
A Finite Mixture Model for Genotype and Environment Interactions: Detecting Latent Population Heterogeneity

Nathan A. Gillespie and Michael C. Neale

Virginia Institute of Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, Richmond, United States of America

Approaches such as DeFries-Fulker extremes regression (LaBuda et al., 1986) are commonly used in genetically informative studies to assess whether familial resemblance varies as a function of the scores of pairs of twins. While useful for detecting such effects, formal modeling of differences in variance components as a function of pairs' trait scores is rarely attempted. We therefore present a finite mixture model which specifies that the population consists of latent groups which may differ in (i) their means, and (ii) the relative impact of genetic and environmental factors on within-group variation and covariation. This model may be considered as a special case of a factor mixture model, which combines the features of a latent class model with those of a latent trait model. Various models for the class membership of twin pairs may be employed, including additive genetic, common environment, specific environment or major locus (QTL) factors. Simulation results based on variance components derived from Turkheimer and colleagues (2003), illustrate the impact of factors such as the difference in group means and variance components on the feasibility of correctly estimating the parameters of the mixture model. Model-fitting analyses estimated group heritability as .49, which is significantly greater than heritability for the rest of the population in early childhood. These results suggest that factor mixture modeling is sufficiently robust for detecting heterogeneous populations even when group mean differences are modest.

The classical twin study yields data which are useful for decomposing both variance within a trait and the covariance between traits. Provided that twins are sampled at random from the population, and that zygosity is correctly diagnosed (or its uncertainty is modeled; Neale, 2003), this method is robust to minor violations in the equal environment assumption (Vandenberg, 1984). However, the presence of genotype by environmental interactions ($G \times E$), genotype and environmental correlations, and nonrandom mating can bias parameter estimates. This article will

focus on estimating heritability in the presence of population heterogeneity.

One popular approach to testing for $G \times E$ is to use a measured environmental variable (Boomsma et al., 1999; Heath et al., 1998; Purcell & Sham, 2002). Another is to use a measured genotype (Eaves & Sullivan, 2001; Martin & Oakeshott, 1983). Other studies have used both (Caspi et al., 2003; Martin et al., 1987). $G \times E$ interaction is observed if there are genetic factors which regulate the overall degree of vulnerability or sensitivity to the pathogenic effects of environmental stress (Kendler & Eaves, 1986; Kendler & Karkowski-Shuman, 1997). All these approaches require that a relevant environmental or genetic interacting factor has been measured. However, it seems likely that such interactions may exist even when the relevant factor has not been measured. The aim of this article is to examine whether such interactions can be detected when direct measurement of the salient interacting variable (be it environmental or genetic or some combination thereof) has not been made.

The usual variance components model employed in the analysis of monozygotic (MZ) and dizygotic (DZ) twin data (see Chapter 8, Neale & Cardon, 1992) partitions trait variance into additive genetic (A), common environment (C) and specific environment (E) components. For the sake of simplicity we will ignore genetic dominance (D) in what follows, although the model we propose would easily be adapted to D as well. The standard univariate ACE model does not explicitly model $G \times E$. If present, $G \times E$ may bias variance component estimates, and the population heterogeneity which it implies would likely go undetected. Variance component estimates are averages across the population in question; the utility of this average is lost if the

Received 31 January, 2006; accepted 5 April, 2006.

Address for correspondence: Nathan A. Gillespie, Virginia Institute of Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, 800 East Leigh Street, Biotech 1, Suite 101, Richmond, VA 23219-1534, USA. E-mail: ngillespie@vcu.edu

population consists of a mixture of subpopulations with widely varying variance components. Akin to the case where there exists a measured environmental factor which affects, for example, the heritability of a trait, we can imagine that this external factor itself is not purely of unique individual specific environment origin, but may be influenced by genetic factors or shared environmental factors, and therefore may be correlated between relatives. Thus the interacting external factor might be binary, ordinal or continuous, and might be influenced by one or more quantitative trait loci (QTLs) or environmental factors. We therefore denote the variance components of the external interacting factor by the subscript F : Q_F , A_F , C_F or E_F . Since any of these external variance components may change the size of any one of the trait's variance components, we can envisage 12 possible types of interaction ($Q_F \times A$, $Q_F \times C$, $Q_F \times E$; $A_F \times A$, $A_F \times C$, $A_F \times E$; $C_F \times A$, $C_F \times C$, $C_F \times E$; and $E_F \times A$, $E_F \times C$, $E_F \times E$). We emphasize that the external factor is not measured, that is, latent, and therefore refer to the model as a latent heterogeneity variance components (LHVC) model. When the external factor (Q_F , A_F , C_F , or E_F) itself generates a mean difference in the trait of interest, it generates part of the trait variance. In other words, the variance component may be partitioned into that due to the external factor (albeit latent) and that due to the residual.

Genetic epidemiologists have shown interest in testing whether the heritability of a trait varies as a function of the value of the trait itself. For example, individuals with extremely high scores on the trait might have a higher heritability than those with lower phenotypic scores. The DeFries-Fulker (DF, DeFries & Fulker, 1985) extremes regression (LaBuda et al., 1986) provides a method of identifying this type of LHVC. In practice, the method involves selecting a particular cut-point on the phenotypic distribution, and testing whether the regression of the twin's phenotype on that of their co-twin differs for subjects above versus below this cut-point. However, the interpretation of such observed differences is obscure, because there is no formal generating model for the observed effect. Indeed, such effects might be produced by any one of the 12 interactions described above, but attempts to identify the particular type of interaction do not appear common except in the animal populations using genetically informative data, for example, milk yield in dairy cows (see Brotherstone & Hill, 1986; Gianola et al., 1992; Foulley & Quaas, 1995).

Unfortunately, in human populations formal modeling of differences in variance components purely as a function of trait scores via mixture distributions is challenging. Even if the technical obstacles to fitting such models have largely been overcome, there remain issues with the possibly large sample sizes required, the statistical properties of tests and the distributional assumptions (see Lubke & Muthén, 2005). For example, nonnormality is known to produce spurious

evidence for multiple classes in the context of growth curves (Bauer & Curran, 2003). Our simulations here assume that there are no scaling issues, though we return to this issue in the discussion.

To our knowledge, variance component mixture modeling has never been attempted within the classical twin study design. In the case of the DF extremes regression method, there is an a priori classification into high and low groups, either based on the trait value itself, or on a variable closely related to it. When there are systematic changes in variance components across the scale, it implies that the grouping factor (whose own variance components are A_F , C_F , and so on) itself has an effect on the trait mean. Therefore, we confine our treatment here to grouping variables which also influence the trait mean, even though it is possible that grouping factors may exist which are not related to the trait under study. The work we present here is intended simply as a first investigation into whether LHVC models are ever identified with univariate data collected from twins, rather than a comprehensive study of the necessary and sufficient conditions and sample sizes that may be required. To illustrate the approach, we consider the results of Turkheimer and colleagues (2003) who reported that the variance components in a measure of cognitive ability differ between high and low socioeconomic status (SES) groups. In particular, shared environmental factors account for a larger proportion of variance among impoverished families. Here we investigate whether this pattern of differences might be detected in the absence of the external interacting SES variable.

Aim

We present a biometrical factor mixture approach which models the population as consisting of two classes which differ in their means as well as the relative impact of genetic and environmental factors on within-group variation and covariation. We also model the genetic and environmental etiology of the latent class membership, or grouping mechanisms. The estimated variance components obtained by Turkheimer and colleagues (2003) for high and low SES groups were used as the basis for the simulation study described below. Using only twins' trait scores from the simulated data, our aims are to determine if biometrical factor mixture modeling (FMM) can accurately recover the following features: (i) the variance components within each class; (ii) the latent class membership probabilities; and (iii) the etiological mechanism for class membership.

Materials and Methods

Model for Class Membership

Throughout this study, only two classes are considered, 'high' and 'low', according to the mean phenotypic value. Although not exhaustive, five mechanisms for an individual's class membership were tested: random effects (nonshared environment 'E');

Table 1
Standardized Components of Variance Within Each Latent Class Based on Turkheimer and Colleagues (2003)

	Class 1			Class 2		
Standardized variance components	A ₁	C ₁	E ₁	A ₂	C ₂	E ₂
	.72	.15	.13	.10	.58	.32
	r_{MZ}			r_{DZ}		
Class 1 concordant	.72 + .15			.5 × .72 + .15		
Discordant	$\sqrt{(.72 \times .10) + \sqrt{(.15 \times .58)}}$			$1/2 \times \sqrt{(.72 \times .10) + \sqrt{(.15 \times .58)}}$		
Class 2 concordant	.10 + .58			.5 × .10 + .58		

Note: Also included are the expected twin pair correlations by zygosity and level of class membership concordance.

cultural effects (shared environment ‘C’); additive genetic differences (additive genetic ‘A’); combined genetic and environmental effects (ACE); and a hypothetical major locus effect caused by differences in allele frequency (QTL dominant).

Under the A, C and dominant QTL class membership models, all MZ twin pairs are concordant for class membership (i.e., there exist only high–high or low–low pairs). DZ pairs are more complicated. Under the purely A class membership model, the pairwise membership probabilities (high–high, high–low, low–high and low–low) for each of the 2 × 2 cells were calculated using a bivariate normal distribution with the correlation equal to one-half. Threshold cut-points on the normal distribution were chosen to obtain the appropriate probability of class membership (e.g., a threshold of 1.282 would be used to give a .9 probability of membership in class 1). Under the purely C model for class membership, DZ twin pairs are always concordant, like the MZ pairs. When class membership is due to E factors, both MZ and DZ twin pairs may be concordant high, concordant low, discordant high–low or low–high. Furthermore, the probabilities that MZ pairs belong to these classes are equal to those for the DZ pairs. Under the QTL model, the class membership probabilities for the DZ pairs were calculated according a simple diallelic dominant locus (e.g., Neale et al., 1999, see Table 1).

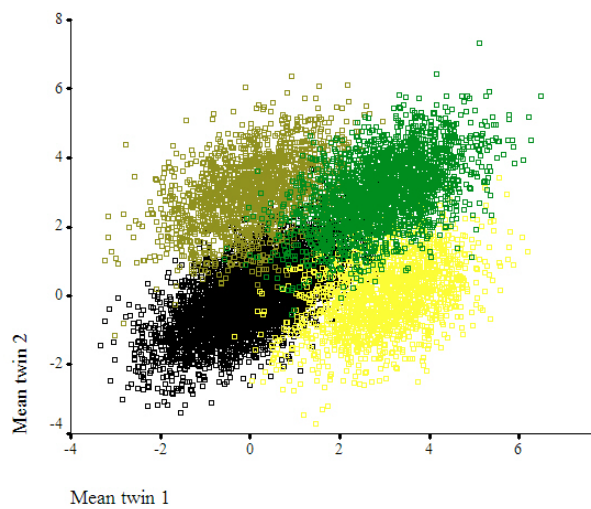
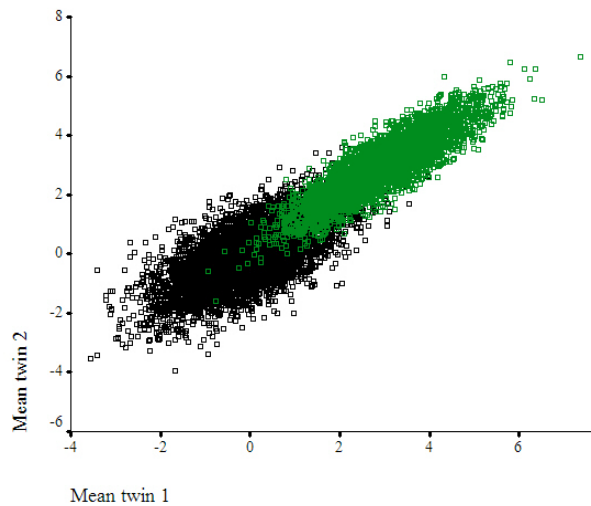
Within-Class Model

Within-class variance components were held constant across all conditions such that in class 1, the standardized variance components were $a_{hi}^2 = .72$, $c_{hi}^2 = .15$, and $e_{hi}^2 = .13$ for additive genetic, shared and nonshared environmental effects respectively. In class 2, the same parameters were $a_{low}^2 = .10$, $c_{low}^2 = .58$ and $e_{low}^2 = .30$. These correspond directly to the variance components for high and low SES groups estimated by Turkheimer and colleagues (2003). Our simulation is based on the assumption that the same genetic and environmental factors operate in the two classes, but their magnitude differs. That is to say, the model parallels the ‘scalar sex-limitation’ model as opposed to the ‘nonscalar sex-limitation’ model (Neale & Cardon, 1992). Therefore, the expected covariance between MZ pairs concordant

for the high class membership is $a_{hi}^2 + c_{hi}^2$; that for the low class is $a_{low}^2 + c_{low}^2$; and that for discordant pairs is $a_{hi} \times a_{low} + c_{hi} \times c_{low}$. The DZ within class covariances are similar, but have only one-half the additive genetic contribution to pair covariance.

Example scatterplots of the simulated data are shown for each of the five class membership models. In these figures, the mean difference between classes was set at 3.0, and the within-class variation was set according to the within-class model just described.

Note that, except for the cases including E as part of the mechanism for group membership, the MZ twin pairs consist of only two classes, leading to a phenotypic mixture distribution consisting of two bivariate normal distributions with different means. In the class with the higher means, the dispersion of the points is slightly closer to the regression line than is the case for the class with the lower mean, reflecting the higher within-class correlation for the high mean group (.87 vs. .68). This pattern is the same regardless of the group membership mechanism. In the E and ACE mechanisms, two additional mixture components are generated, corresponding to cases where the MZ twin pairs are discordant for group membership. The dispersion of the MZ pairs discordant for class membership around their regression line is greater than for concordant pairs, reflecting the lower within-class correlation (.56) for this group. Overall, the pattern of the scatterplots for the DZ data is similar to that of the MZ pairs, except that for the DZ pairs the correlation for pairs concordant for the higher mean group is *less* than that for the lower mean group (.51 vs. .63) which is opposite to the MZ pattern. Again, the class-discordant pairs have a lower correlation than either of the same-class groups. Note that this pattern is not due to different genetic or environmental factors operating in the high versus low mean groups, but simply due to the dramatic contrast between the variance components (high A and low C vs. low A and high C) in the two groups. Clearly, given a purely AE model for one group, and a purely CE for the other, the prediction would be of zero correlation between twins (either MZ or DZ) for pairs discordant for group membership.

**Figure 1a**

Distribution of simulated data under an additive genetic (A) class membership mechanism for MZ twin pairs (above) and DZ twin pairs (below).

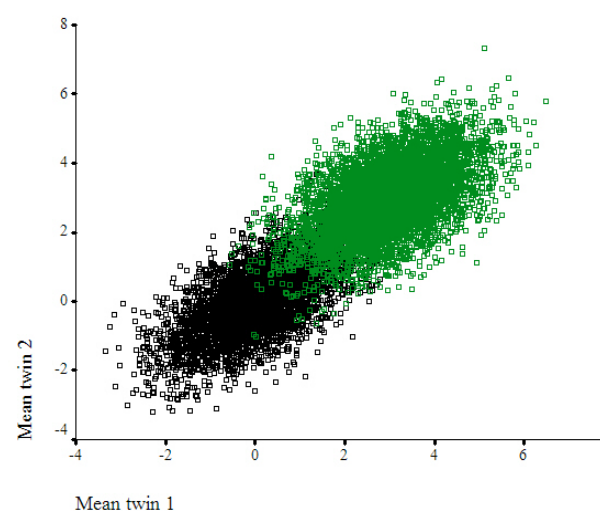
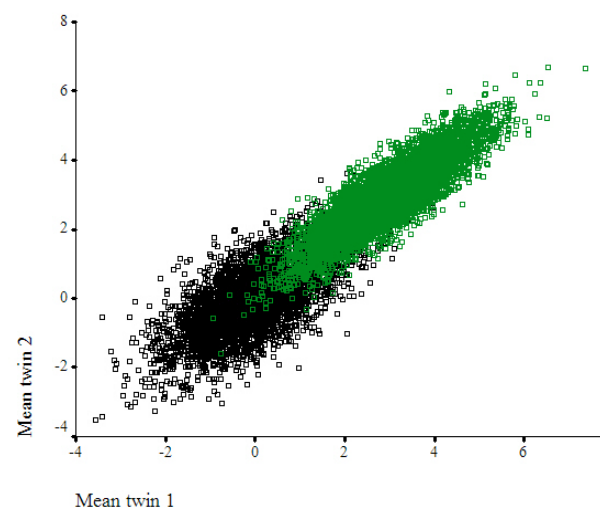
Note: Simulated data of 10,000 MZ and 10,000 DZ twin pairs, with a 3 SD mean difference. Class membership probability of .60 and DZ twin pair correlation of .50.

Simulation Design Parameters

Each mechanism consisted of six levels of varying group membership probability (or allele frequency): .60, .65, .70, .75, .80, and .85. For each level we simulated data for 20,000 twin pairs (10,000 MZ, 10,000 DZ) using SAS software (Version 8 of the SAS System for Unix). Throughout the simulation we used a two class model with simulated mean differences of 1½ and 3 standard deviations. The simulation exercise was repeated 1000 times for each of these conditions, so in total $5 \times 6 \times 20,000 \times 2 \times 1000 = 1.2 \times 10^6$ scores were simulated.

Biometrical Factor Mixture Modeling

With increases in computing power, the last decade has seen a growing interest in FMM. FMM combines the features of *latent class* with those of *latent trait*

**Figure 1b**

Distribution of simulated data under a shared environmental (C) class membership mechanism for MZ twin pairs (above) and DZ twin pairs (below).

Note: Simulated data of 10,000 MZ and 10,000 DZ twin pairs with a mean difference of 3 SD class membership probability of .60.

modeling. Latent class analysis identifies subtypes, or mixture components, of related cases from multivariate categorical data. Latent trait, or structure analysis (see Lazarsfeld & Henry, 1968), is a form of factor analysis which may be used for binary or categorical data. In the case of continuous data, as considered here, the latent profile model is the homologue of latent class analysis. The name factor model is commonly used regardless of whether the input data are continuous, ordinal or binary. Thus latent profile models are mixture models with zero common factors but differing group means and/or variances, while factor models have common factors but consider the population to be homogeneous (i.e., only one component in the mixture). Factor mixture models are a generalization of both model types, including both factors and mixture components (Lubke & Muthén, 2005).

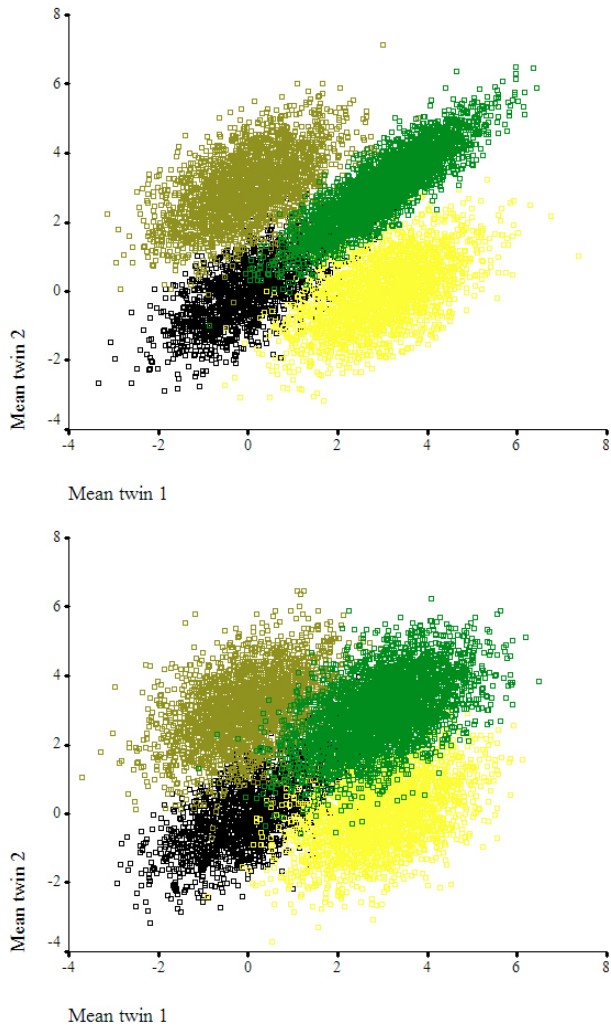


Figure 1c

Distribution of simulated data under a nonshared environmental (E) class membership mechanism for MZ twin pairs (above) and DZ twin pairs (below).

Note: Simulated data of 10,000 MZ and 10,000 DZ twin pairs with a mean difference of 3 SD, and class membership probability of .60.

We implemented FMM which employed the standard biometrical genetic variance components to account for within-class covariation between relatives (Neale & Cardon, 1992). Thus the total variance in the simulated trait scores is decomposed into additive (A) genetic, as well as common (C) and unique (E) environmental variance. MZ co-twins are genetically identical, so additive genetic effects correlate 1.0 in MZ twin pairs but only .5 in DZ twin pairs. An important assumption of this biometrical model is that shared environmental effects correlate to an equal extent in MZ and DZ twin pairs (Kendler et al., 1994). Nonshared environmental effects are by definition uncorrelated and also reflect measurement error including short-term fluctuations.

An assumption of the usual biometrical model is that the entire population is homogeneous with respect to variance components. This assumption was relaxed in

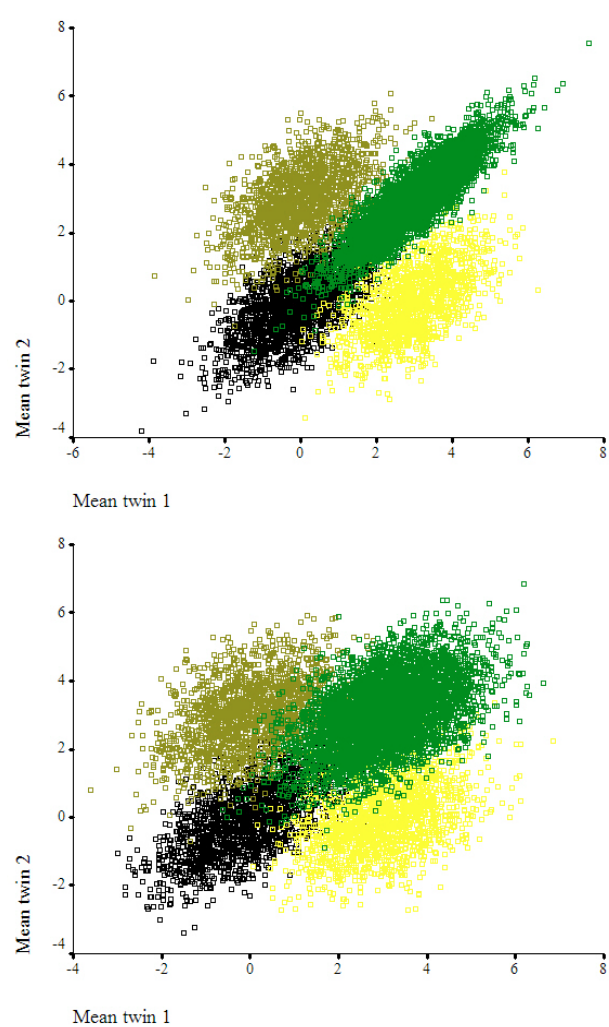


Figure 1d

Distribution of simulated data under an additive genetic, shared and nonshared environmental (ACE) class membership mechanism for MZ twin pairs (above) and DZ twin pairs (below).

Note: Simulated data of 10,000 MZ and 10,000 DZ twin pairs with a mean difference of 3 SD, MZ twin pair correlation of .50, and class membership probability of .65. This corresponds to multivariate cell probabilities of .4964 (High-High), .1536 (Low-High), .1536, (High-Low), and .1964 (Low-Low). The DZ twin pair correlation and the probability of class membership were fixed at .40 and .65 respectively. This corresponds to multivariate cell probabilities of .4803 (High-High), .1697 (Low-High), .1697, (High-Low), and .1803 (Low-Low).

the biometric FMM. Thus the model represents a hybrid latent class model for group membership (the mixture component) and a latent trait model for covariance between twins within group. This model was implemented in Mx (Neale, 1999) using its mixture distribution features NModel and Weight. NModel controls the number of hypothetical models for the twin pairs, which was set to four: low-low; low-high; high-low; and high-high. Per Mx syntax for mixture distributions, the user supplies predicted means and covariance matrix formulae which vertically stack the components of the mixture. In this case a 1×2 vector of means (for twin 1 and twin 2) and a 2×2 predicted

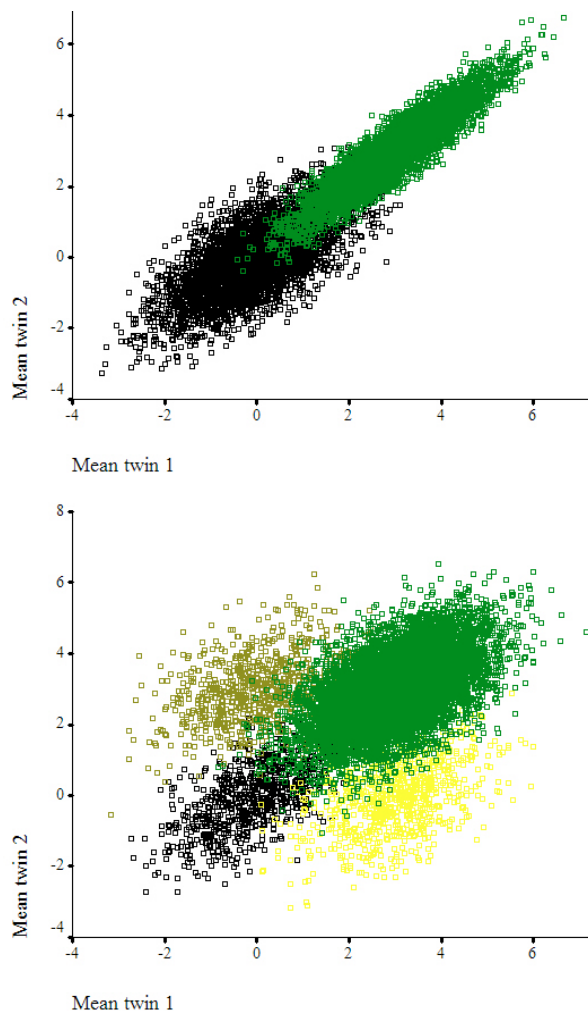


Figure 1e. Distribution of simulated data under a QTL (dominant) class membership mechanism for MZ twin pairs (above) and DZ twin pairs (below).

Note: Simulated data of 10,000 MZ and 10,000 DZ twin pairs with a 3 SD mean difference. An allele frequency of .50/.50 corresponds to DZ twin pair multivariate cell probabilities of .5625 (High–High), .1875 (Low–High), .1875 (High–Low), and .0625 (Low–Low).

covariance is provided for each of the four components of the mixture distribution (low–low, low–high, high–low, high–high), which resulted in a 4×2 predicted mean vector and 8×2 predicted covariance matrix. The Weight command provides a vector (4×1 in this case) of probabilities of pairwise class membership. In general, the likelihood of an observed data vector under a mixture distribution can be written:

$$\ln L_{Nmodel} = \sum_{i=1}^{Nmodel} \ln (w_i L_i)$$

where w_i is the weight, and L_i is the likelihood for the i^{th} model. In the present case, the weights are free parameters which estimate the class membership probabilities. However, in the case of the dominant QTL model of class membership, the obtained weights are a function of the QTL allele frequency which have been calculated elsewhere (see Neale et al., 1999).

All models were fitted to the simulated raw continuous data by full information maximum likelihood (FIML). Departure of parameter estimates from the values used for the simulation was assessed by comparing the likelihood of the model with parameters freely estimated to that of a nested submodel in which the variance components and class membership probabilities (or QTL frequencies) under the simulated conditions were fixed to the values used to simulate the data. Twice the negative difference in log-likelihoods between these models is, under certain regularity conditions, asymptotically distributed as a chi-square statistic with degrees of freedom equal to the number of parameters in the model (9). p values and change in the log-likelihood were averaged across the entire simulated data to evaluate overall model fit under each of the five conditions. This exercise was repeated for the simulations and examined, via QQ-plot, for departure from the expected distribution of statistics (these plots are not shown but are available on request from the first author). Posterior class membership probabilities for each pair can also be obtained from Mx. Column 11 from the individual likelihood file (generated using the command Option Mx%P = filename) contains the weighted likelihood $w_i L_i$ for twin pair j under each of the $i = 1 \dots 4$ models. The posterior class probability that a twin pair belongs to class i is given by $w_i L_i / \sum_{i=1} (w_i L_i)$

Typically, it is better to conduct hypothesis testing and data analysis via model fitting, which retains the uncertainty about twin pair class membership. However, there may be occasions — such as the attempt to locate a QTL or an environmental factor which gives rise to group membership — where classification of the twin pairs may prove useful.

Results

Regardless of the class membership mechanism, biometrical FMM reliably recovered the simulated population parameters. This was true even for smaller group differences of 1.5 standard deviations, as well as with diminishing probabilities of class membership, that is, p approaching .85. Maximum and minimum ranges in the population parameters were only marginally wider under the condition of smaller mean separation. However, the statistical tests for the difference between the recovered parameters and those used for the simulation did not perform accurately under several of the simulation conditions. Typically, the proportion of tests reaching the .05 significance level exceeded that nominal expected proportion. On the whole, the test performed better for the larger than the smaller class differences. These results suggest that care should be used when drawing conclusions about the statistical significance of empirical findings obtained with this method.

Table 2a

Comparison of Simulated and Recovered Parameter Estimates Under the Additive Genetic (A) Class Membership Mixture Model

Class membership probabilities and corresponding thresholds for class membership (τ) [†]													
		.60, $\tau = -.2533$.65, $\tau = -.3853$.70, $\tau = -.5244$.75, $\tau = -.6745$.80, $\tau = -.8416$.85, $\tau = -1.036$	
True values		μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>
A ₁	.72	.7179	.0213	.7183	.0220	.7193	.0159	.7182	.0234	.7184	.0257	.7179	.0289
C ₁	.15	.1520	.0207	.1516	.0213	.1507	.0155	.1518	.0228	.1517	.0251	.1521	.0281
E ₁	.13	.1301	.0035	.1301	.0038	.1300	.0029	.1300	.0045	.1299	.0050	.1300	.0058
A ₂	.10	.1002	.0230	.0999	.0217	.1002	.0146	.0991	.0191	.0989	.0185	.0988	.0175
C ₂	.58	.5795	.0207	.5797	.0194	.5799	.0129	.5805	.0167	.5807	.0159	.5808	.0150
E ₂	.32	.3203	.0062	.3203	.0059	.3200	.0041	.3204	.0056	.3204	.0055	.3204	.0053
μ_1	.00	.0001	.0142	.0001	.0149	-.0001	.0082	-.0003	.0170	-.0007	.0189	-.0001	.0219
μ_2	3.00	2.9998	.0117	2.9998	.0115	3.0001	.0112	2.9999	.0109	2.9998	.0107	2.9999	.0107
τ		-.2531	.0099	-.3852	.0099	.5245	.0070	-.6744	.0103	-.8414	.0109	-1.0359	.0121
p^{\dagger}		.5653		.5648		.4823		.5729		.5637		.5558	
$p < .05$.08		.09		.08		.08		.08		.08	

Class membership probabilities and corresponding thresholds for class membership (τ) [†]													
		.60, $\tau = -.2533$.65, $\tau = -.3853$.70, $\tau = -.5244$.75, $\tau = -.6745$.80, $\tau = -.8416$.85, $\tau = -1.036$	
True values		μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>
A ₁	.72	.7166	.0286	.7173	.0289	.7178	.0304	.7178	.0307	.7175	.0337	.7192	.0264
C ₁	.15	.1533	.0273	.1526	.0275	.1521	.0290	.1521	.0291	.1523	.0319	.1508	.0246
E ₁	.13	.1301	.0049	.1301	.0055	.1301	.0061	.1302	.0070	.1302	.0081	.1300	.0069
A ₂	.10	.1000	.0283	.0998	.0275	.0996	.0255	.0990	.0245	.0986	.0236	.0991	.0158
C ₂	.58	.5796	.0255	.5799	.0246	.5800	.0225	.5806	.0213	.5810	.0202	.5815	.0143
E ₂	.32	.3204	.0069	.3203	.0067	.3204	.0065	.3204	.0063	.3204	.0061	.3194	.0064
μ_1	.00	.0000	.0250	-.0002	.0270	-.0002	.0292	-.0001	.0334	-.0003	.0381	-.0143	.0927
μ_2	1.50	1.4998	.0188	1.4996	.0178	1.4998	.0167	1.5000	.0158	1.5000	.0148	1.5155	.0869
τ		-.2531	.0304	-.3855	.0311	-.5245	.0317	-.6742	.0341	-.8410	.0367	-1.0341	.0311
p^{\dagger}		.5486		.5479		.5464		.5547		.5513		.4785	
$p < .05$.09		.09		.09		.10		.09		.09	

Note: Shaded areas are simulated population parameters.

A₁, C₁, and E₁ additive genetic, shared and nonshared environmental standardized variance components for class 1.

A₂, C₂, and E₂ additive genetic, shared and nonshared environmental standardized variance components for class 2.

μ_1 and μ_2 class 1 and 2 means.

[†] Average p value based on 9 *df*.

$p < .05$ Proportion of sample with p values less than .05.

[†] The Mx `allint` function was used to estimate the 2 x 2 cell frequencies for DZ twin pairs using group membership probabilities (ranging .60 to .85 by .05) to estimate standardized thresholds, while assuming a twin pair correlation of .5.

Discussion

When based simply on twins' trait scores, it was uncertain a priori whether biometrical factor mixture model would be able to detect population heterogeneity reliably, regardless of the mechanism of class membership. The results presented here suggest that under certain conditions this is possible. Biometrical FMM thus offers a practical means of inferring population heterogeneity when analysing genetically informative data. We simulated models of random, shared environmental, additive genetic, combined genetic and environmental, as well as a dominant QTL mechanisms responsible for class membership. We also varied the mean group difference from 1.5 to 3 standard deviations. However, even a 1.5 *SD* difference between groups is quite substantial. Latent heterogeneity with

less dramatic group differences would likely be more difficult to detect and estimate accurately.

Part of FMM's ability to recover population parameters derives from the fact that we imposed a mixture constraint such that the variance components within each class were constrained to equal one. When we relaxed this equality of variance constraint, optimization still managed to recover parameter estimates in all instances. As a further test, we simulated variance components exceeding 1 by augmenting the genetic variance in group one (previously $a^2 = .72$) by increments of 5%. Under these conditions, Mx did not recover the population parameters reliably. This behavior, though disappointing, is not surprising, because the likelihood surface over parameters of a mixture distribution is known to contain strong local minima (McLachlan & Basford, 1988; Redner & Walker, 1984). However, by

Table 2b

Comparison of Simulated and Recovered Parameter Estimates Under the Shared Environmental (C) Class Membership Mixture Model

True values		Class membership probabilities											
		.60		.65		.70		.75		.80		.85	
		μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>
A_1	.72	.7201	.0206	.7203	.0197	.7201	.0188	.7204	.0181	.7204	.0174	.7203	.0165
C_1	.15	.1496	.0206	.1494	.0196	.1497	.0187	.1494	.0179	.1494	.0171	.1495	.0161
E_1	.13	.1303	.0032	.1302	.0030	.1302	.0028	.1303	.0027	.1303	.0026	.1302	.0025
A_2	.10	.0994	.0230	.0988	.0248	.0992	.0275	.0987	.0302	.0987	.0342	.0987	.0386
C_2	.58	.5806	.0198	.5810	.0214	.5806	.0239	.5813	.0260	.5814	.0295	.5813	.0344
E_2	.32	.3200	.0096	.3202	.0106	.3202	.0116	.3200	.0131	.3200	.0149	.3200	.0175
μ_1	.00	-.0007	.0141	-.0006	.0135	-.0005	.0128	-.0005	.0123	-.0006	.0118	-.0006	.0114
μ_2	3.00	2.9990	.0177	2.9989	.0196	2.9986	.0215	2.9985	.0241	2.9979	.0277	2.9973	.0342
$P(g)$.5998	.0049	.6499	.0048	.6999	.0046	.7500	.0045	.7999	.0042	.8499	.0039
p^1		.3551		.3535		.3465		.3467		.3482		.3502	
$p < .05$.16		.17		.16		.16		.15		.16	

True values		.60		.65		.70		.75		.80		.85	
		μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>
A_1	.72	.7216	.0313	.7209	.0209	.7204	.0189	.7214	.0261	.7210	.0169	.7214	.0224
C_1	.15	.1483	.0314	.1491	.0209	.1497	.0187	.1485	.0258	.1489	.0167	.1487	.0219
E_1	.13	.1301	.0043	.1301	.0029	.1299	.0027	.1301	.0036	.1301	.0024	.1299	.0032
A_2	.10	.0991	.0333	.0992	.0264	.0993	.0299	.0982	.0457	.1000	.0369	.0978	.0608
C_2	.58	.5807	.0270	.5806	.0213	.5801	.0241	.5814	.0374	.5795	.0314	.5804	.0523
E_2	.32	.3202	.0161	.3202	.0128	.3206	.0143	.3204	.0226	.3205	.0179	.3218	.0306
μ_1	.00	-.0021	.0350	-.0010	.0230	-.0004	.0206	-.0014	.0270	-.0008	.0172	-.0079	.0631
μ_2	1.50	1.4976	.0462	1.4987	.0370	1.4999	.0414	1.4969	.0667	1.5000	.0537	1.5058	.1119
$P(g)$.5985	.0241	.6492	.0171	.6997	.0164	.7485	.0226	.7993	.0146	.8365	.1025
p^1		.3493		.3465		.3936		.3425		.3383		.3480	
$p < .05$.16		.17		.14		.18		.18		.17	

Note: Shaded areas are simulated population parameters.

A_1 , C_1 , and E_1 additive genetic, shared and nonshared environmental standardized variance components for class 1.

A_2 , C_2 , and E_2 additive genetic, shared and nonshared environmental standardized variance components for class 2.

μ_1 and μ_2 class 1 and 2 means.

$P(g)$ probability of group (class) membership.

¹ Average p value based on 9 *df*.

$p < .05$ proportion of sample with p values less than .05.

using the THard option in Mx to generate random starting values for all parameters before refitting the model, we were able to obtain estimates consistent with those used to generate the data, simply by selecting the solution with the lowest log-likelihood.

Limitations

Our findings must be interpreted in the context of at least three potential limitations. First, our data were simulated under a standard normal distribution. Likewise, our probabilities of class membership were generated from random numbers between 0 and 1 with a uniform distribution. FMMs are capable of modeling skewed distributions (Bauer & Curran, 2003; Pearson, 1895; Titterington et al., 1985). However, unlike DF regression which is robust in the face of severe skewing of the data (Bishop, 2005; Lubke & Muthén, 2005), in such

instances it is possible that the observed ‘mixtures’ are simply capturing the skew rather than classes per se within a population.

Second, we also did not vary the size of the simulated data sets. It is quite likely that our ability to infer latent classes when based on smaller sample sizes, or with smaller group differences, will decline. Preliminary analyses suggest that this is likely especially with sample sizes less than 1000. Additional work is required to provide power estimates which are simultaneously a function of sample size, mean differences between classes, varying heterogeneity and class membership mechanisms. Lubke and Muthén (2005) have reported that multivariate Mahalanobis distances of 1.5 to 2 are required to obtain proportions of recovered parameters above .90. In the current study, the Mahalanobis distance is not a sufficient statistic to

Table 2c

Comparison of Simulated and Recovered Parameter Estimates Under the Nonshared Environmental (E) Class Membership Mixture Model

		Class membership probabilities											
		.60		.65		.70		.75		.80		.85	
True values		μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>
<i>A</i> ₁	.72	.7198	.0153	.7193	.0143	.7194	.0138	.7202	.0138	.7195	.0130	.7195	.0126
<i>C</i> ₁	.15	.1501	.0149	.1506	.0138	.1504	.0133	.1498	.0134	.1504	.0126	.1504	.0121
<i>E</i> ₁	.13	.1301	.0029	.1301	.0026	.1302	.0024	.1300	.0022	.1301	.0020	.1301	.0019
<i>A</i> ₂	.10	.1003	.0131	.1000	.0125	.1003	.0123	.1008	.0131	.1003	.0129	.1009	.0143
<i>C</i> ₂	.58	.5798	.0144	.5798	.0154	.5794	.0171	.5793	.0201	.5790	.0247	.5770	.0355
<i>E</i> ₂	.32	.3200	.0096	.3202	.0106	.3203	.0124	.3199	.0150	.3206	.0190	.3221	.0277
μ_1	.00	.0001	.0086	-.0002	.0082	-.0002	.0079	-.0002	.0077	-.0001	.0073	-.0001	.0072
μ_2	3.00	3.0000	.0098	2.9996	.0098	2.9996	.0104	2.9997	.0114	2.9997	.0117	2.9997	.0129
P(g)		.6000	.0029	.6500	.0028	.7000	.0027	.7500	.0025	.8001	.0022	.8501	.0020
<i>p</i> ¹		.2867		.3035		.3019		.2855		.2994		.2983	
<i>p</i> < .05		.24		.22		.21		.24		.21		.22	

		.60		.65		.70		.75		.80		.85	
True values		μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>
<i>A</i> ₁	.72	.7189	.0193	.7203	.0183	.7201	.0177	.7199	.7496	.7196	.0163	.7195	.0152
<i>C</i> ₁	.15	.1509	.0182	.1497	.0175	.1500	.0170	.1501	.0164	.1505	.0156	.1506	.0146
<i>E</i> ₁	.13	.1302	.0055	.1300	.0047	.1300	.0042	.1300	.0037	.1299	.0033	.1299	.0029
<i>A</i> ₂	.10	.1003	.0156	.1010	.0155	.1008	.0154	.1010	.0157	.1023	.0176	.1039	.0241
<i>C</i> ₂	.58	.5799	.0203	.5789	.0237	.5789	.0294	.5792	.0380	.5728	.0534	.5703	.0761
<i>E</i> ₂	.32	.3199	.0176	.3201	.0209	.3202	.0265	.3199	.0335	.3249	.0463	.3258	.0631
μ_1	.00	-.0001	.0138	-.0002	.0127	-.0002	.0113	-.0002	.0102	.0000	.0091	.0001	.0085
μ_2	1.50	1.5000	.0235	1.4990	.0250	1.4988	.0270	1.4988	.0298	1.4983	.0334	1.4989	.0396
P(g)		.6000	.0104	.6496	.0097	.6996	.0090	.7496	.0081	.7997	.0072	.8498	.0062
<i>p</i> ¹		.3225		.3783		.3752		.3077		.3053		.3033	
<i>p</i> < .05		.20		.13		.13		.18		.20		.21	

Note: Shaded areas are simulated population parameters.

*A*₁, *C*₁, and *E*₁ additive genetic, shared and nonshared environmental standardized variance components for class 1.

*A*₂, *C*₂, and *E*₂ additive genetic, shared and nonshared environmental standardized variance components for class 2.

μ_1 and μ_2 class 1 and 2 means.

P(g) probability of group (class) membership.

¹ Average *p* value based on 9 *df*.

p < .05 proportion of sample with *p* values less than .05.

describe the distance between the population means because the population covariance structures differ between the groups, and because there are effectively four groups (twins concordant high, concordant low, or discordant for group membership). However, under conditions of independence and assuming identical components of variance, which yields the maximum Mahalanobis distance for nonnegatively correlated traits, our mean differences of 1.5 for both twin pairs would yield a distance of 2.12; in practice the correlations are closer to .5 which gives a Mahalanobis distance of 1.73 between the concordant pair groups. This distance is therefore close to that in the Lubke and Muthén study. It is interesting to note that twin pairs may be more informative than singletons for identifying subgroups in the population.

Third, although our parameter coverage was good, it was not exhaustive. For example, under the ACE class membership, we held the probability of group membership constant while only adjusting the difference between the MZ and DZ twin pair correlations. With the QTL model, we only considered the case of a single dominant locus. Nevertheless, we modeled a likely range of class membership mechanisms, which provides a good starting point for further exploration of the parameter space.

Conclusion

Under certain conditions, it is possible to identify population heterogeneity based only on the observed trait scores of twin pairs. This heterogeneity may consist of mean differences between groups, and may feature different variance components as a function of group

Table 2d

Comparison of Simulated and Recovered Parameter Estimates Under the Genetic and Environmental (ACE) Class Membership Mixture Model

MZ twin pair correlations and thresholds for class membership (τ)													
		.60, $\tau = .3853$.65, $\tau = .3853$.70, $\tau = .3853$.75, $\tau = .3853$.80, $\tau = .3853$.85, $\tau = .3853$	
True values		μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>
A_1	.72	.7199	.0196	.7199	.0196	.7199	.0195	.7199	.0195	.7202	.0200	.7198	.0206
C_1	.15	.1502	.0190	.1502	.0190	.1502	.0189	.1502	.0189	.1500	.0196	.1503	.0199
E_1	.13	.1299	.0034	.1299	.0033	.1299	.0033	.1299	.0032	.1299	.0031	.1299	.0030
A_2	.10	.1013	.0183	.1014	.0185	.1015	.0187	.1014	.0191	.1009	.0198	.1005	.0209
C_2	.58	.5785	.0196	.5784	.0197	.5783	.0199	.5784	.0200	.5794	.0206	.5799	.0210
E_2	.32	.3202	.0112	.3202	.0110	.3202	.0107	.3202	.0104	.3196	.0103	.3196	.0100
μ_1	.00	-.0001	.0114	-.0001	.0114	.0000	.0114	.0000	.0114	-.0002	.0115	.0000	.0118
μ_2	3.00	3.0000	.0137	3.0001	.0139	3.0000	.0139	3.0001	.0141	2.9994	.0144	3.0001	.0142
τ		.3855	.0088	.3855	.0089	.3856	.0090	.3856	.0090	.3852	.0091	.3853	.0091
p^{\dagger}		.5527		.5520		.5544		.5530		.5524		.5613	
$p < .05$.05		.04		.04		.04		.04		.04	
		.60, $\tau = .3853$.65, $\tau = .3853$.70, $\tau = .3853$.75, $\tau = .3853$.80, $\tau = .3853$.85, $\tau = .3853$	
True values		μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>	μ	<i>SD</i>
A_1	.72	.7189	.0384	.7189	.0384	.7189	.0385	.7188	.0387	.7187	.0390	.7185	.0400
C_1	.15	.1508	.0378	.1509	.0377	.1509	.0378	.1510	.0381	.1511	.0384	.1514	.0395
E_1	.13	.1302	.0047	.1302	.0046	.1302	.0045	.1302	.0044	.1302	.0043	.1301	.0042
A_2	.10	.1026	.0349	.1028	.0353	.1028	.0356	.1028	.0367	.1031	.0386	.1033	.0406
C_2	.58	.5763	.0392	.5761	.0389	.5763	.0384	.5765	.0388	.5763	.0401	.5768	.0415
E_2	.32	.3212	.0219	.3211	.0209	.3209	.0198	.3207	.0189	.3206	.0181	.3198	.0175
μ_1	.00	-.0001	.0171	-.0003	.0171	-.0002	.0172	-.0001	.0171	-.0002	.0171	-.0001	.0176
μ_2	1.50	1.5002	.0244	1.5000	.0246	1.5001	.0246	1.5001	.0248	1.4999	.0248	1.4997	.0251
τ		.3851	.0283	.3849	.0286	.3851	.0288	.3852	.0289	.3850	.0293	.3851	.0295
p^{\dagger}		.3835		.3874		.3932		.3946		.3975		.4006	
$p < .05$.15		.14		.14		.15		.15		.15	

Note: Shaded areas are simulated population parameters.

A_1 , C_1 , and E_1 additive genetic, shared and nonshared environmental standardized variance components for class 1.

A_2 , C_2 , and E_2 additive genetic, shared and nonshared environmental standardized variance components for class 2.

μ_1 and μ_2 class 1 and 2 means.

\dagger Average p value based on 9 *df*.

$p < .05$ Proportion of sample with p values less than .05.

*The ALLINT function was used to estimate cell frequencies for MZ twin pairs using twin pair correlations (ranging .60 to .85 by .05) each with a group membership probability of .65 ($\tau = .3853$). The DZ twin pair correlations was held constant at .40 with a class membership probability of .65 ($\tau = .3853$).

membership. Depending on the mechanism that determines group membership, this change of the effect of the variance components may reflect $G \times G$, $G \times C$, $G \times E$ or $G \times QTL$ interaction. Although we have confined our exploration to a fairly narrow region of the parameter space, and only to continuous data, the method shows some potential to detect latent $G \times E$ effects.

The practical utility of FMM for the quantitative genetic analysis of twin data remains open to question. On the whole, we would not envisage a standard approach to test for heterogeneity by FMM, and then only proceeding to the nonmixture (i.e., standard) variance component analysis when the hypothesis of a mixture is not supported. Rather, these methods would seem most appropriate when there exists an a priori hypothesis that the population is not homogeneous — as was the case for the Turkheimer et al. situation. In

other cases, distributional evidence might make the method attractive. Such evidence does, however, raise the issue of scaling. It is well known that, for example, skewness can be interpreted as a mixture distribution (Bauer & Curran, 2003; Pearson, 1895) — and indeed vice versa. This source of potential bias is problematic for measurement of behavioral traits or other traits where observed scores consist of summaries of items assessed at the ordinal level. In this case, it would seem more appropriate to attempt mixture distribution modeling at the level of the latent trait (or factor), while modeling the relationship between the factor and the items at their original level of measurement. Although presently, such modeling would present some practical difficulties, it seems likely that these will be overcome in the near future.

Table 2e

Comparison of Simulated and Recovered Parameter Estimates Under the QTL Dominant Class Membership Mixture Model

		Allele frequency (p/q)											
		.60 / .40		.65 / .35		.70 / .30		.75 / .25		.80 / .20		.85 / .15	
True values		μ	SD	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD
A_1	.72	.7202	.0127	.7202	.0122	.7202	.0119	.7206	.0110	.7201	.0108	.7192	.0109
C_1	.15	.1498	.0123	.1497	.0118	.1497	.0114	.1494	.0105	.1499	.0103	.1507	.0102
E_1	.13	.1300	.0019	.1300	.0018	.1300	.0017	.1300	.0017	.1301	.0017	.1301	.0016
A_2	.10	.1005	.0299	.1028	.0332	.1052	.0419	.1078	.0503	.1128	.0767	.1280	.1220
C_2	.58	.5798	.0302	.5776	.0350	.5750	.0459	.5731	.0540	.5686	.0810	.5538	.1282
E_2	.32	.3197	.0109	.3196	.0124	.3199	.0155	.3192	.0181	.3186	.0245	.3181	.0337
μ_1	.00	.0003	.0229	.0007	.0266	.0009	.0330	.0018	.0415	.0053	.0581	.0101	.0848
μ_2	3.00	3.0001	.0088	3.0000	.0087	3.0001	.0086	3.0001	.0081	3.0002	.0083	3.0003	.0081
p^a		.5998	.0037	.6498	.0038	.6998	.0041	.7498	.0045	.7993	.0050	.8493	.0059
p^1		.3597		.3596		.3648		.3720		.3472		.3526	
$p < .05$.17		.16		.15		.15		.18		.18	

		.60 / .40		.65 / .35		.70 / .30		.75 / .25		.80 / .20		.85 / .15	
True values		μ	SD	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD
A_1	.72	.7207	.0144	.7213	.0140	.7192	.0190	.7208	.0119	.7201	.0121	.7194	.0129
C_1	.15	.1493	.0136	.1487	.0133	.1501	.0152	.1493	.0113	.1500	.0113	.1507	.0120
E_1	.13	.1299	.0023	.1300	.0022	.1307	.0121	.1299	.0019	.1299	.0019	.1299	.0021
A_2	.10	.1062	.0488	.1101	.0612	.1160	.0895	.1286	.1082	.1403	.1381	.1564	.1647
C_2	.58	.5745	.0488	.5704	.0651	.5652	.0878	.5525	.1161	.5430	.1441	.5325	.1774
E_2	.32	.3193	.0188	.3195	.0218	.3188	.0279	.3189	.0348	.3167	.0481	.3111	.0705
μ_1	.00	.0049	.0554	.0041	.0676	.0135	.1268	.0104	.1166	.0310	.1730	.0831	.2902
μ_2	1.50	1.5010	.0169	1.5009	.0156	1.4942	.1031	1.5015	.0130	1.5018	.0133	1.5028	.0138
p^a		.5988	.0168	.6488	.0181	.6950	.0480	.7468	.0230	.7932	.0295	.8360	.0423
p^1		.3754		.3716		.3643		.3654		.3528		.3626	
$p < .05$.15		.15		.18		.17		.18		.19	

Note: Shaded areas are the simulated allele frequencies (p/q) and population parameters.

$A_1, C_1,$ and E_1 additive genetic, shared and nonshared environmental standardized variance components for class 1.

$A_2, C_2,$ and E_2 additive genetic, shared and nonshared environmental standardized variance components for class 2.

μ_1 and μ_2 class 1 and 2 means.

p^a Frequency of p allele.

¹ Average p value based on 9 *df*.

$p < .05$ Proportion of sample with p values less than .05.

Acknowledgments

The project was supported by NIH grants MH-65322 and DA-18673 and NHMRC Sidney Sax Postdoctoral Fellowship.

References

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods, 8*, 338–363.

Bishop, D. V. (2005). DeFries-Fulker analysis of twin data with skewed distributions: Cautions and recommendations from a study of children’s use of verb inflections. *Behavior Genetics, 35*, 479–490.

Boomsma, D. I., de Geus, E. J., van Baal, G. C., & Koopmans, J. R. (1999). A religious upbringing reduces

the influence of genetic factors on disinhibition: Evidence for interaction between genotype and environment on personality. *Twin Research, 2*, 115–125.

Brotherstone, S., & Hill, W. G. (1986). Heterogeneity of variances amongst herds for milk production. *Animal Production, 48*, 283–291.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A., & Poulton, R. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science, 301*, 386–389.

DeFries, J. C., & Fulker, D. W. (1985). Multiple regression analysis of twin data. *Behavior Genetics, 15*, 467–473.

Eaves, L. J., & Sullivan, P. (2001). Genotype-environment interaction in transmission disequilibrium tests. *Advances in Genetics, 42*, 223–240.

- Foulley, J. L., & Quaas, R. L. (1995). Heterogeneous variances in Gaussian linear mixed models. *Genetics, Selection, Evolution*, 27, 211–228.
- Gianola, D., Foulley, J. L., Fernando, R. L., Henderson, C. R., & Weigel, K. A. (1992). Estimation of heterogeneous variances using empirical Bayes methods: Theoretical considerations. *Journal of Dairy Science*, 75, 2805–2823.
- Heath, A. C., Eaves, L. J., & Martin, N. G. (1998). Interaction of marital status and genetic risk for symptoms of depression. *Twin Research*, 1, 119–122.
- Kendler, K. S., & Eaves, L. J. (1986). Models for the joint effect of genotype and environment on liability to psychiatric illness. *American Journal of Psychiatry*, 143, 279–289.
- Kendler, K. S., & Karkowski-Shuman, L. (1997). Stressful life events and genetic liability to major depression: Genetic control of exposure to the environment? *Psychological Medicine*, 27, 539–547.
- Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1994). Parental treatment and the equal environment assumption in twin studies of psychiatric illness. *Psychological Medicine*, 24, 579–590.
- LaBuda, M. C., DeFries, J. C., & Fulker, D. W. (1986). Multiple regression analysis of twin data obtained from selected samples. *Genetic Epidemiology*, 3, 425–433.
- Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Houghton Mifflin: Boston.
- Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, 10, 21–39.
- Martin, N. G., Eaves, L. J., & Heath, A. C. (1987). Prospects for detecting genotype X environment interactions in twins with breast cancer. *Acta Geneticae Medicae et Gemellologiae*, 36, 5–20.
- Martin, N. G., & Oakeshott, J. G. (1983). Is Pi a selectively balanced polymorphism? *Human Heredity*, 33, 24–28.
- McLachlan, G. J., & Basford, K. E. (1988). *Mixture models: Inference and applications to clustering*. New York: Marcel Dekker.
- Neale, M. C. (1999). *Mx: Statistical modeling* (5th ed.). Box 126 MCV, Richmond, VA 23298: Department of Psychiatry.
- Neale, M. C. (2003). A finite mixture distribution model for data collected from twins. *Twin Research*, 6, 235–239.
- Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publishers.
- Neale, M. C., Cherny, S. S., Sham, P. C., Whiteld, J. B., Heath, A. C., Birley, A. C., & Martin, N. G. (1999). Distinguishing population stratification from genuine allelic effects with Mx: Association of ADH2 with alcohol consumption. *Behavior Genetics*, 29, 233–244.
- Pearson, K. (1895). Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London A*, 343–414.
- Purcell, S., & Sham, P. (2002). Variance components models for gene–environment interaction in quantitative trait locus linkage analysis. *Twin Research*, 5, 572–576.
- Redner, R. A., & Walker, H. F. (1984). Mixture densities, maximum likelihood, and the EM algorithm. *SIAM Review*, 26, 195–239.
- Titterton, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical analysis of finite mixture distributions*. Chichester: John Wiley & Sons.
- Turkheimer, E., Haley, A., Waldron, M., D’Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 14, 623–628.
- Vandenberg, S. G. (1984). Does a special twin situation contribute to similarity for abilities in MZ and DZ twins? *Acta Geneticae Medicae et Gemellologiae*, 33, 219–222.