability to simultaneously estimate both V_{NA} and V_C, as opposed to the fixing of $\hat{r}$ to ¼, generally has greater potential to bias parameter estimates (Keller & Coventry, 2005).

Fortunately, V_{NA} and V_C can be simultaneously estimated using another type of twin design, the *extended twin-family design* (ETFD; Truett et al., 1994). By including not only MZ and DZ twins, but also their parents, nontwin siblings, spouses, and children, the ETFD increases the number of unique covariance observations to 80. For example, one such covariance observation is between the children of two MZ co-twins, who are genetically related by 50% but socially are cousins. The drastic increase in information in the ETFD allows not only for the simultaneous estimation of both V_{NA} and V_C, but also allows V_C to be disaggregated into a component due to parental influence ('vertical transmission'), a component due to living with a co-twin, and a component that is shared between any two siblings (the residual common environment).

The ETFD has two additional important advantages: effects of genotype–environment covariation (CV_{GE}) and assortative mating ('like marrying like'; V_{AM}) are not confounded with estimates of V_C as they are in the CTD. Genotype–environment covariation occurs when parental phenotypes, which are partly genetic in origin, modify their offsprings' environment. For example, intelligent parents providing not only more 'intelligence genes' to their offspring, but also enriched home environments. The CTD assumes no genotype–environment covariation and as a consequence, such effects mimic $\hat{V}_C$ in the CTD (Eaves et al., 1989; Purcell, 2002; Tambs et al., 1992).[2] Assortative mating tends to increase the genetic covariance between DZ twins as well as other siblings on a trait given that a) the trait is genetically influenced; and b) the correlation is due to initial assortment rather than convergence between spouses over time (Eaves et al., 1978). Because an increase in genetic covariation cannot occur between MZ twins, the genetic effects of assortative mating inflate DZ similarity relative to MZ similarity and thus mimic $\hat{V}_C$ in the CTD.[3]

In our companion report, we noted that assortative mating mimics $\hat{V}_C$, which is generally the case,[3] and we did not address effects of genotype–environment covariation, which also tend to mimic $\hat{V}_C$. Because both assortative mating and gene–environment covariation, as measured by the ETFD, are factors that increase DZ similarity relative to MZ similarity, then the treatment of biases in our companion report are essentially correct, so long as the parameter 'V_C' in our companion report is interpreted as including the effects of the common environment, assortative mating, and gene–environment covariation. However, we can disaggregate these three components to more accurately illustrate how each effect unfolds to bias CTD parameters. In summary, assuming that r = ¼, in ADE models, $\hat{V}_A$ is overestimated by $3V_C$ and $\hat{V}_{NA}$ is underestimated by $2V_C$ (Keller & Coventry, 2005), with both biases

exaggerated by assortative mating and genotype–environment covariation. In ACE models, $\hat{V}_A$ is overestimated by $1½V_{NA}$ (Keller & Coventry, 2005), but this bias is attenuated by assortative mating. Similarly, in ACE models, $\hat{V}_C$ is underestimated by $½V_{NA}$ (Keller & Coventry, 2005), but because assortative mating and genotype–environment covariation mimic V_C in the CTD, these factors serve to bias $\hat{V}_C$ back upwards, such that on balance, it is difficult to know whether $\hat{V}_C$ is underestimated or overestimated in the CTD. Particularly in the presence of high assortative mating, V_C may actually be overestimated in the CTD (worked from the $C\hat{V}_{MZ}$ and $C\hat{V}_{DZ}$ equations presented in Eaves et al., 1989; Tambs et al., 1992).

The purpose of the present paper is to empirically assess the degree of bias in CTD parameter estimates caused by 1) its inability to simultaneously estimate both V_{NA} and V_C; 2) assuming that mating is random; and 3) ignoring genotype–environment covariation. Because ETFD parameter estimates are not biased by these three factors, comparing CTD and ETFD parameter estimates offers one way to isolate the effects of these biases on CTD estimates.

We do not argue, in our comparison against the ETFD, that this design is infallible or that this method of estimating the bias in CTD estimates is perfect. Cohort effects, for example, can decrease correlations between nontwins and thus might inflate nonadditive genetic effects in the ETFD. Moreover, like the CTD, the ETFD does not allow a way to estimate r (both designs traditionally fix $\hat{r}$ = ¼), and thus we cannot gauge the effects of this source of bias by comparing estimates from the two designs.[4] Nevertheless, the present method allows a way to gauge what are likely to be major sources of bias in the CTD (Keller & Coventry, 2005).

Method

Sample of ETFD Phenotypes and Parameter Estimates

We include seven ETFD studies that report on a total of 17 phenotypes, representing nearly all ETFD papers published,[5] (see Table 1). Most of the seven studies presented different solutions for males versus females (Eaves, Heath, Martin, Maes, et al., 1999; Eaves, Heath, Martin, Neale, et al., 1999; Lake et al., 2000; Maes et al., 1997; Truett et al., 1994) and two presented different solutions for different data sets (Kendler et al., 1994; 1995), leading to a total of 38 phenotypes available for comparing ETFD and CTD parameter estimates.

Derivation of CTD Parameter Estimates

We derived the CTD estimates by fitting sex-limited ACE and ADE models in Mx (Neale, 2004) to the five distinct standardized covariances (i.e., correlations) of MZ and DZ twin pairs reported in the ETFD papers. Following convention, we fit ACE models when $C\hat{V}_{DZ}/C\hat{V}_{MZ} \geq$ ½ and ADE models when $C\hat{V}_{DZ}/C\hat{V}_{MZ} \leq$ ½. We fitted the common effects sex limitation model, which constrains the male–female
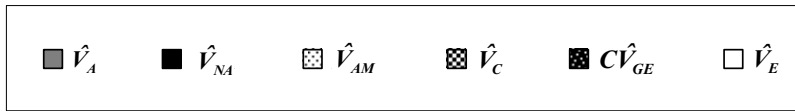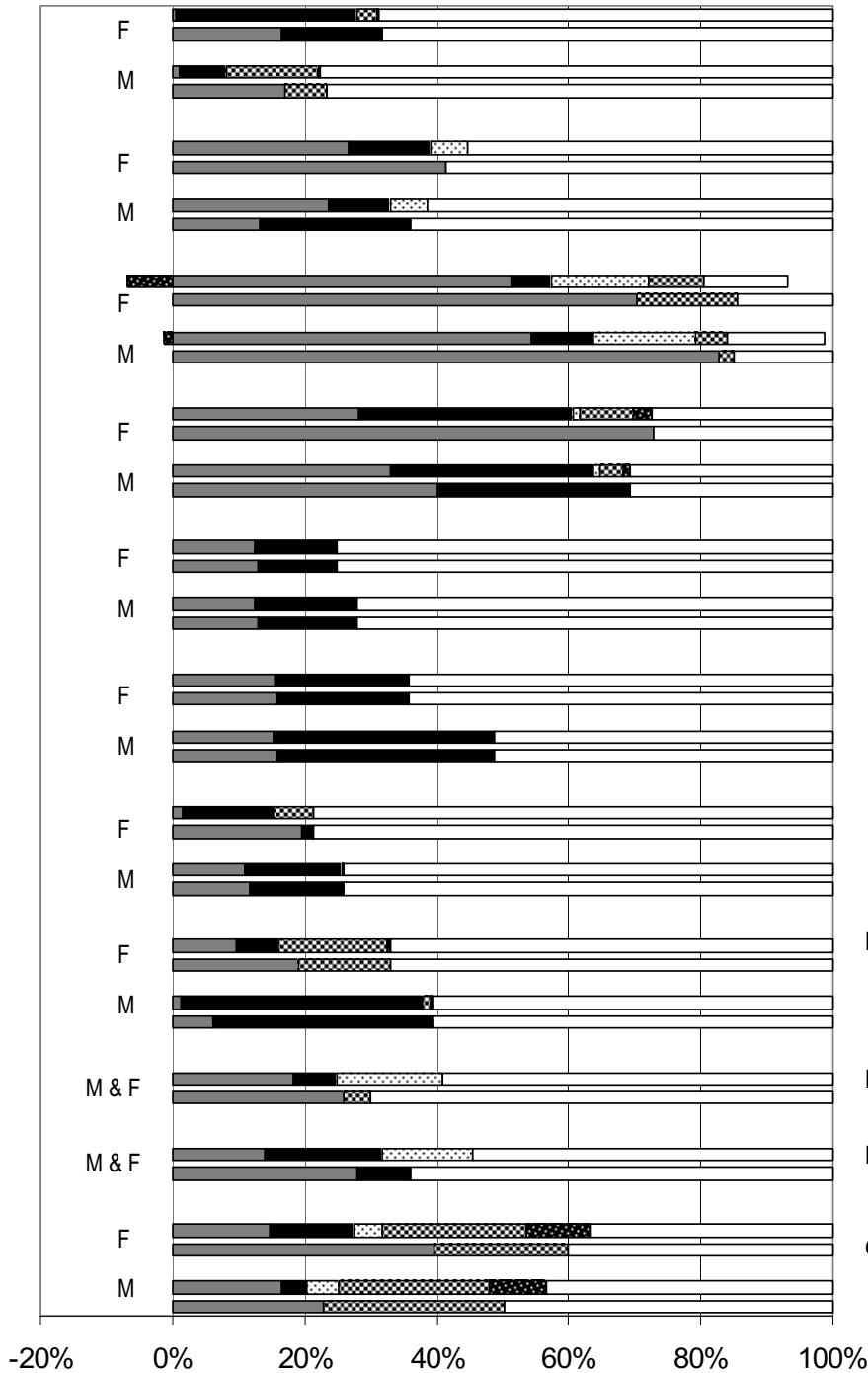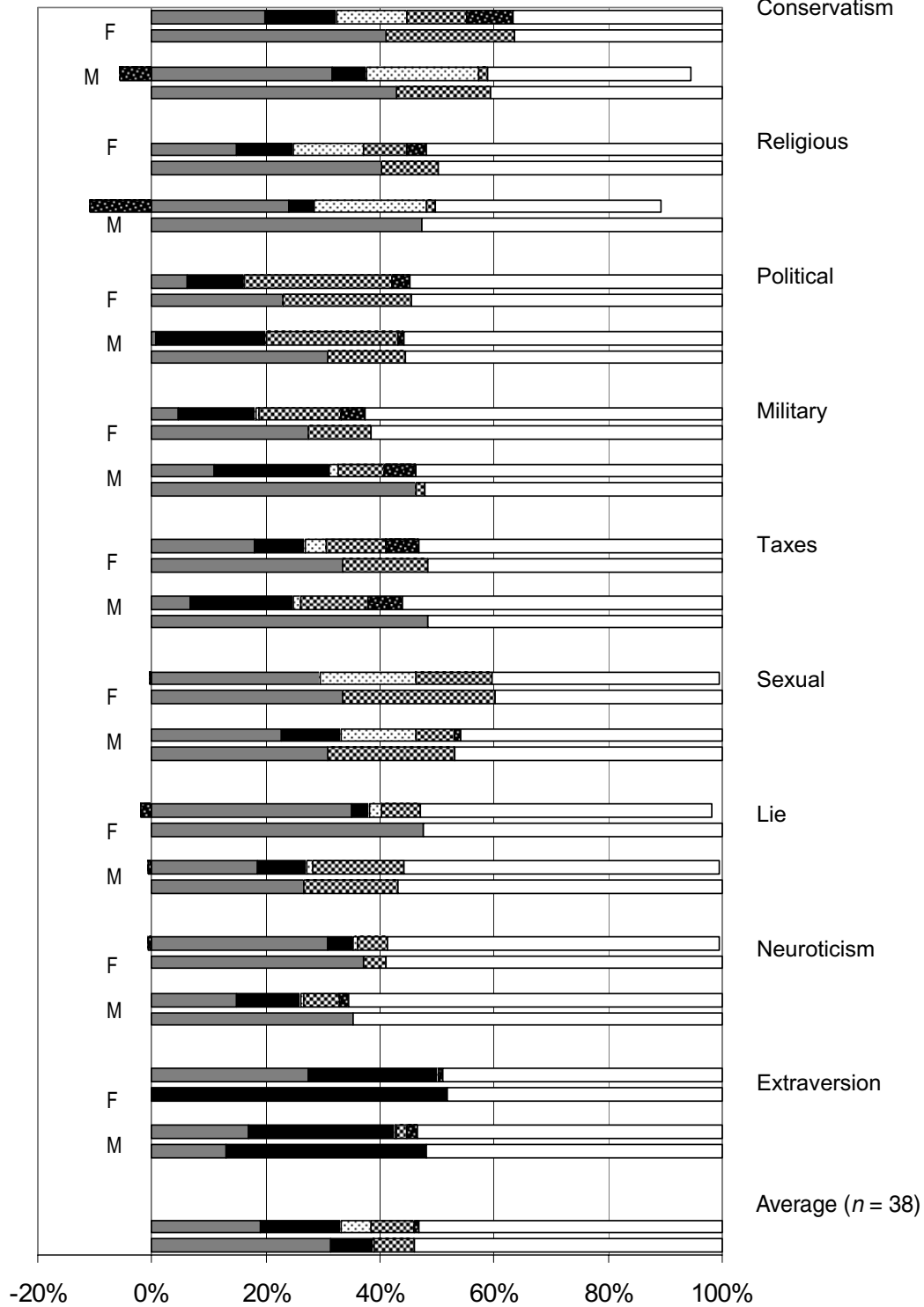
Figure 1

Comparison of the proportion of total variance (%) for published extended twin-family design (ETFD) estimates (higher of the paired bars) versus classical twin design (CTD) estimates (lower of the paired bars). On occasion, the G–E covariance estimates were negative (bars to the left of 0%).

Table 1

Standardized Estimates for Published Extended Twin-Family Designs (ETFD) Relative to ACE or ADE Models of the Classical Twin Design (CTD)

| Authors (data) Phenotype	Sex	Model	$\hat{V}_A$	$\hat{V}_{NA}$	$\hat{V}_G$	$\hat{V}_C$	$C\hat{V}_{GE}$	$\hat{V}_{AM}$	$\hat{V}_E$
Truett et al. 1994 (AARP & Virginia data)									
Church attendance	M	ETFD	.165	.039	.204	.228	.087	.048	.434
		ACE	.230		.230	.272			.498
	F	ETFD	.148	.127	.275	.218	.096	.043	.367
		ACE	.396		.396	.204			.400
Kendler et al. 1994 (AARP & Virginia data)[1]									
Depressive symptoms	M & F	ETFD	.140	.177	.316			.138	.546
(AARP data)		ADE	.280	.080	.360				.640
Depressive symptoms	M & F	ETFD	.183	.066	.249			.160	.591
(Virginia data)		ACE	.260		.260	.040			.700
Kendler et al. 1995 (AARP & Virginia data)[1]									
Panic-Phobia	M	ETFD	.013	.365	.378	.012	.004	.001	.606
(AARP data)		ADE	.061	.333	.394				.606
	F	ETFD	.097	.063	.160	.165	.006	.000	.668
		ACE	.191		.191	.140			.669
Panic-Phobia	M	ETFD	.111	.147	.258	.001	.000	.001	.743
(Virginia data)		ADE	.117	.141	.259				.741
	F	ETFD	.016	.137	.153	.060	.000	.000	.787
		ADE	.196	.017	.213				.787
Somatization	M	ETFD	.154	.334	.488		.000	.001	.511
(AARP data)		ADE	.158	.329	.488				.512
	F	ETFD	.154	.204	.358		.000	.000	.642
		ADE	.158	.200	.358				.642
Somatization	M	ETFD	.125	.154	.278		.000	.002	.720
(Virginia data)		ADE	.130	.148	.278				.722
	F	ETFD	.125	.125	.250		.000	.000	.749
		ADE	.131	.119	.250				.750
Maes et al. 1997 (AARP & Virginia data)[1]									
BMI	M	ETFD	.390	.270	.660	.030	.010	.010	.270
		ADE	.400	.293	.693				.307
	F	ETFD	.280	.320	.600	.080	.030	.010	.270
		ADE	.730	.000	.730				.270
Eaves, Heath, Martin, Maes, et al. 1999 (AARP & Virginia data)									
Stature	M	ETFD	.558	.094	.652	.047	−.012	.161	.153
		ACE	.827		.827	.024			.149
	F	ETFD	.596	.069	.665	.096	−.079	.172	.149
		ACE	.703		.703	.152			.145
Lake et al. 2000 (ATR, AARP & Virginia data)[1]									
Neuroticism	M	ETFD	.236	.094	.330			.057	.613
		ADE	.132	.227	.360				.640
	F	ETFD	.264	.123	.387			.057	.547
		ADE	.414	.000	.414				.586

Table 1 (continued)

Standardized Estimates for Published Extended Twin-Family Designs (ETFD) Relative to ACE or ADE Models of the Classical Twin Design (CTD)

Authors (data) Phenotype	Sex	Model	$\hat{V}_A$	$\hat{V}_{NA}$	$\hat{V}_G$	$\hat{V}_C$	$C\hat{V}_{GE}$	$\hat{V}_{AM}$	$\hat{V}_E$
Eaves, Heath, Martin, Neale, et al. 1999 (AARP & Virginia data)									
Psychoticism[2]	M	ETFD	.010	.072	.082	.140	.001	.000	.776
		ACE	.170		.170	.063			.767
	F	ETFD	.006	.270	.276	.030	.002	.000	.679
		ADE	.165	.153	.318				.682
Extraversion[2]	M	ETFD	.169	.258	.427	.019	.019	.001	.534
		ADE	.131	.349	.480				.520
	F	ETFD	.274	.228	.502	.000	.006	.002	.490
		ADE	.000	.519	.519				.481
Neuroticism[2]	M	ETFD	.149	.113	.262	.065	.015	.003	.654
		ACE	.352		.352	.001			.647
	F	ETFD	.313	.044	.357	.054	−.005	.007	.588
		ACE	.372		.372	.038			.590
Lie[2]	M	ETFD	.188	.087	.275	.162	−.005	.010	.557
		ACE	.267		.267	.164			.569
	F	ETFD	.362	.035	.397	.070	−.019	.020	.531
		ACE	.475		.475	.000			.525
Sexual[3]	M	ETFD	.226	.105	.331	.067	.012	.129	.455
		ACE	.307		.307	.224			.469
	F	ETFD	.297	.000	.297	.134	−.004	.169	.404
		ACE	.334		.334	.268			.398
Taxes[3]	M	ETFD	.070	.187	.257	.123	.063	.015	.582
		ACE	.485		.485	.000			.515
	F	ETFD	.180	.088	.194	.106	.057	.038	.532
		ACE	.334		.334	.150			.516
Military[3]	M	ETFD	.110	.202	.312	.080	.056	.015	.537
		ACE	.464		.464	.014			.522
	F	ETFD	.047	.135	.182	.145	.040	.006	.627
		ACE	.274		.274	.110			.616
Political[3]	M	ETFD	.006	.195	.201	.231	.011	.000	.557
		ACE	.307		.307	.137			.556
	F	ETFD	.062	.099	.161	.258	.032	.002	.549
		ACE	.230		.230	.224			.546
Religious[3]	M	ETFD	.307	.056	.363	.017	−.139	.254	.506
		ADE	.473	.000	.473				.527
	F	ETFD	.148	.099	.247	.078	.033	.123	.519
		ACE	.404		.404	.099			.498
Conservatism	M	ETFD	.355	.067	.422	.016	−.062	.222	.401
		ACE	.428		.428	.165			.407
	F	ETFD	.198	.125	.323	.106	.081	.124	.366
		ACE	.410		.410	.229			.363

Note: Values omitted were fixed to zero; AARP = American Association of Retired Persons (referred to as 'Virginia 30,000' when combined with the Virginia data); ATR = Australian NHMRC Twin Register.

[1] These ETFD estimates differ from those presented in the original papers because $\hat{V}_{AM}$ was not included in the total phenotypic variation in these papers. The parameters presented here have been corrected to include $\hat{V}_{AM}$ in the total phenotypic variation.

[2] Subscales of the Eysenck Personality Questionnaire–Short.

[3] Subscales within conservatism.

additive genetic correlation to $\frac{1}{2}$, for all phenotypes except four, church attendance, neuroticism (from Eaves, Heath, Martin, Maes, et al., 1999) and attitudes towards sex and conservatism, where we fitted the general sex limitation model with the male–female additive genetic correlation free (Neale & Cardon, 1992). For eight phenotypes (from Kendler et al., 1995), $C\hat{V}_{MZ}$ and $C\hat{V}_{DZ}$ were not reported separately for males and females. Parameter estimates differed significantly between the sexes on these phenotypes, so we could not use the combined $C\hat{V}_{MZ}$ and $C\hat{V}_{DZ}$ to accurately estimate the CTD parameters. Rather than excluding these eight phenotypes from the analysis, we treated the ETFD parameter estimates as 'population parameters', substituting them into the equations for CV_{MZ} and CV_{DZ} presented in Eaves et al. (1989) and Tambs et al. (1992). These algebraically derived $C\hat{V}_{MZ}$ and $C\hat{V}_{DZ}$ parameters were computed separately for males and females and then modeled in Mx to find the CTD parameters, as described above. To double check our method, we also algebraically derived $C\hat{V}_{MZ}$ and $C\hat{V}_{DZ}$ for several other phenotypes and compared these derivations to the $C\hat{V}_{MZ}$ and $C\hat{V}_{DZ}$ parameters reported in the ETFD papers. The algebraically derived estimates were similar and, importantly, did not appear to differ in any systematic way.

Results and Discussion

For each of the 17 phenotypes, Figure 1 presents the ETFD parameters $\hat{V}_A$, $\hat{V}_{NA}$, $\hat{V}_C$, $\hat{V}_E$, $\hat{V}_{AM}$, and $C\hat{V}_{GE}$ on the higher of the paired bars and the CTD parameters $\hat{V}_A$, $\hat{V}_{NA}$, $\hat{V}_C$, and $\hat{V}_E$, derived from model fitting on the lower of the paired bars. The exact values of these parameter estimates are also presented in Table 1. The combined genetic parameters in the CTD, $\hat{V}_A + \hat{V}_{NA} = \hat{V}_G$ (M = .39 standard units), were higher than in the ETFD (M = .33), although by only 18%. The most striking differences between the ETFD and CTD parameters were for $\hat{V}_A$ and $\hat{V}_{NA}$, which is not surprising given the high correlation between these two genetic parameters (Martin et al., 1978). As shown in Table 2, on average, $\hat{V}_A$ in the CTD (M = .31) was 63% larger than $\hat{V}_A$ in the ETFD (M = .19). Similarly, $\hat{V}_{NA}$ in the CTD (M = .08) was on average 43% smaller than $\hat{V}_{NA}$ in the ETFD (M = .14). There were three unusual cases, on male and female extraversion and male neuroticism (from the Lake et al., 2000 paper), where CTD estimates of $\hat{V}_A$ were smaller and $\hat{V}_{NA}$ were larger than equivalent ETFD estimates. These situations might reflect sampling error or some type of contrast effect unique to DZ twins. Our algebraic estimates for these three cases were also more reasonable, further supporting this conclusion.

On average, $\hat{V}_C$ was only slightly smaller in the CTD (M = .071) compared to the ETFD (M = .075). However, this small average difference conceals variability between phenotypes. While the presence of V_{NA} deflates $\hat{V}_C$ in ACE CTD models, the presence of V_{AM} and CV_{GE} will serve to inflate $\hat{V}_C$ in ACE CTD

models. For the traits that showed negligible effects of V_{AM} and/or CV_{GE} (political attitudes in males and females, attitudes towards lying and the military in females, and psychotocism and attitudes towards taxes in males), $\hat{V}_C$ was smaller in the CTD (M = .09) compared to the ETFD (M = .16). Similarly, for traits that were apparently influenced by V_{AM} and/or CV_{GE} (conservatism in males and females, church attendance and attitudes towards the military in males, and stature, attitudes towards sex and attitudes towards taxes in females), $\hat{V}_C$ (M = .18) was larger in the CTD relative to the ETFD (M = .11). In these cases, what initially would have appeared to be a strong effect of the common environment with the CTD was resolved to be due to either assortative mating or (to a smaller degree) genotype–environment covariance. This effect was also apparent from a significant correlation (r = .77, p = .001) between ETFD $\hat{V}_{AM}$ and the difference scores between ETFD $\hat{V}_C$ and CTD $\hat{V}_C$ parameters; higher $\hat{V}_{AM}$ was associated with upward biases in $\hat{V}_C$ in the CTD. In conclusion, while V_C from the CTD appears to sometimes underestimate $\hat{V}_C$, particularly in the presence of $\hat{V}_{NA}$, it may also overestimate V_C depending upon the degree to which assortative mating and genotype–environment covariation affect the trait.

We observed similar biases in $\hat{V}_A$, $\hat{V}_{NA}$, and $\hat{V}_C$ CTD parameters when we separately compared parameters estimated using ADE CTD models against those estimated using ACE CTD models (Table 2). However, the effects appeared to be less pronounced in the ADE models. This was because, for phenotypes modeled using the ADE method, where estimates of $\hat{V}_A$ and $\hat{V}_{NA}$ will be biased upwards and downwards respectively by V_C, CV_{GE}, and V_{AM}, the average ETFD estimates for $\hat{V}_C$ (M = .02), $\hat{V}_{AM}$ (M = .03) and $C\hat{V}_{GE}$ (M = −.01) were very low.

Limitations

Our study has four primary limitations. First, unfortunately, some studies (Kendler et al., 1994; 1995; Lake et al., 2000) only reported parameter estimates from the final, reduced models (where nonsignificant parameter estimates for $\hat{V}_C$ and $C\hat{V}_{GE}$ had been fixed to zero; see Table 1). In such cases, we used values of 'zero', even though these ETFD parameters may have potentially had a positive (albeit small and nonsignificant) value. This introduces a bias in the ETFD means and parameter estimates that we present. Nonsignificant parameters may not be significantly different from zero, but their values are nevertheless unbiased estimates of the true population parameters. Thus, the means reported here for $\hat{V}_C$ from the ETFD could be slightly underestimated. However, based on the large statistical power in ETFD studies and the likely low true values of ETFD parameters, these effects are unlikely to substantively alter our conclusions.

Table 2

Comparison of Standardized Parameters from the Classical Twin Design (CTD) and Extended Twin-Family Design (ETFD)

Parameter	CTD		ETFD		CTD–ETFD		Bias in CTD as % of ETFD[1]
	M	SD	M	SD	M	SD	
All traits (n = 38)							
$\hat{V}_A$	.31	.18	.19	.14	.12	.14	.63
$\hat{V}_{NA}$	.08	.13	.14	.09	−.06	.10	−.43
$\hat{V}_C$	.07	.09	.07	.08	.00	.06	.00
$\hat{V}_G$	.39	.15	.33	.14	.06	.06	.18
Traits in which $C\hat{V}_{DZ}/C\hat{V}_{MZ} > \frac{1}{2}$ (n = 22)							
$\hat{V}_A$	.37	.16	.20	.16	.17	.10	.85
$\hat{V}_{NA}{}^2$	.00	.00	.10	.05	−.10	.05	
$\hat{V}_C$	.12	.09	.12	.07	−.01	.08	.00
$\hat{V}_G$	.37	.16	.30	.14	.08	.07	.23
Traits in which $C\hat{V}_{DZ}/C\hat{V}_{MZ} < \frac{1}{2}$ (n = 16)							
$\hat{V}_A$	.23	.19	.17	.11	.06	.16	.28
$\hat{V}_{NA}$	.18	.15	.20	.09	−.02	.13	−.10
$\hat{V}_C{}^3$	.00	.02	.00	.02	−.02	.02	
$\hat{V}_G$	.41	.15	.38	.13	.04	.04	.08

Note: [1]When CTD > ETFD, CTD/ETFD -1; when CTD < ETFD, 1- CTD/ETFD. [2]Parameter is fixed to 0 in the CTD (ACE models). [3]Parameter is fixed to 0 in the CTD (ADE models). See text for explanation of parameters.

Second, there have only been a limited number of ETFD studies conducted, and because of this, our analysis was based on only 17 traits. It would be unwise to extrapolate these findings too widely because these 17 traits are almost certainly not a representative sample of traits studied using the CTD. Many of the traits (e.g., subscales of attitudes) reported here are highly correlated with one another. Moreover, some traits appear to have been selected specifically as test cases for whether assortative mating could explain apparent common environmental effects. We expect that a random sample of traits studied by behavior geneticists might show less effects of $\hat{V}_{AM}$ and thus less downward bias in $\hat{V}_{NA}$ and greater downward bias in $\hat{V}_C$ from the CTD.

Third, our biases were estimated by comparing CTD estimates derived only from MZ and DZ twin correlations to ETFD estimates derived from 80 correlations among twin-family kinships. Some of the differences between CTD and ETFD parameters were therefore not only due to parameter bias, but also to differences between the samples and sampling errors. This effect adds 'noise' to our results; it should not have presented systematic bias.

A final limitation to the present study is that the degree to which genetic nonadditivity is due to higher-order epistasis (parameterized by the value r) could not be reliably estimated in either the CTD or the ETFD. To the degree that $r < \frac{1}{4}$ (that higher-order epistasis is involved in trait variation), $\hat{V}_C$ and $\hat{V}_{AM}$ parameters will be biased downwards while $\hat{V}_{NA}$ parameters will be biased upwards in both CTD and ETFD studies. Therefore, our findings probably underestimate the

degree of bias in $\hat{V}_C$ and $\hat{V}_{AM}$ and overestimate the degree of bias in $\hat{V}_{NA}$ from CTD studies. For similar reasons, the broad-sense heritability parameters from both the ETFD and (especially) the CTD should generally be regarded as overestimates. However, in a companion report (Keller & Coventry, 2005), it is shown that the source of bias in the CTD that stems from fixing $\hat{r} = \frac{1}{4}$ tends to have much less *potential* to bias CTD parameter estimates than the bias that stems from fixing either $\hat{V}_C$ or $\hat{V}_{NA}$ to 0. Therefore, although the CTD biases derived in this report are probably smaller than their true values for the reasons spelled out above, our method has probably captured the majority of the bias that exists in CTD parameter estimates.

Conclusion

These comparisons between CTD and ETFD parameter estimates are consistent with the biases predicted to exist in CTD parameters (Eaves et al., 1978; Grayson, 1989; Keller & Coventry, 2005). While $\hat{V}_G$ parameters from CTD studies tend to be only moderately overestimated, the estimates for $\hat{V}_A$ and $\hat{V}_{NA}$ from CTD studies appear to be substantially biased. Neither ACE nor ADE models of the CTD were able to accurately assess the genetic architectures of traits. When both $\hat{V}_A$ and $\hat{V}_{NA}$ are estimated (ADE models), their combined value would be better considered an estimate of broad-sense heritability. Similarly, when $\hat{V}_{NA}$ is not included in CTD studies (ACE or AE models), $\hat{V}_A$ estimates would be more appropriately considered estimates of broad-sense rather than narrow-sense heritability.

Our results also show that V_{NA} conceals factors that increase DZ similarity, such as the effects of the

common environment, assortative mating, and to a lesser degree, genotype–environment covariation. In general, therefore, V_C from the CTD should be considered an underestimate of the combined effects of V_C, V_{AM}, and CV_{GE}. For certain phenotypes, assortative mating or genotype–environment covariation accounts for most of what would otherwise appear to be effects of the common environment in CTD studies. For other phenotypes, effects of the common environment are indeed underestimated by the CTD, although the magnitudes of this bias do not seem to be large enough to call into question the general finding from CTD studies that the common environment plays surprisingly little role in the variation of most traits.

For both the CTD and the ETFD, rather than the normal practice of reporting only those estimates that result from fixing parameters at assumed and ultimately arbitrary values, it seems preferable to report the full range of estimated parameter values that are possible given different values of the fixed parameters (Keller & Coventry, 2005). In doing so, a clear sense of possible parameter values is conveyed to readers. In our companion report, we demonstrate a graphical approach for presenting parameter indeterminacy in the CTD. Such a presentation is more easily accomplished using the ETFD because only a single value, r, needs to be varied. Along these lines, we suggest that researchers, who use the ETFD, report parameter estimates from the full model in addition, perhaps, to those from reduced models. Parameter estimates from full models may have wider confidence intervals, but they are also expected to be less biased (Keller & Coventry, 2005). Presenting both types of estimates seems preferable.

The ETFD is an enormously useful, although underutilized, method for understanding the causes of variation in a trait. The CTD is more limited in its ability to resolve various aspects of phenotypic variation. Nevertheless, researchers using the CTD should be heartened by the present results which suggest that previous CTD estimates of overall genetic and common environmental effects are unlikely to be very inaccurate. This conclusion must be tempered, of course, by the fact that these results may not generalize to all phenotypes, and by the fact that a potentially important source of bias, the makeup of genetic nonadditivity in phenotypes, remains poorly understood.

Endnotes

1 We use 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 The presence of genotype–environment correlations also serve to attenuate, but to a lesser extent, the effects that assortative mating have when they mimic V_C in biasing CTD estimates (Eaves et al., 1989; Tambs et al., 1992).

3 The effects of assortment can also act in a way that does not mimic V_C, but, essentially, this is only to serve as a subtle downward bias on V_A in ACE models of the CTD. In virtually all other instances, the effects of assortment mirror the effects of V_C (Eaves et al., 1989; Tambs et al., 1992).

4 While it is possible to differentiate additive-by-additive epistasis from dominance by comparing parent–offspring to sibling–sibling correlations (Heath et al., 1984), this still does not enable an estimation of r (Keller & Coventry, 2005).

5 We omitted the Tambs et al. (1992) paper (and others by the same first author) as they did not report variance components. Two studies (Eaves, Heath, Martin, Maes, et al., 1999; Eaves, Heath, Martin, Neale, et al., 1999) estimated parameters for conservatism from the Virginia data, so we only consider the estimates from one (Eaves, Heath, Martin, Neale, et al., 1999). Eaves, Heath, Martin, Maes, et al. (1999) and Lake et al. (2000) both derive estimates for neuroticism but from different data sets so we include both.

References

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

Eaves, L. J., Eysenck, H. J., & Martin, N. G. (1989). *Genes, culture and personality: An empirical approach.* London: Academic Press.

Eaves, L. J., Heath, A. C., Martin, N. G., Maes, H. H., Neale, M. C., Kendler, K. S., Kirk, K. M., & Corey, L. A. (1999). Comparing the biological and cultural inheritance of personality and social attitudes in the Virginia 30,000 study of twins and their relatives. *Twin Research, 2*, 62–80.

Eaves, L. J., Heath, A. C., Martin, N. G., Neale, M. C., Meyer, J. M., Silberg, J. L., Corey, L. A., Truett, K., & Walters, E. (1999). Biological and cultural inheritance of stature and attitudes. In C. R. Cloninger (Ed.), *Personality and psychopathology*. Washington, DC: American Psychiatric Association.

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.

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. (2005). Quantifying and addressing parameter indeterminacy in the Classical Twin Design. *Twin Research, 8*, 201–213.

Kendler, K. S., Walters, E. E., Truett, K. R., Heath, A. C., Neale, M. C., Martin, N. G., & Eaves, L. J. (1994). Sources of individual differences in depressive symptoms: Analysis of two samples of twins and their families. *American Journal of Psychiatry, 151*, 1605–1614.

Kendler, K. S., Walters, E. E., Truett, K. R., Heath, A. C., Neale, M. C., Martin, N. G., & Eaves, L. J. (1995). A twin-family study of self-report symptoms of panic-phobia and somatization. *Behavior Genetics, 25*, 499–515.

Lake, R. I. E., Eaves, L. J., Maes, H. H., Heath, A. C., & Martin, N. G. (2000). Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. *Behavior Genetics, 30*, 223–233.

Maes, H. H. M., Neale, M. C., & Eaves, L. J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior Genetics, 27*, 325–351.

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

Neale, M. C. (2004). *Mx: Statistical modeling* (Version 1.5.01). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Neale, M. C., & Cardon, L. R. (2004). *Methodology for genetic studies for twins and families*. Dordrecht, the Netherlands: Kluwer Academic.

Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Research, 5*, 554–571.

Tambs, K., Moum, T., Holmen, J., Eaves, L. J., Neale, M. C., Lund-Larsen, P. G., & Naess, S. (1992). Genetic and environmental effects on blood pressure in a Norwegian sample. *Genetic Epidemiology, 9*, 11–26.

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