

Variance Components Models for Gene–Environment Interaction in Twin Analysis

Shaun Purcell

Social, Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, King's College, London, UK

Gene–environment interaction is likely to be a common and important source of variation for complex behavioral traits. Often conceptualized as the genetic control of sensitivity to the environment, it can be incorporated in variance components twin analyses by partitioning genetic effects into a mean part, which is independent of the environment, and a part that is a linear function of the environment. The model allows for one or more environmental moderator variables (that possibly interact with each other) that may i) be continuous or binary ii) differ between twins within a pair iii) interact with residual environmental as well as genetic effects iv) have nonlinear moderating properties v) show scalar (different magnitudes) or qualitative (different genes) interactions vi) be correlated with genetic effects acting upon the trait, to allow for a test of gene–environment interaction in the presence of gene–environment correlation. Aspects and applications of a class of models are explored by simulation, in the context of both individual differences twin analysis and, in a companion paper (Purcell & Sham, 2002) sibpair quantitative trait locus linkage analysis. As well as elucidating environmental pathways, consideration of gene–environment interaction in quantitative and molecular studies will potentially direct and enhance gene-mapping efforts.

Simple quantitative genetic models average over any group differences within a population. The presence of gene–environment interaction ($G \times E$) will mean that a single statistic is no longer adequate to describe a whole population, as genetic effects will now depend on individuals' environments. A heritability of 50%, for example, might equally entail scenario S_1 where, for all individuals, differences in the trait are equally due to genetic and environmental factors or scenario S_2 where, for half the population, the trait is completely genetically determined, whereas for the other half the trait is completely environmentally determined. In the context of twin analysis, consideration of $G \times E$ aims to distinguish between scenarios such as S_1 and S_2 . This requires that the E component of the $G \times E$ is a measured variable that indexes the differential etiologies present in S_2 . For example, if the 50:50 split reflected males and females, this would represent a $G \times \text{sex}$ interaction.

It is possible to detect $G \times E$ within various study designs (Heath et al., 2002); G and E can be either latent variables (e.g., additive genetic variance, shared environment) or measured variables (e.g., DRD4 genotype, age, place of residence). When both G and E are latent variables, it is possible to detect $G \times E$ as a heteroscedastic

bivariate twin distribution, where twin pair difference correlates with twin pair sum (Jinks & Fulker, 1970). However, as well as suffering from low power, this test also is sensitive to non-normality in the trait. More importantly, beyond indicating that *some* form of interaction is occurring, it sheds no light on underlying processes. Having both G and E as measured variables provides the most power for detecting $G \times E$; the results will potentially be very informative also, beginning to map onto the underlying biology. For example, sex moderates the effect of the *APOE* $\epsilon 4$ allele on cognitive decline, where women show higher $\epsilon 4$ -associated risk than men (Yaffe et al., 2000). Additionally, the $\epsilon 4$ allele moderates the impact of estrogen in women on cognitive decline, as the estrogen use is associated with less cognitive decline only in women without the $\epsilon 4$ risk allele.

In the present paper we consider the case of latent $G \times$ measured E , which is most relevant to the classical twin study. For example, additive genetic effects on depression symptoms interact with marital status in women, where unmarried women show greater levels of genetic influence (Heath et al., 1998). Another example of latent $G \times$ measured E is that a religious upbringing seems to attenuate genetic influences on the personality trait of disinhibition (Boomsma et al., 1999). Testing for $G \times E$ with a binary moderator such as marital status is straightforward. The parameters of interest (e.g., a^2 , c^2 and e^2) are estimated for “exposed” and “unexposed” individuals separately. A test of $G \times E$ is given by equating the parameters across exposure group and observing the associated decline in model fit (i.e., testing for heterogeneity).

The concepts of $G \times E$ apply equally to quantitative trait locus (QTL) studies as well as twin studies. A companion paper (Purcell & Sham, 2002) looks at two ways in which interactions can be incorporated into the variance components approach to sib-pair linkage analysis: testing for environmentally-moderated QTL effects and utilizing information about any residual interactions (i.e., not involving the QTL) detected prior to analysis.

Address for correspondence: Shaun Purcell, SGDP Research Centre, IoP, Box PO 080, 111 Denmark Hill, London, SE5 8AF. Email: s.purcell@iop.kcl.ac.uk

Continuous Moderator Variables

Complex human traits are often best defined in quantitative terms, to avoid the potential loss in power associated with artificial dichotomization of a continuous variable. Many potential moderator variables are also most naturally defined in quantitative terms: some obvious examples include age, gestational age, socioeconomic status, educational level, consumption of food, drugs or alcohol. Although typical approaches to $G \times E$ are often limited to binary moderators, it is equally possible to allow for continuous moderating variables that may differ between twins in a pair.

The most basic $G \times E$ interaction involving a continuous moderating E variable implies that genetic effects increase or decrease as a linear function of the moderator. Although this formulation covers a large class of $G \times E$, a second nonlinear class is also considered, where genetic effects may, for example, be attenuated at extreme high and extreme low levels of a moderator.

Gene–Environment Correlation

$G \times E$ is often conceptualized as genetic control of sensitivity to different environments. A related phenomenon, $G-E$ correlation (r_{GE}) represents genetic control of exposure to different environments (Kendler & Eaves, 1986). Equivalently, of course, $G \times E$ is the environmental control of differential gene effects, whereas r_{GE} is the environmental control of gene frequency. A recent example of r_{GE} showed that genetic influences on alcohol and drug misuse are correlated with various aspects of the family and school environment (Jang et al., 2001) and we might expect r_{GE} to feature in many other complex traits. Typical approaches to $G \times E$ in twin analyses involving stratification of a sample by the environmental moderator variable (Neale & Cardon, 1992) have been unable to disentangle $G \times E$ and r_{GE} in a single analysis, however. For example, if individuals in a certain environment show greater genetic influence, this could be due to either (1) the environment modifying the effects of certain genes or (2) certain trait-influencing genes being more likely to be present in that environment. As described below, the present method is able to discriminate between these alternatives and to allow analysis of $G \times E$ in the presence of r_{GE} .

Unmodelled $G \times E$ and r_{GE}

Before considering the modelling of $G \times E$ it is worth reviewing the impact of $G \times E$ and r_{GE} on standard twin models, in terms of biased parameter estimates. In short, interaction between A and C acts like A ; interaction between A and E acts like E . Correlation between A and C acts like C ; correlation between A and E acts like A . For example, in the case of $A \times C$ interaction, if a standardized trait $T = aA + cC + iAC + eE$ then the expected variance is $Var(T) = a^2 + c^2 + i^2 + e^2$, assuming that the latent variables A , C and E have unit variance. The expected twin covariances are

$$\begin{aligned} Cov(T_1, T_2) &= a^2 Cov(A_1, A_2) + c^2 Cov(C_1, C_2) \\ &\quad + e^2 Cov(E_1, E_2) + i^2 Cov(A_1 C_1, A_2 C_2) \\ &= a^2 + c^2 + i^2 \text{ for MZ twins} \\ &= a^2/2 + c^2 + i^2/2 \text{ for DZ twins} \end{aligned}$$

as $Cov(A_1, A_2)$ is 1 for MZ twins, 0.5 for DZ twins; $Cov(C_1, C_2) = 1$ and $Cov(E_1, E_2) = 0$ for all twins; also $Cov(A_1 C_1, A_2 C_2) = Cov(A_1, A_2) Cov(C_1, C_2) = Cov(A_1, A_2)$. Similar covariance algebra can show that $A \times E$ interaction contributes to the E component. If A is correlated with (rather than interacting with) an environmental variable, say C , with correlation r_{AC} then the expected trait variance is $Var(T) = a^2 + c^2 + 2ac \times r_{AC} + e^2$ and the expected twin covariances are

$$\begin{aligned} Cov(T_1, T_2) &= a^2 Cov(A_1, A_2) + c^2 Cov(C_1, C_2) \\ &\quad + e^2 Cov(E_1, E_2) + ac Cov(A_1, C_2) + ac Cov(A_2, C_1) \\ &= a^2 + c^2 + 2ac \times r_{AC} \text{ for MZ twins} \\ &= a^2/2 + c^2 + 2ac \times r_{AC} \text{ for DZ twins} \end{aligned}$$

as $Cov(A_1, C_2) = Cov(A_2, C_1) = r_{AC}$. Similarly, if A and E are non-independent then

$$\begin{aligned} Cov(T_1, T_2) &= a^2 + c^2 + 2ae \times r_{AE} \text{ for MZ twins} \\ &= a^2/2 + c^2 + ae \times r_{AE} \text{ for DZ twins} \end{aligned}$$

Current Aims

The initial framework for the analysis of $G \times E$ in the context of the twin study has existed for some time. For example, Martin et al. (1987) describe a model to handle continuous moderator variables and interactions between both latent and measured genetic and environmental effects, as well as documenting the power of such tests. The present study aims to extend this seminal work and to provide more comprehensive simulation results. This should help in interpreting the results of continuous moderator models now appearing in the literature (e.g., Mustanski et al., 2002; Wichers et al., in press). This paper is split into five main sections: (1) the basic $G \times E$ model involving a continuous moderator variable that can interact with latent genetic and environmental effects (2) nonlinear $G \times E$ using a quadratic approximation (3) $G \times E$ in the presence of r_{GE} (4) scalar (different magnitudes) and qualitative (different genes) interactions (5) the impact of distributional factors on $G \times E$ analysis.

This paper introduces some notation in order to clarify different moderating effects. Standard $G \times E$ will be called $A \times M$: the G is replaced by A to refer specifically to additive genetic effects; E is replaced by M (moderator), to distinguish it from the latent nonshared twin environment. Other types of interaction are $C \times M$ and $E \times M$, where the latent shared and nonshared environments, respectively, interact with a measured moderator and, in the companion paper, $Q \times M$ interaction, where a specific QTL interacts with a moderator. The term $G \times E$ will still be used to refer to the whole class of these effects.

$G \times E$ with Continuous Moderator Variables

A naive treatment of continuous moderation might proceed as follows: stratify the sample into a number of groups on the basis of the moderator, calculate heritability within each strata, equate parameters across strata or test

for a linear trend in heritability across strata. There are, however, several problems with such an approach. Firstly, the stratification procedure will effectively reduce the sample size, especially if the moderator is not shared between twins. Secondly, the use of heritability essentially assumes equal variance across strata, whereas what is of interest is whether the absolute magnitude of genetic effects changes, not only the proportion. Thirdly, although it is logical to initially assume a linear $G \times E$ interaction, this linearity should be at the level of *effect* rather than the level of *variance component*, as variance is a second-order statistic.

Consider the basic biometric model for a hypothetical additive biallelic trait locus, with additive genetic value a and increaser allele frequency p . The locus' contribution to the variance, $2p(1-p)a^2$, is a function of both the square of magnitude of effect and how common it is. A linear $A \times M$ interaction implies that the additive genetic value is a linear function of the moderator M , namely $a + \beta M$ where β is an unknown parameter to be estimated. If β is significantly non-zero, this is evidence of a $A \times M$ interaction. The contribution to the variance is $2p(1-p)(a + \beta M)^2$, indicating that variance is a quadratic function of the moderator under linear interaction. Figure 1 illustrates a linear interaction effect for a single hypothetical QTL.

This hypothetical QTL model directly translates into the twin model. Path coefficients represent the magnitude of effect and so we express the path coefficients as linear functions of a moderator. In other words, the additive genetic path coefficient is no longer a , it is now $a + \beta_x M$. Therefore, if β_x is significantly non-zero, this represents an $A \times M$ interaction. The moderator may be obligatorily shared or it can be specified separately for each twin (e.g., age and parental income are obligatorily shared; weight and exposure to violence are not). Binary moderators can be coded as "0/1", in which case the model reduces to the standard "stratify by environment" approach.

Any variable which has a moderating, or interactive, effect on a trait may also have a mediating, or main, effect. Therefore, the moderator can also be entered in the means

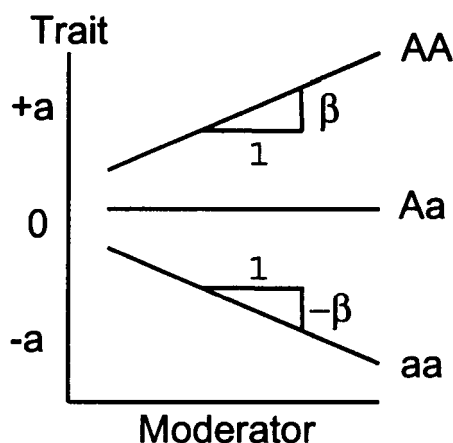


Figure 1

The biometric model incorporating linear $A \times M$ interaction; the coefficient β assess the extent of interaction.

model, where the parameter β_M represents the standard phenotypic regression coefficient. Additionally, we allow for $C \times M$ and $E \times M$ interaction: that is, of the measured moderator with either the residual latent shared or nonshared environmental variables, assessed by β_Y and β_Z respectively.

For each twin pair conditional on the twins' moderator M , the expected trait mean for twin i is $\mu + \beta_M M_i$ and the expected trait variance is

$$\text{Var}(T_i) = (a + \beta_x M_i)^2 + (c + \beta_Y M_i)^2 + (e + \beta_Z M_i)^2$$

for $i = 1, 2$. The expected MZ covariance is

$$\text{Cov}_{\text{MZ}}(T_1, T_2) = (a + \beta_x M_1)(a + \beta_x M_2) + (c + \beta_Y M_1)(c + \beta_Y M_2)$$

the expected DZ covariance is

$$\text{Cov}_{\text{DZ}}(T_1, T_2) = 0.5(a + \beta_x M_1)(a + \beta_x M_2) + (c + \beta_Y M_1)(c + \beta_Y M_2)$$

This is equivalent to the model used by Martin et al. (1987), in which, for example, variance due to additive genetic effects and $G \times E$ is $a^2 \cdot (1 + \beta M)^2$ (i.e., in the current formulation, their interaction coefficient is β_x/a).

Seven parameters (unmoderated components a , c and e ; moderated components β_x , β_Y and β_Z ; main effect β_M) are now estimated under the full model, *ACE-XYZ-M*. Figure 2 shows a partial path diagram representing the *ACE-XYZ-M* model. The best-fitting model can be obtained by successively dropping either moderating, main effect and/or unmoderated components. Assuming that at least one moderated parameter remains estimated in the model, the results must be considered in the context of a sensible range of moderator variable values. The expected variance components representing additive genetic, shared environmental and nonshared environmental effects can be plotted as a function of M . For example, the additive genetic component is $(a + \beta_x M)^2$ for a sensible range of M . Clearly, to extrapolate beyond the range of M observed in the data could be misleading: an approach to a clearer visualization of moderated variance components is outlined further below. Table 1 presents a simple example of calculating the additive genetic variance and heritability given parameter values for the *ACE-XYZ-M* model — this indicates what an interaction of $\beta_x = 0.2$ (a value subsequently used in many of the simulations) actually "looks like".

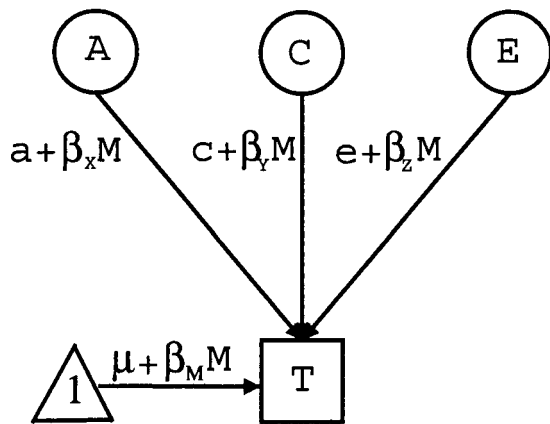
An Example

A normally-distributed trait was simulated with A , C and E components representing 25%, 25% and 50% of the trait variance respectively. In addition, an obligatorily shared moderator variable was created with (1) a substantial main effect on the trait and (2) a marked interactive effect on the A component of the trait. The $A \times M$ was such that genetic effects were attenuated at intermediate values of the moderator but exaggerated at extreme high or extreme low values.

Table 1

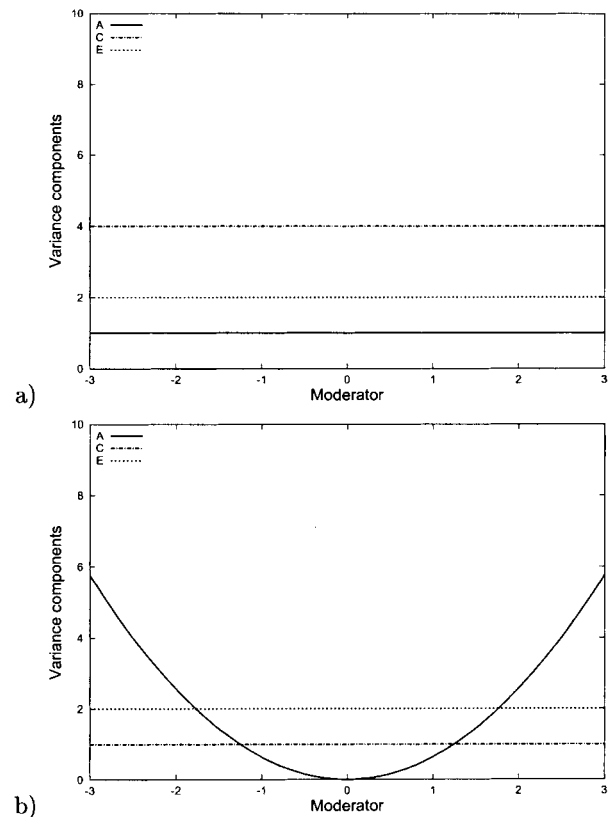
An $A \times M$ Interaction: Additive Genetic Variance and Heritability is Tabulated Against Different Values of the Moderating Variable. Parameter Values are $a = c = e = 1$ and $\beta_x = 0.2$, $\beta_y = \beta_z = 0$

SD from mean	$(a + \beta_x M)$	$(a + \beta_x M)^2$	$Var(T)$	h^2
-3	0.4	0.16	2.16	0.07
-2	0.6	0.36	2.36	0.15
-1	0.8	0.64	2.64	0.24
0	1.0	1.00	3.00	0.33
1	1.2	1.44	3.44	0.42
2	1.4	1.96	3.96	0.49
3	1.6	2.56	4.56	0.56


Figure 2

Partial path diagram for the $ACE\text{-}XYZ\text{-}M$ model, shown for one twin only. Latent variables have unit variance.

Fitting the various models starting with the full $ACE\text{-}XYZ\text{-}M$ model, the best-fitting model was the $ACE\text{-}X\text{-}M$ model, which correctly represents the simulation procedure described above. Figure 3 represents the variance components under the basic ACE model (i.e., equivalent to looking only at the trait and completely ignoring the moderator) and the best-fit model $ACE\text{-}X\text{-}M$. The signatures of mediation and moderation are clearly visible. Note that under the ACE model the C component is much greater than the A component, even though both residual components were simulated to account for 25% of the variance, because C includes the variance due to the main effect of the moderator (the moderator was obligatorily shared between twins). When the main effect is explicitly accounted for by the M component in the best-fitting $ACE\text{-}X\text{-}M$ model, the C component drops to the appropriate level (i.e., half of E). As the β_x parameter is nonzero in the best-fitting model, we observe that the additive genetic variance varies as a function of the moderator, in a way which directly corresponds to the simulated properties (i.e., no genetic effects at intermediate levels of the moderator, exaggerated genetic effects at extreme values of the moderator).


Figure 3

Modelling moderating and main effects: a) ACE model
b) $ACE\text{-}X\text{-}M$ model.

If one were to standardize the variance components (for example, by plotting

$$\frac{(a + \beta_x M)^2}{(a + \beta_x M)^2 + (c + \beta_y M)^2 + (e + \beta_z M)^2}$$

the results will indicate proportions of variance. In the current example, genetic influences increase at extreme values of the moderator whilst environmental effects are constant. Proportionally, however, the environmental effects necessarily get smaller at extreme values, relative to genetic effects, as illustrated in Figure 4 which plots the

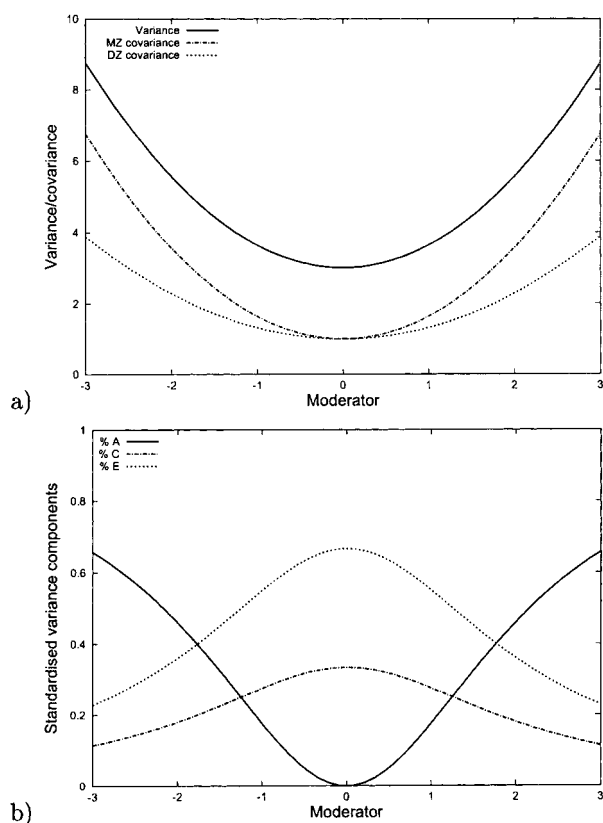


Figure 4

The impact of standardization: a) the expected variance and twin covariance as a function of a moderator; in this example, the total trait variance differs across the range of the moderator; b) the standardized variance components.

expected variance and twin covariances as well as the standardized variance components for the current example. Arguably, this is somewhat misleading, and plotting only unstandardized results is encouraged.

Further Simulations

A more comprehensive set of simulations was conducted in order to explore some of the properties of this model. A moderately large sample size of 500 MZ pairs and 500 DZ pairs was used under all models. Twin data were simulated for a continuous, normally-distributed trait and moderator variable. In all cases, the unmoderated parameter values were set at $a = c = e = 1$ (to give a variance of 3 excluding moderating and main effects). Table 2 gives the average parameter estimates and fit statistics for several conditions. Data were simulated under three true model conditions; each condition was replicated 50 times; each replicate was analysed under 8 nested models. The three true models were *ACE-X*, *ACE-Y* and *ACE-Z*, representing $A \times M$, $C \times M$ and $E \times M$ interactions. The β interaction coefficients were either set at 0 (if not in the model) or 0.2. In all cases the moderator variable was set to have a twin correlation (for both MZ and DZ twins) of 0.5. Results not shown here indicate a very similar pattern for other values, including 0 and 1 for these simulations (also, the next set of simulations varies the moderator twin correlation). No main effects of the moderator are simulated or included in the model in this first set of simulations.

Table 2 shows the average best-fit parameter estimates as well as the averaged minus twice log-likelihood of the data and AIC index. The last column, “% selected” refers to the percentage of the 50 replicates that were selected

Table 2

Average Parameter Estimates and Fit Statistics for Twin Models of Linear $A \times M$, $C \times M$ and $E \times M$ Interaction

True model	Analytic model	β_x	β_y	β_z	-2LL	AIC	% selected
<i>ACE-X</i>	<i>ACE-XYZ</i>	0.18	0.01	0.01	7433.95	-552.05	4
	<i>ACE-YZ</i>	—	0.16	0.04	7438.00	-550.00	6
	<i>ACE-XZ</i>	0.20	—	0.00	7434.88	-553.13	14
	<i>ACE-XY</i>	0.20	0.00	—	7435.08	-552.92	8
	<i>ACE-X</i>	0.20	—	—	7435.99	-554.01	62
	<i>ACE-Y</i>	—	0.18	—	7440.42	-549.58	6
	<i>ACE-Z</i>	—	—	0.07	7447.48	-542.52	0
	<i>ACE</i>	—	—	—	7455.44	-536.56	0
<i>ACE-Y</i>	<i>ACE-XYZ</i>	0.05	0.16	0.00	7436.76	-549.24	2
	<i>ACE-YZ</i>	—	0.21	0.00	7437.95	-550.05	10
	<i>ACE-XZ</i>	0.20	—	-0.02	7440.31	-547.69	6
	<i>ACE-XY</i>	0.04	0.16	—	7437.63	-550.37	6
	<i>ACE-X</i>	0.18	—	—	7441.45	-548.55	18
	<i>ACE-Y</i>	—	0.21	—	7438.85	-551.15	58
	<i>ACE-Z</i>	—	—	0.05	7453.59	-536.41	0
	<i>ACE</i>	—	—	—	7457.70	-534.30	0
<i>ACE-Z</i>	<i>ACE-XYZ</i>	-0.03	0.03	0.21	7415.41	-570.59	6
	<i>ACE-YZ</i>	—	0.00	0.21	7416.57	-571.43	12
	<i>ACE-XZ</i>	-0.01	—	0.21	7416.55	-571.45	4
	<i>ACE-XY</i>	0.26	-0.06	—	7452.27	-535.73	0
	<i>ACE-X</i>	0.21	—	—	7453.87	-536.13	0
	<i>ACE-Y</i>	—	0.15	—	7462.24	-527.76	0
	<i>ACE-Z</i>	—	—	0.21	7417.56	-572.44	78
	<i>ACE</i>	—	—	—	7472.70	-519.30	0

from the 8 analytic models on the basis of AIC. The parameter estimates for the mean and unmoderated parameters are not shown: they were all very close to simulated values.

Under the full *ACE-XYZ* model, the model generally recovers the interaction parameters quite well. For example, for data simulated under the *ACE-X* model, the average values of β_x , β_y and β_z are 0.18, 0.01 and 0.01 (i.e., true values 0.20, 0.00 and 0.00). In this case, the average $-2LL$ is 7433.95, whereas under the *ACE-X* model it is 7435.99. The average difference, from dropping the *Y* and *Z* components is only 2.04, which is not significant for a χ^2_2 at $\alpha = 0.05$, suggesting that these terms can be dropped from the model.

There is clearly an issue of specificity here, however. For example, note that the β_y coefficient of the *ACE-Y* model is 0.18 even when the data were simulated under the *ACE-X* model. That is, these interaction parameters are quite highly correlated (as are *a* and *c*), which can lead to some reduction in power to detect one in the presence of the other. However, on the basis of lowest AIC, the correct model was selected the majority of the time under all three conditions (62%, 58% and 78%). Typically, the second most selected model contained the correct interaction term, also. In none of the cases was the basic *ACE* model with no interaction terms supported.

Figure 5 illustrates the relationship between variance components and the expectations for twin variance and covariance as a function of a moderator variable. Quite different patterns of interaction can give rise to quite similar patterns of variance and covariance. Given that analysis moves from the observed variances and covariances to the inferred variance components, the relative indistinguishability of the models is unsurprising. Generally, however, the parameter estimates under the full model can serve as a guide to the true model. One strategy, therefore, is to plot the variance components using the parameter estimates of the full model — the general outline of this plot should not change greatly under nested submodels. This does highlight the danger of only testing for $A \times M$ interaction within this framework, however. For example, for data simulated under the *ACE-Y* model, the average difference in $-2LL$ between the *ACE-X* and *ACE* models is 16.25, which is highly significant for 1 degree of freedom.

Several other properties are explored in the next set of simulations, the results of which are given in Table 3. Data were simulated under four models, and under different twin correlations for the moderator variable ($r = 0, 0.5$ and 1). The first model is simply the null model with no moderating or mediating effects. The second model represents combined interactive effects, with $\beta_x = \beta_y = \beta_z = 0.2$.

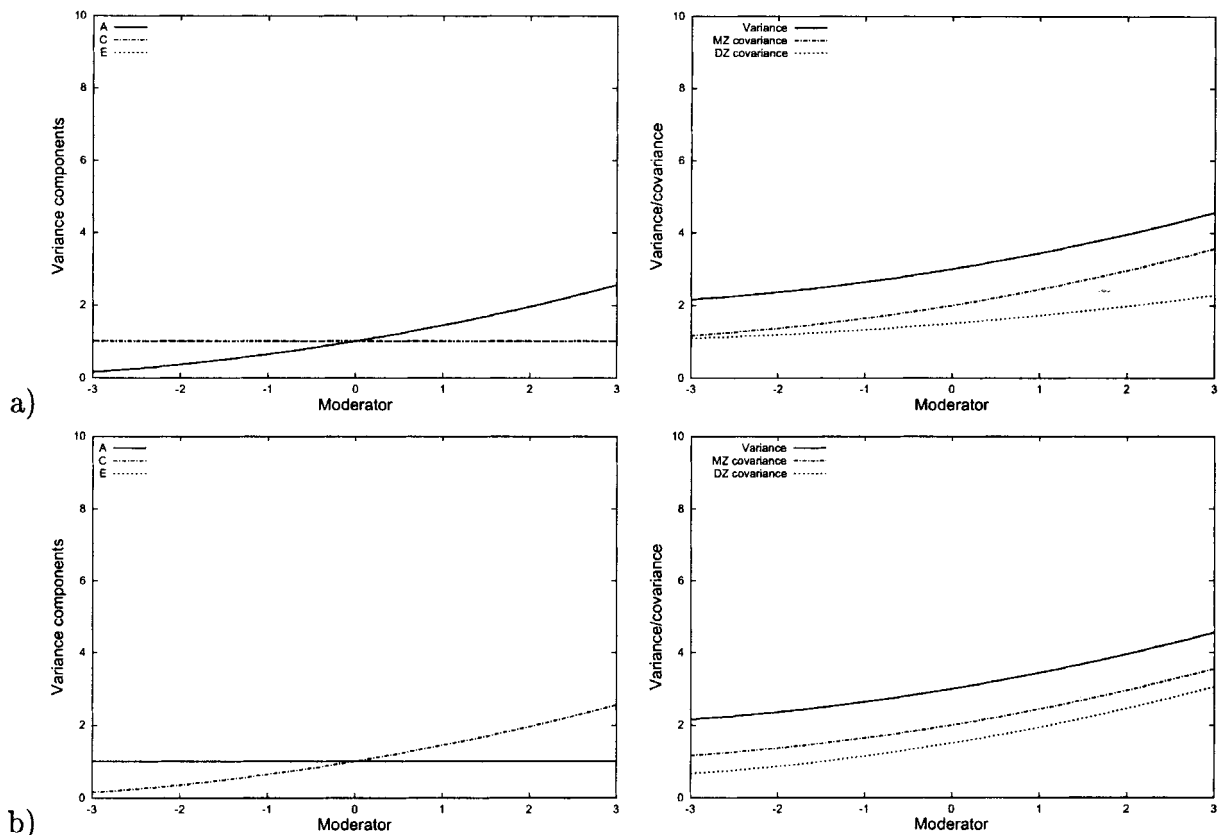


Figure 5

Relationship between variance components and expected variance, twin covariance. In both cases $a = c = e = 1$. In panel (a), $\beta_x = 0.2$ whilst $\beta_y = \beta_z = 0$. In panel (b), $\beta_y = 0.2$ whilst $\beta_x = \beta_z = 0$. Despite the marked difference in etiology (lefthand graphs), the expected variances and covariances are remarkably similar (righthand graphs).

The third model also includes a main effect, $\beta_M = 0.2$. The final model has two opposing interactive effects, $\beta_X = 0.2$ and $\beta_Y = -0.2$.

In these simulations there is no genetic component to the moderator (i.e., the MZ correlation always equals the DZ correlation for the moderator). However, genetic effects on the moderator should not have any great impact, unless the genetic effects are also shared with the trait (this scenario is investigated further below, $G \times E$ in the presence of r_{GE}).

The likelihood ratio test (LRT) column of Table 3 represents a 4 degree of freedom test between ACE-XYZ-M and ACE models. As can be seen, the average parameter estimates all fall very close to the simulated values. The LRT is as expected under the null (around 4 for a 4 degree of freedom test). For the combined interactive effects, the LRT values are very high, indicating that a joint test of all moderating effects is very powerful in this case. It appears that moderator variables that are uncorrelated between twins offer slightly more resolving power in $G \times E$ analysis. Finally, note that when the interactive effects are in opposition and almost canceling each other out, as in the fourth model, the power to detect them jointly is much smaller (although the power to detect them individually would presumably be greater than usual).

Multiple Moderator Variables

Within this framework, it is possible to incorporate more than one moderator variable along with any interactions between them. As a concrete example, age might moderate genetic effects but only in males. In this case, considering only $A \times M$, two moderator variables, age (M_{Age} , continuous) and sex (M_{Sex} , binary) would have their own interaction coefficients, β_{Age} and β_{Sex} ; additionally, an interaction parameter $\beta_{Age \times Sex}$ captures any difference in age-moderated genetic effects between sexes. The additive genetic variance component is now

$$(a + \beta_{Age} M_{Age} + \beta_{Sex} M_{Sex} + \beta_{Age \times Sex} M_{Age} M_{Sex})^2$$

This kind of extension should probably be limited to cases where prior knowledge or analyses have at least suggested a

moderating effect for both variables. The results will be easiest to visualize when one of the moderators is binary (i.e., two plots of variance components as a function of the continuous moderator one for each level of the binary moderator). A significant interaction parameter for the two moderators means that the slope for a particular variance component will differ between plots.

As an example, a single dataset involving two-variable moderation was simulated. For 500 MZ and 500 DZ twin pairs, continuous (C) and binary (B) moderators were simulated for each individual (only moderation of additive genetic effects is considered in this example). For the "0/1" binary parameter, additive genetic effects were moderated with $\beta_B = 0.5$. The continuous moderator variable had a coefficient of 0.2, but only for individuals scoring "1" on the binary moderator (i.e., $\beta_C = 0$ and $\beta_{B \times C} = 0.2$). Residual components were set at $a = c = e = 1$. The estimates were as follows: $\beta_B = 0.558$, $\beta_C = 0.046$, $\beta_{B \times C} = 0.204$, with $-2LL = 7603.025$. Fixing the interaction moderating parameter, $\beta_{B \times C}$ to 0, the minus twice log-likelihood rose to 7610.854 — a significant difference for 1 degree of freedom.

Summary

This section described a basic $G \times E$ model which allows for one or more continuous or binary moderating variables to have main effects on a trait as well as interacting with any or all of the genetic and environmental latent variables. In general, simulation results suggest that the model will perform well although there may be issues of specificity, e.g., distinguishing between $A \times M$ and $C \times M$.

Nonlinear $G \times E$ with Continuous Moderator Variables

So far we have assumed that, at the level of effect, all $G \times E$ interactions are linear. In order to more accurately characterize a conceptually interesting class of $G \times E$ models, however, it is necessary to extend the basic model to allow for certain nonlinear interactions. Imagine that at least some level of exposure to an environmental risk is *necessary* to develop disease, whilst high levels of exposure are *sufficient* to cause disease, and that otherwise disease status is

Table 3

Linear $G \times E$ Interaction in Twins. The LRT Represents the Difference in Model Fit Between the ACE and ACE-XYZ-M Models (i.e., Fixing β_X , β_Y , β_Z and β_M to 0) Which Is Distributed as a χ^2 on 4 Degrees of Freedom

Simulated					Estimated				
β_X	β_Y	β_Z	β_M	r	β_X	β_Y	β_Z	β_M	LRT
—	—	—	—	0	0.01	0.00	-0.01	0.00	3.87
				0.5	0.00	0.00	0.00	0.00	3.94
				1	-0.01	0.01	0.00	0.00	4.20
0.2	0.2	0.2	—	0	0.20	0.20	0.21	0.00	211.28
				0.5	0.19	0.21	0.20	0.00	185.22
				1	0.20	0.21	0.20	0.00	168.43
0.2	0.2	0.2	0.2	0	0.20	0.20	0.21	0.20	249.62
				0.5	0.21	0.19	0.21	0.20	217.14
				1	0.22	0.20	0.20	0.20	185.95
0.2	-0.2	—	—	0	0.18	-0.18	0.00	0.00	11.14
				0.5	0.16	-0.17	0.01	-0.01	9.14
				1	0.19	-0.18	0.00	0.00	9.43

influenced by genes. At both extreme low levels of the moderator (nobody has the disease) and extreme high levels (everybody has the disease) there is no genetic variation; at intermediate values of the moderator there is variation due to genes. Although this scenario has been expressed in terms of a binary disease for clarity, similar principles would apply to quantitative traits, as Figure 6 illustrates. The leftmost figure illustrates variation (due to genes) only occurring at intermediate values of the moderator, due to the above kind of process. The second figure illustrates the residual variation after the main effect of the moderator has been partialled out in the means model (β_M). This shows the characteristic pattern of genetic variation being attenuated at the extremes of the moderator. The rightmost figure returns to the biometrical model for the hypothetical QTL – there is a nonlinear interaction at the level of effect. This kind of nonlinear interaction can be well approximated by a quadratic term being added to the equation describing the additive genetic effect. In other words, the additive genetic path now becomes $(a + \beta_X M + \beta_{X^2} M^2)$. The full model is now $ACE\text{-}XYZ\text{-}X^2 Y^2 Z^2\text{-}M$. The variance is

$$\begin{aligned} \text{Var}(T) = & (a + \beta_X M + \beta_{X^2} M^2)^2 \\ & + (c + \beta_Y M + \beta_{Y^2} M^2)^2 + (e + \beta_Z M + \beta_{Z^2} M^2)^2 \end{aligned}$$

This kind of model might be of interest in a wide variety of circumstances. For example, it is possible that exposure to combat and post-traumatic stress disorder would follow a similar pattern. Eaves et al. (1977) noted that there are many situations in which we might find significant nonlinear trends: “Society may react in a uniform way to extreme deviations on either side of the population mean. This would produce a pattern ... which shows greater environmental variation in the middle of the scale than either end,” which would represent nonlinear $E \times E$ in our terminology. The authors continue: “In practice, this kind of interaction is common in psychometric data because of floor and ceiling effects.” Alternatively, we may expect genes to operate maximally only in their “average expected environment” (Scarr, 1992) such that genetic variation is attenuated at both environmental extremes.

An Example

Twin data were simulated under an AE model such that $r_{MZ} = 0.8$ and $r_{DZ} = 0.4$. On top of these residual components, a

moderator variable was simulated with two properties, (1) a main effect on the trait, such that the phenotypic trait-moderator correlation was $r = 0.25$ and (2) a moderating effect such that all genetic effects were exaggerated at intermediate values of the moderator but attenuated at extreme values of the moderator.

This moderating effect is the converse of that in the example linear $G \times E$ simulation, where intermediate values were attenuated and extreme values exaggerated. Note that although the two scenarios display superficial similarity, a nonlinear term is required in this case. For the linear $G \times E$ models, it is clear that (1) the variance components cannot be negative and (2) they are a function of the moderator up to a second-order term. Consequentially, any curve will always be “U-shaped” where the stationary point is also the global minimum. This is not a constraint as such — it follows naturally from assuming a linear $G \times E$ interaction, but it also illustrates the need for the nonlinear models.

Figure 7 displays the results for this simulation (leftmost figure), as well as illustrating a further problem with visualizing the variance components (other two figures). The best-fitting model is the $AE\text{-}X^2\text{-}M$ which precisely recovers the simulated architecture. However, the leftmost figure shows the expected variance components plotted as a function of the moderator, revealing a pattern that only partially corresponds to our intuitions regarding the simulated properties (i.e., exaggerated at intermediate levels, attenuated at extreme levels). That is, the curve describing the A component seems to suggest that, moving in either direction away from the moderator mean, genetic influences decrease and then sharply increase at even more extreme values. Although this is merely an artifact of over-extrapolation and over-fitting, it raises the question of the precise range of the moderator used to visualize the results. In this case, the x-axis corresponds exactly to the observed range of the moderator, which seems a sensible choice (i.e., rather than artificially truncating the moderator distribution). How would one interpret these results in the absence of prior knowledge (i.e., on non-simulated data)?

Consideration of the distribution of the moderator (shown in the middle figure) is useful. As hardly any cases score at those extreme values of the moderator, there is little or no power to estimate the true location of the curve at these points. Although the estimates have over-fit to the

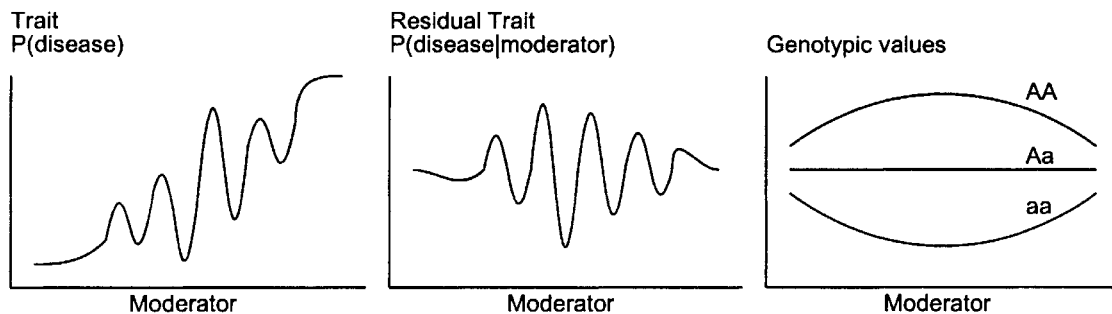


Figure 6

Nonlinear interaction and the biometric model. Please see the text for a full description.

data somewhat, there would be very little change to the sample log-likelihood if the curve were drastically altered at these extreme values. A method of visualization is proposed, and illustrated in the rightmost figure, such that the intensity of the line is directly proportional to the frequency of the moderator within x bins across the distribution. In this way, the visibility of the line is related to the contribution to the sample log-likelihood for that portion of the distribution: if the line is invisible it is because there is little or no power to place it. Although this is not an exact method, it should help to interpret results more clearly, as illustrated in this example, where the curve now approximates the simulated architecture. (For ease of presentation, standard plots will be utilized for the rest of this paper, however.)

Further Simulations

A different approach was adopted for this set of simulations, with only 3 replicate datasets simulated under 4 different models in order to allow a closer inspection of the solutions. All results are presented graphically in Figure 8. The four rows of plots represent the four models. The leftmost column of plots represent the true parameter values used to simulate the data. In all cases 500 MZ and 500 DZ twins were simulated, with a moderator variable correlated 0.5 between twins. The middle column represents the variance components estimated in three simulated samples, superimposed upon each other, from the $ACE-XYZ-X^2Y^2Z^2$ model. The third column represents the superimposed $ACE-X-X^2$ model estimates. In all 12 simulations, the $ACE-X-X^2$ model was selected as the best-fitting model, which corresponds to the simulated values chosen.

The four models were chosen to represent different scenarios that involve nonlinear interaction terms. In all cases, the interactions involved only the A additive genetic component; also $c = 1$ and $e = 1.5$. The first model (row a) $a = 2$, $\beta_x = 0$ and $\beta_{x^2} = -0.2$ is similar to the last example. The second model (row b) is similar to the first linear example although involving a quadratic interaction term: $a = 1$, $\beta_x = 0$ and $\beta_{x^2} = 0.2$. The third model (row c) represents a kind of “threshold effect” where only above a certain critical value on the moderator does the genetic variance shoot up, in this case around 1.5 standard deviations above the

mean): $a = 0$, $\beta_x = -0.8$ and $\beta_{x^2} = -0.2$. The final model (row d) represents a similar scenario: a threshold type effect for extreme low scorers, but also a linear increasing trend above the mean that plateaus out above 2 SD ($a = 1$, $\beta_x = 1$ and $\beta_{x^2} = -0.2$).

Although not shown on Figure 8, other models were fit to the data: in all cases, there was no significant reduction in model fit from dropping the non-genetic interaction terms (i.e., $ACE-XYZ-X^2Y^2Z^2$ versus $ACE-X-X^2$). In contrast, comparing $ACE-XYZ-X^2Y^2Z^2$ and $ACE-YZ-Y^2Z^2$ models, in 11 out of 12 cases the genetic interaction terms could not be dropped. None of the genetic interaction terms could be dropped from the $ACE-X-X^2$ model, however. Using AIC criterion, the $ACE-X-X^2$ model was the best fit model in all cases, of the models compared: $ACE-XYZ-X^2Y^2Z^2$, $ACE-YZ-Y^2Z^2$, $ACE-X-X^2$, $ACE-X$ and ACE .

As can be seen from Figure 8, the parameters are recovered quite well, allowing for sample-to-sample variation. Under the full model the plots are a little messy, but under the $ACE-X-X^2$ model (the best-fitting model in all cases), the simulated structure is recovered very well indeed. The simulated effects are quite large, although nonlinear effects have been found in real, modestly-sized datasets also (unpublished work). Again, it is important to remember whilst looking at the plots, that most of the sample will fall with 1–2 SD above and below the mean, so the models are not quite as extreme as they first seem.

Although generally robust with this sample size, these problems are sensitive to starting values and prone to local minima, as well as being computationally expensive. Care must be taken when fitting these models.

Summary

In order to characterize a large class of potential $G \times E$, in which an effect is attenuated at both high and low extremes of a moderator, a quadratic term was entered in the model. Simulation results suggest that it is possible to discriminate between the nonlinear models and to estimate the interaction coefficients (of which there are up to 6) quite well using only moderately large sample sizes.

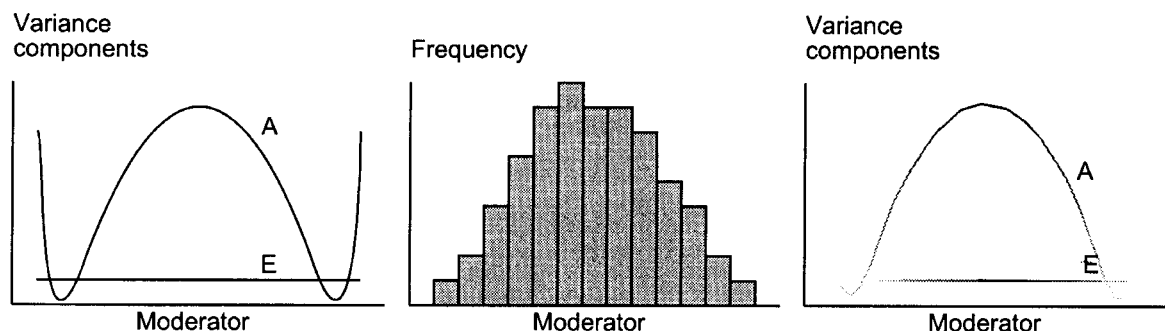


Figure 7

Visualization of variance components for the nonlinear $G \times E$ example.

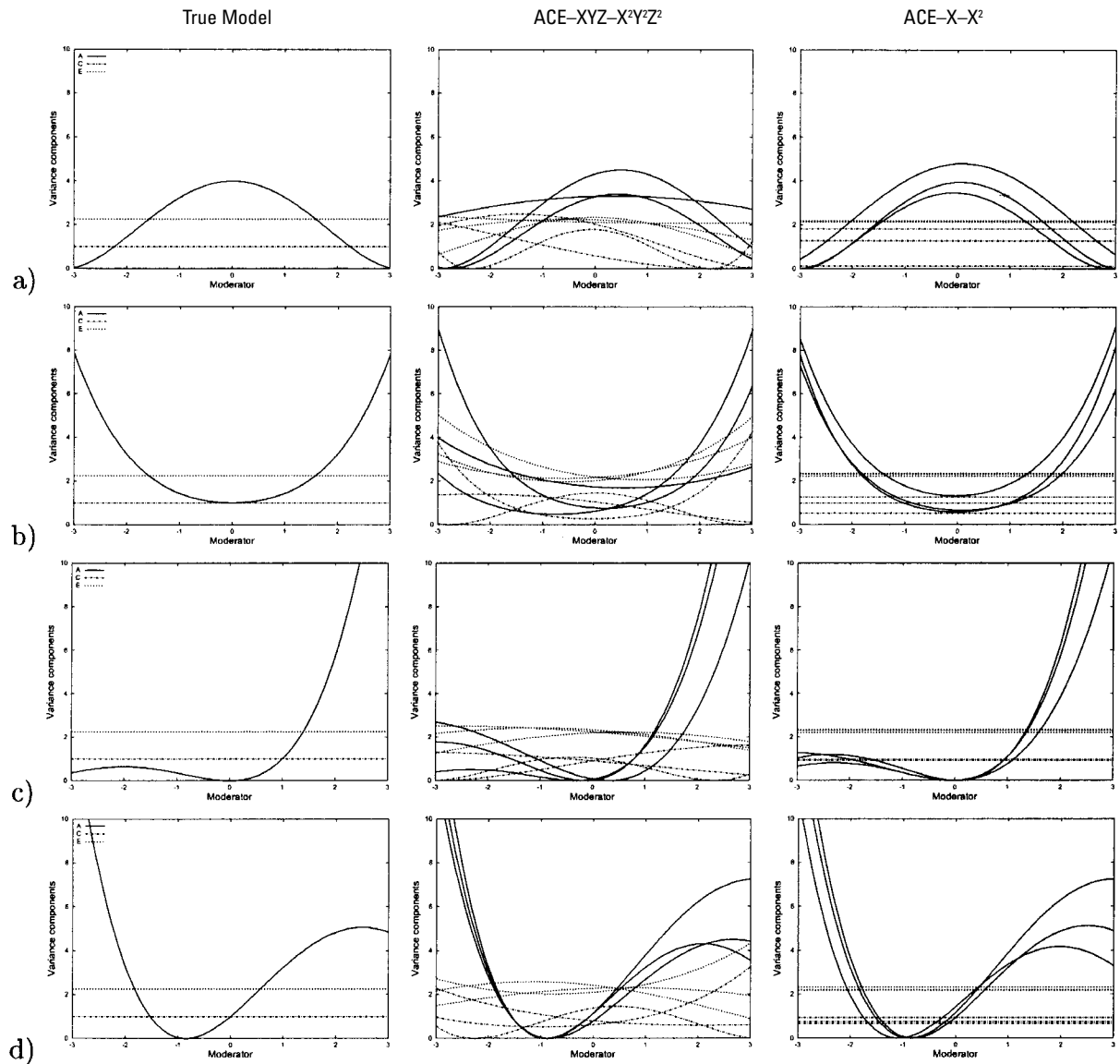


Figure 8

Nonlinear models: simulation under four genetic models (a–d). The left column of figures represent the true model; the middle column represents the variance components estimated in three samples of 500 MZ and 500 DZ twins simulated under each model, superimposed on the graph, from the $ACE-XYZ-X^2Y^2Z^2$ model. The third column represents the $ACE-X-X^2$ model estimates from three replicates superimposed — this was the best fitting model in each case.

Gene–Environment Correlation

Potential moderators will typically be correlated with the trait — in the absence of strong *a priori* reasons, it is likely to be the phenotypic association that flagged up the variable as a potential moderator in the first case. Although this correlation may be due to trait-mediating effects of the moderator, it may alternatively be due to other shared causes, which includes the possibility of shared genes. It is well known that many “environmental” variables demonstrate substantial heritable components (Plomin et al., 2001). Many environmental variables may in fact be *correlated* with the genetic effects on the trait (r_{GE}) rather than *modifying* the genetic effects on the trait ($G \times E$).

As mentioned earlier, r_{GE} can appear as $G \times E$ in typical analyses of $G \times E$. However, the current approach can be easily extended to model $G \times E$ in the presence of r_{GE} .

Entering the moderator in the means model to allow for a main effect will effectively remove from the covariance model any genetic effects that are shared between trait and moderator. That is, r_{GE} will appear as a main effect, β_M , due to the moderator acting as a proxy measure for the additive genetic effects on the trait. Any interactions detected will not be due to r_{GE} , but rather will be interactions between the moderator and variance components specific to the trait. In this way, evidence for $G \times E$ will never reflect a “false-positive” claim for interaction. However, this approach will also fail to detect $G \times E$ interaction where the

moderated genetic component is common between trait and moderator (i.e., $G \times E$ in the presence of r_{GE}).

Table 4 illustrates the application of the basic $G \times E$ presented in this paper so far in the presence of a genetic correlation between moderator and trait. Again, 500 MZ and 500 DZ twins were simulated 50 times under each condition; $a = c = e = 1$ and the twin correlation for the moderator is 0.5. Note that the estimate of β_M is inflated due to the shared genetic effects. More importantly, the tests of $G \times E$ not allowing for any main effect (the last column) show inflated test statistics when there is no interaction but there is a gene–environment correlation. That is, the third and sixth rows have average values of 4.01 and 11.06 for this test, both of which are greater than the critical value for this 1 degree of freedom test. Note however that the test of an interaction that allows for a main effect (second to last column) does not show such an effect.

$G \times E$ in the Presence of r_{GE}

In the previous simulations, the interactive effect was simulated for the genetic effects specific to the trait. If in fact the interaction was with only the genetic effect shared with the moderator, the above model would have failed to detect it. As mentioned, this is because these effects have already been partialled out in the means model. For a continuous moderator, the current model can be re-formulated as a bivariate model of both trait and moderator in order to detect these effects of $G \times E$ in the presence of r_{GE} , however. Figure 9 shows the partial path diagram for one twin to illustrate this approach. Here the moderator features twice in the model — as a dependent variable to be modelled as well as a moderator variable to define the paths to the trait. The main effect in the means model has been replaced with a path indicating shared genetic effects. The trait is now influenced by two sources of genetic influence: that which is shared with the moderator, and that which is not (common and unique paths a_C and a_U). Each path can interact with the moderator, represented by the coefficients β_{X_C} and β_{X_U} . The C and E components (not shown on the

path diagram) follow the standard bivariate Cholesky parameterization, but without moderation. From these parameters, r_{GE} can be calculated as

$$a_M a_C / (a_M \sqrt{a_C^2 + a_U^2})$$

if there is no $G \times E$ but otherwise gives r_{GE} at the mean value of the moderator (assuming a zero-centered moderator). In the presence of $G \times E$, r_{GE} must be calculated conditional on M as it will also vary as a function of the moderator

$$r_{GE} | M = \frac{a_M(a_C + \beta_{X_C}M)}{a_M \sqrt{(a_C + \beta_{X_C}M)^2 + (a_U + \beta_{X_U}M)^2}}$$

whilst calculating the average r_{GE} for the sample involves integrating over the distribution of the moderator.

The following set of simulations illustrates this model's ability to distinguish between the two types of interaction, whether or not there is a genetic correlation. Results are shown in Table 5. The conditions (each consisting of 25 replicate samples of 500 MZ and 500 DZ twins) varied the genetic correlation between moderator and trait and the presence or absence of moderating effects on the common and unique genetic paths, as described above. The genetic correlations correspond to the correlation of unmoderated effects only. That is, $r_{GE} = 0$ corresponds to $a_M = 1$, $a_C = 0$ and $a_U = 1$; $r_{GE} = 0.5$ corresponds to $a_M = 1$, $a_C = \sqrt{0.25}$ and $a_U = \sqrt{0.75}$; $r_{GE} = 1$ corresponds to $a_M = 1$, $a_C = 1$ and $a_U = 0$. Shared and nonshared environmental effects were simulated for each component to have a variance of 1 and be uncorrelated between trait and moderator (i.e., $c_M = e_M = 1$, $c_C = e_C = 0$ and $c_U = e_U = 1$). The 15 conditions are arranged in five blocks: (1) no moderation (2) moderation of common path (3) moderation of unique path (4) moderation of common and unique path, similar effects (5) moderation of common and unique paths, opposing effects. Four models were analysed: $ACE-X_C X_U$, $ACE-X_C$, $ACE-X_U$ and ACE . Three likelihood ratio test statistics were constructed (final three columns in Table 5): in order

Table 4

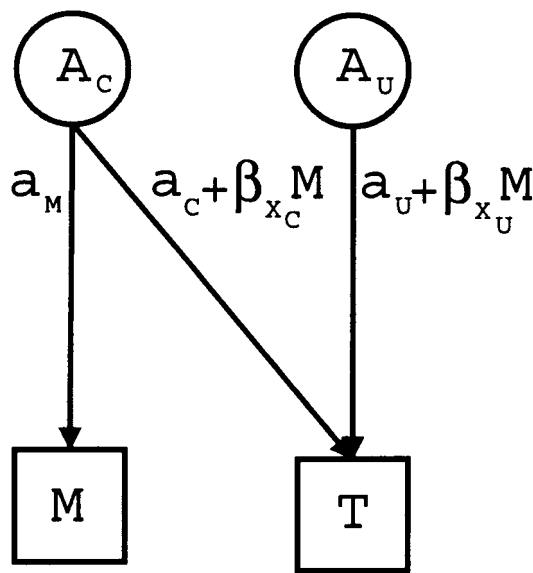
Performance of the Basic $G \times E$ model in the presence of r_{GE} (i.e., the Moderator M has Shared Genetic Influence with the Trait, Measured by the Genetic Correlation, r_{GE}). The Table Presents Parameter Estimates for the $ACE-X-M$ Model and Three Likelihood Ratio Tests: of a Main Effect in the Presence of an Interaction; of Interaction in the Presence of a Main Effect; of Interaction not Allowing for any Main Effect

Simulated	Estimated								Likelihood ratio tests		
	r_{GE}	β_X	β_M	a	c	e	β_{X_C}	β_{X_U}	$ACE-X-M$	$ACE-X-M$	$ACE-X$
									$ACE-X$	$ACE-M$	ACE
0	—	—	—	0.99	0.99	0.99	-0.01	0.00	0.83	1.03	1.03
0.5	—	—	—	0.95	0.98	1.01	0.00	0.14	37.86	1.04	1.07
1	—	—	—	0.66	1.05	1.04	0.00	0.28	155.51	1.76	4.01
0	—	0.2	—	0.99	1.01	1.00	0.00	0.20	79.66	0.99	0.95
0.5	—	0.2	—	0.93	1.02	1.00	0.00	0.34	215.67	1.06	2.07
1	—	0.2	—	0.65	1.06	1.04	0.00	0.48	406.74	1.36	11.06
0	0.2	—	—	0.99	1.00	1.00	0.21	0.00	0.97	49.60	49.73
0.5	0.2	—	—	0.94	1.00	1.00	0.20	0.12	25.72	46.54	55.07
1	0.2	—	—	0.68	1.05	1.02	0.25	0.25	100.40	46.19	88.90
0	0.2	0.2	—	1.01	0.98	0.99	0.21	0.20	74.03	52.61	56.26
0.5	0.2	0.2	—	0.94	1.00	1.00	0.20	0.31	167.17	44.99	77.03
1	0.2	0.2	—	0.70	1.04	1.02	0.25	0.45	308.27	43.88	146.27

Table 5

Performance of the Extended $G \times E$ model in the presence of r_{GE} . The Table presents Parameter Estimates for the $ACE-X_cX_u$ Model and Three Likelihood Ratio Tests: Of Moderation for Both Common and Unique Genetic Effects; of Moderation for Common Effects Only; of Moderation for Unique Effects Only

r_{GE}	Simulated			Estimated					Likelihood ratio tests		
	β_{X_c}	β_{X_u}		a_M	a_c	a_u	β_{X_c}	β_{X_u}	$ACE-X_cX_u$ ACE	$ACE-X_cX_u$ $ACE-X_u$	$ACE-X_cX_u$ $ACE-X_c$
0	—	—		0.99	-0.01	0.97	0.01	0.01	2.83	1.77	1.03
0.5	—	—		0.98	0.48	0.78	0.01	0.00	1.72	0.92	0.83
1	—	—		1.04	0.98	0.26	0.00	0.03	1.92	0.87	1.04
0	0.2	—		1.01	-0.01	0.94	0.20	0.01	59.74	57.72	1.27
0.5	0.2	—		1.04	0.49	0.84	0.19	0.00	63.82	53.67	1.00
1	0.2	—		1.03	0.94	0.29	0.20	0.03	86.77	53.99	0.70
0	—	0.2		1.02	0.00	0.99	0.00	0.19	46.28	0.65	26.81
0.5	—	0.2		1.03	0.45	0.88	0.01	0.20	41.68	1.04	22.80
1	—	0.2		1.00	0.94	0.26	0.00	0.18	8.43	1.14	6.85
0	0.2	0.2		1.05	-0.07	0.97	0.19	0.22	99.28	44.28	19.64
0.5	0.2	0.2		1.02	0.48	0.88	0.19	0.20	123.36	45.12	8.82
1	0.2	0.2		1.00	0.93	0.20	0.19	0.17	86.29	40.42	5.88
0	0.2	-0.2		1.04	-0.01	0.99	0.18	-0.20	94.53	40.40	15.33
0.5	0.2	-0.2		1.02	0.44	0.88	0.21	-0.19	80.46	49.79	15.49
1	0.2	-0.2		0.99	1.00	0.18	0.20	-0.20	86.30	47.22	4.48

**Figure 9**

Extended $G \times E$ model to allow for gene–environment correlation.

of the columns (1) moderation for both common and unique genetic effects (2) moderation for common effects only (3) moderation for unique effects only. The parameter estimates under the full $ACE-X_cX_u$ model are also shown in the Table.

Under the null (first three rows), the models perform as expected: the β coefficients are all near zero, and the tests of moderation show average χ^2 values close to their expected values under no moderation. The average unmoderated genetic parameter values are close to their simulated values, with the exception of a_u when $r_{GE} = 1$, which is simulated at

0, but has an average estimated value of 0.26. This artifact, which also exists in the other $r_{GE} = 1$ conditions, is explained further below.

The second three conditions simulate a moderating effect of shared genetic influence between the trait and moderator. This effect is recovered well and detected, no matter what the background genetic correlation between trait and moderator. The specific tests of β_{X_c} (the second column of likelihood ratio tests in the Table) show highly significant average values, whereas the β_{X_u} parameters average near zero. Power seems to increase as r_{GE} increases. The third three conditions simulate a moderating effect of genetic influence specific to the trait. Again, the parameters are recovered well, although power to detect β_{X_u} drops off as r_{GE} increases. The final six rows of Table 5 show that the model works when both β_{X_c} and β_{X_u} are nonzero.

When $r_{GE} = 1$ some subtle properties of the model emerge — they are worth considering in further detail. The contribution to the variance of total unique genetic effects is $(a_u + \beta_{X_u}M)^2 = a_u^2 + 2a_u\beta_{X_u}M + \beta_{X_u}^2M^2$. As mentioned above, the power to detect β_{X_u} decreases with increasing r_{GE} , because when $r_{GE} = 1$ then $a_u \rightarrow 0$ and so $2a_u\beta_{X_u}M \rightarrow 0$ which reduces the impact of β_{X_u} on the variance by canceling this cross-product term. A similar logic applies to the relationship between a_c and r_{GE} .

Additionally, as a_u and therefore $2a_u\beta_{X_u}M$ approach 0, then β_{X_u} only makes squared contributions to the variance. Therefore, when $r_{GE} = 1$, the estimate of a_u is likely to be near zero, which reduces the power to identify the *sign* of β_{X_u} although the *absolute value* can still be identified. Although the contribution to the variance will be the same (and so this is not an issue for the analysis of real data), taking the average of the unsquared parameter in repeated simulation would lead to an apparent bias in parameter estimate for β_{X_u} when $r_{GE} = 1$. The average values for β_{X_u} (not shown in the Table) were in fact -0.03, 0.02 and

–0.12, in the 9th, 12th and 15th rows, respectively: making the signs all positive (or all negative in row 15) produces the unbiased average parameter estimates as shown in the Table (0.18, 0.17 and –0.20).

As noted above, there is also an apparent bias in the estimates of a_U when $r_{GE} = 1$. This parameter has a large standard error, and it can be fixed to 0 when $r_{GE} = 1$ with no significant reduction in fit, on average. Again, under certain conditions the sign of a_U is not identified; however, optimization favors the positive values, probably due to a positive starting value being specified. This apparent bias is therefore not important in real analysis, it is only a consequence of taking averages.

The current model of $G \times E$ in the presence of r_{GE} could be extended in a number of ways. For example, a third variable that is a potential index of genetic sensitivity to an environmental factor can be incorporated, to produce models similar to recent Markov Chain Monte Carlo methods which handle $G \times E$ in the presence of r_{GE} (Eaves & Erkanli, 2002).

Summary

In the basic model, any genetic effects that are shared between the trait and the moderator will be modelled as main effects of the moderator. An extension to the basic model explicitly models shared genetic effects, as well as any interactions between these effects and the moderator, allowing for the analysis of $G \times E$ in the presence of r_{GE} .

Qualitative $G \times E$ with Continuous Moderator Variables

All the previous models of $G \times E$ have implicitly addressed *scalar*, or *quantitative*, moderation, meaning that the *magnitude* of polygenic effect has varied as a function of the moderator. However, it is also possible that *different* polygenic effects operate at different points along the moderating continuum. The same distinction is found in “sex-limitation” models, where males and females may have

different magnitudes of genetic (or environmental) effects but may also differ in which genes operate in males and females. Evidence for the “different genes” hypothesis is reduced covariance between twins discordant for the moderating variable (i.e., sex).

To allow for qualitative $G \times E$ with continuous moderator variables, we adopt the most simple biological model: that there are two independent sets of polygenes, $A1$ and $A2$, which show different profiles of scalar interaction with the moderator. Figure 10 illustrates this concept. The left panel depicts a standard moderated variance component, which is consistent with at least some genetic effects being amplified at higher values of the moderator. This curve could also have come about as a result of different genes operating at higher levels of the moderator, however, as shown in the right panel. Here we see that the $A1$ set of polygenes is not moderated, whereas the $A2$ set only have an effect at high levels of the moderator. In this way, individuals high on the moderator have a different profile of genes operating compared to individuals low on the moderator (not just greater or lesser effects of the same genes).

It is worth drawing a distinction between qualitative interaction and r_{GE} . Qualitative interaction implies that different loci have an effect depending on the value of the moderator. Gene–environment correlation implies that certain alleles of certain loci are present depending on the value of the moderator. In the latter case, an association between an individual’s genetic loading and the moderator ensues, which has to be explicitly modelled. This is not the case for qualitative interaction however.

Noting that a model with both sets of polygenes showing scalar interaction is not identified, the expected additive genetic variance for twin i is now $a_1^2 + (a_2 + \beta_{x_2} M_i)^2$; the additive genetic component of the MZ covariance is $a_1^2 + (a_2 + \beta_{x_2} M_1)(a_2 + \beta_{x_2} M_2)$; the additive genetic component of the DZ covariance is $a_1^2/2 + (a_2 + \beta_{x_2} M_1)(a_2 + \beta_{x_2} M_2)/2$. The formulation implies that the effective coefficients of genetic relatedness will be attenuated for twin

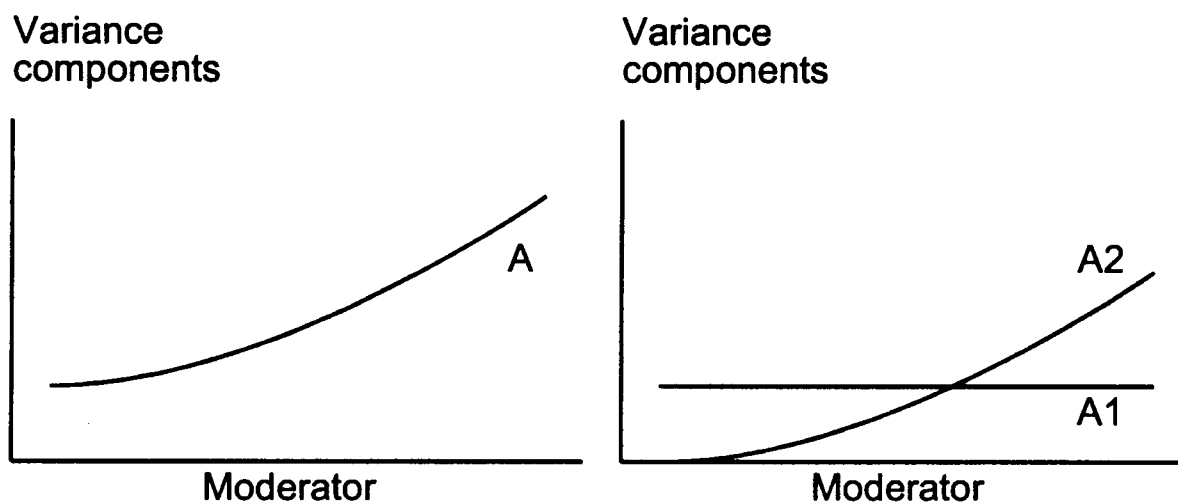


Figure 10

Schematic illustrating scalar (left figure) and qualitative (right figure) $G \times E$. See text for further explanation.

pairs discordant on the moderator under qualitative interaction. That is, the effective coefficients (normally 1 and 0.5 for MZ and DZ pairs respectively) are for MZ pairs

$$\alpha_{MZ} = \frac{a_1^2 + (a_2 + \beta_{X_2} M_1)(a_2 + \beta_{X_2} M_2)}{\sqrt{a_1^2 + (a_2 + \beta_{X_2} M_1)^2} \sqrt{a_1^2 + (a_2 + \beta_{X_2} M_2)^2}}$$

and for DZ twins

$$\alpha_{DZ} = \frac{a_1^2/2 + (a_2 + \beta_{X_2} M_1)(a_2 + \beta_{X_2} M_2)/2}{\sqrt{a_1^2 + (a_2 + \beta_{X_2} M_1)^2} \sqrt{a_1^2 + (a_2 + \beta_{X_2} M_2)^2}}$$

This model will be referred to as the $A_1A_2CE-X_2$ model (assuming that shared and nonshared environmental components are also included). It can be seen that when $M_1 = M_2$ then $\alpha_{MZ} = 1$ and $\alpha_{DZ} = 0.5$ for any values of a_1 , a_2 and β_{X_2} . Figure 11 shows the attenuation for MZ and DZ pairs as a function of M_1 and M_2 . Note that the exact shape of this surface will depend on a_1 , a_2 and β_{X_2} and can go negative under certain conditions. Clearly, this model is not applicable for obligatorily shared moderators.

Initial simulation results suggest poor power to discriminate between scalar and qualitative $G \times E$, however. Table

6 shows the results fitting the $ACE-X$ (scalar) and $A_1A_2CE-X_2$ (qualitative) models to six example datasets simulated under either scalar $G \times E$ (first three rows) or qualitative $G \times E$ (second three rows). The likelihood ratio test statistic (LRT column) is the χ^2 increase in fit from qualitative to scalar models. Simulating under a population value of $a_1 = 0$ implies scalar interaction (i.e., there is only one set of polygenes). Also, note that $c = e = 1$ and that 1000 pairs of each zygosity were simulated (twice the usual sample size). As can be seen, the qualitative model is correctly rejected in all three scalar cases (first three rows). However, there is at best only very weak evidence to support the qualitative model in the second three datasets, with only 1 of the three being significant at the 5% significance level. More extensive simulation work is required to properly evaluate the power of this test under a range of conditions.

Summary

An interaction may involve the same genes having different effects (scalar interaction) or different genes operating (qualitative interaction) at different levels of the moderator. A simple model of qualitative interaction for continuous

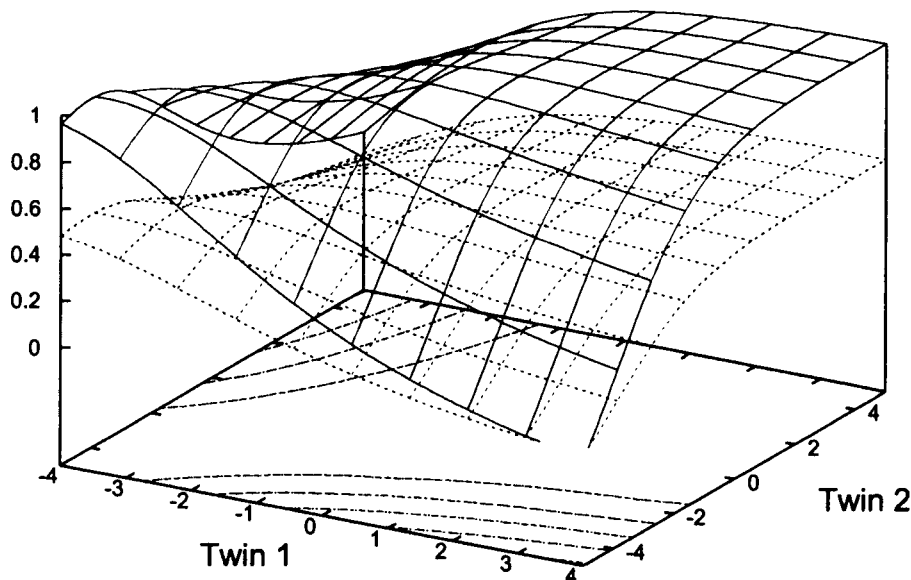


Figure 11

Plot of α_{MZ} (solid grid) and α_{DZ} (dotted grid) for $a_1 = a_2 = 1$ and $\beta_{X_2} = 0.3$. Along the diagonal $M_1 = M_2$, $\alpha_{MZ} = 1$ and $\alpha_{DZ} = 0.5$.

Table 6

Scalar and Qualitative $G \times E$: Results from Six Simulated Example Datasets

Simulated			$A_1A_2CE-X_2$			$ACE-X$		LRT	
a_1	a_2	β_{X_2}	a_1	a_2	β_{X_2}	a	β_X	χ^2	p
0	1	0.2	0.21	1.04	0.20	1.06	0.19	0.00	0.95
			0.00	1.10	0.17	1.10	0.17	0.00	1.00
			0.00	0.83	0.26	0.83	0.26	0.00	1.00
1	0.5	0.25	1.08	0.44	0.22	1.16	0.09	2.56	0.11
			1.14	0.45	0.19	1.22	0.07	1.51	0.22
			0.95	0.39	0.28	1.02	0.11	5.95	0.01

variables was presented, although power to discriminate between scalar and qualitative interaction empirically looks likely to be low.

Other Distributional Factors Influencing $G \times E$ Analysis

The effect of mismatching continuous and binary moderators

Although many moderator variables may act continuously, it is also entirely reasonable that some moderators act in a more discrete manner, even if they can be measured on a continuous scale. In this section we consider the impact of “misclassifying” a moderator variable: either falsely dichotomizing what is actually a continuous moderator or using a continuous measure when the moderating effect is really a threshold effect (e.g., only the top 10% individuals show an increased genetic effect).

Samples were simulated under two kinds of model: continuous or binary moderation. For the continuous case, $\beta_x = 0.2$; for the binary case, the continuous moderator was transformed to a binary scale, with individuals more than 1.28 standard deviations above the mean scored “1”, all others scored “0” (corresponding to a 9:1 ratio of “0”:“1”), with $\beta_x = 0.8$ (no direct comparison can be made between the magnitude of interaction in the continuous and binary cases in terms of β_x alone however).

Similarly, analysis adopted either a continuous or binary approach towards the moderator. The correctly classified scenarios are therefore when the data were generated using a continuous moderator and also analyzed using a continuous moderator; also, when the data were generated using a binary moderator and analyzed using the same binary moderator. The misclassified scenarios are when the data were generated using a continuous moderator, which was subsequently dichotomized for analysis; also, when the data were generated using a binary variable but the analysis used the underlying continuous “liability” instead.

In addition, some further analytic conditions were considered. A binary moderating effect could rightly be described as “nonlinear” in terms of the underlying continuous dimension — the nonlinear model was therefore included when analyzing a continuous moderator to see how well a quadratic function performs at approximating the “step function” of a threshold effect. Finally, although it

is common for experimenters to dichotomize continuous or semi-continuous variables in analysis (e.g., taking the top 5% of high scorers on the moderator) the chosen threshold may not correspond to the actual threshold (e.g., actually the top 10% show a moderated effect). Two further analysis conditions were included to represent this kind of misclassification of binary variables: the true threshold was 10% and the classification was either too harsh (5%) or too inclusive (25%). In all cases, $a = c = e = 1$ for 500 MZ and 500 DZ twin pairs; 50 replicates were generated under each condition.

Table 7 shows that, as expected, a continuous analysis model works much better when the moderation is truly continuous; likewise, a correctly specified binary moderator in analysis performs best when the moderation is truly binary. For continuously-moderated data, the average test statistic under the continuous analysis models is typically at least double the binary analysis models. As expected, allowing for a nonlinear continuous effect adds nothing. For binary-moderated data, the 10% binary model in analysis works best. However, the nonlinear model seems to offer a good approximation, capturing around three-quarters of the available information. Furthermore, when the binary analysis model is misspecified (i.e., the dichotomy is either too harsh or too inclusive), then performance is worse than the nonlinear model and equal to the linear continuous model.

From these results it seems to be a good strategy to adopt continuous moderators whenever available, allowing for nonlinear moderation to model any threshold effects. Figure 12 shows the average estimated variance components as a function of the moderator under different analysis models for the case of a binary moderating effect in the data, based on a liability with a sib correlation of 0.5.

Non-normal Trait Distributions

The current method relies on often relatively subtle differences in the variance, MZ covariance and DZ covariance across the range of the moderator variable to infer the presence of any interactive effect. Whilst it would be expected that deviations from multivariate normality may obscure these subtle effects, it is also possible that certain forms of measurement bias and error could lead to spurious evidence for $G \times E$.

Table 7

Continuous and Binary Moderators: Effects of Misspecifying Moderator Type. The LRT Represents the Difference in Model Fit Between the ACE and ACE-X (or ACE-X-X') Models. All LRT are Distributed as a χ^2 on 1 Degree of Freedom, Except for the Nonlinear Test Which is on 2 Degrees of Freedom

		Model											
		Continuous						Binary					
		Linear		Nonlinear				10%		5%		25%	
Data	r	β_x	LRT	β_x	β_{x^2}	LRT	β_x	LRT	β_x	LRT	β_x	LRT	
Continuous	0	0.20	22.05	0.20	0.00	23.36	0.36	8.17	0.41	6.88	0.33	12.58	
	0.5	0.20	18.05	0.20	0.01	19.16	0.36	7.72	0.40	5.67	0.32	10.82	
	1	0.22	17.04	0.23	-0.02	18.00	0.37	6.53	0.39	4.74	0.35	10.03	
Binary (10%)	0	0.17	16.12	0.15	0.10	27.40	0.79	41.30	0.71	17.95	0.37	16.12	
	0.5	0.24	18.57	0.18	0.10	28.80	0.84	42.84	0.73	18.62	0.39	16.36	
	1	0.20	12.63	0.15	0.10	19.88	0.81	28.77	0.73	13.30	0.35	10.71	

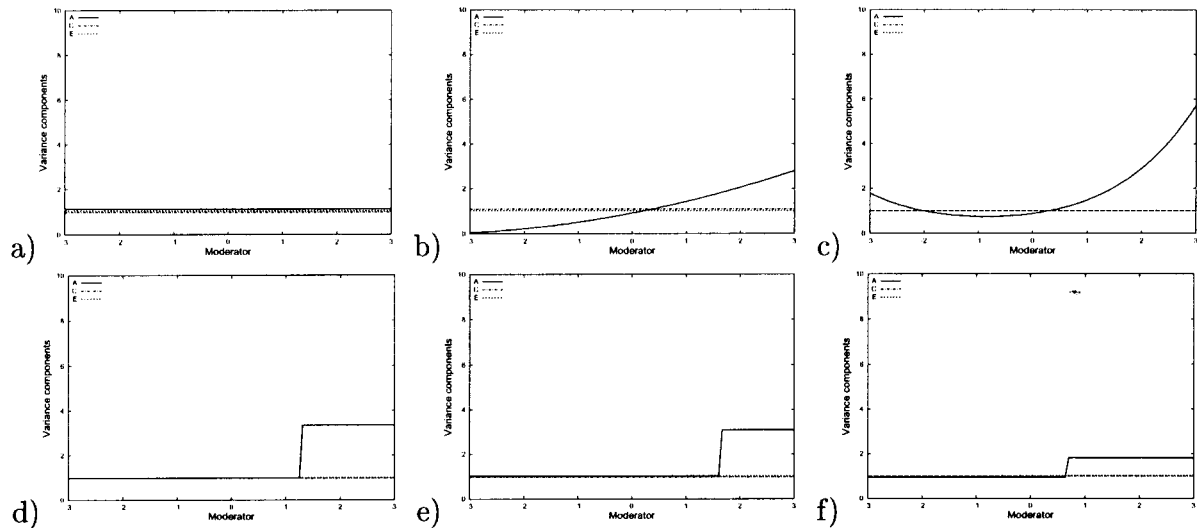
Many behavioral measurements have skewed, or J-shaped, distributions. For example, on a six-point symptom scale, the majority of individuals might score 0 or 1, whilst only a handful of individuals score above 4. If such a measure does in fact represent of normally-distributed liability, then the low end of the scale distribution is less informative than the high end. If a second variable correlates with the trait, then the second variable will also correlate with the “informativeness” of the first measure. This would be detected as an interactive effect. For example, the second variable would predict that twins with similar low scores on the liability are more likely to have identical scores on the measurement than twins with similar high scores on the liability. This would be an example of heteroscedasticity.

A set of simulations investigated this effect. In all cases, $a = c = e = 1$ for 500 MZ and 500 DZ twin pairs. A continuous covariate was simulated with a sibling correlation of 0.5. Three conditions were assessed: (1) no moderation and no main effect, $\beta_X = \beta_M = 0$ (2) a main effect only, $\beta_X = 0$, $\beta_M = 0.5$ and (3) a true moderating effect and a main effect, $\beta_X = 0.2$, $\beta_M = 0.5$. Twenty-five replicate datasets were

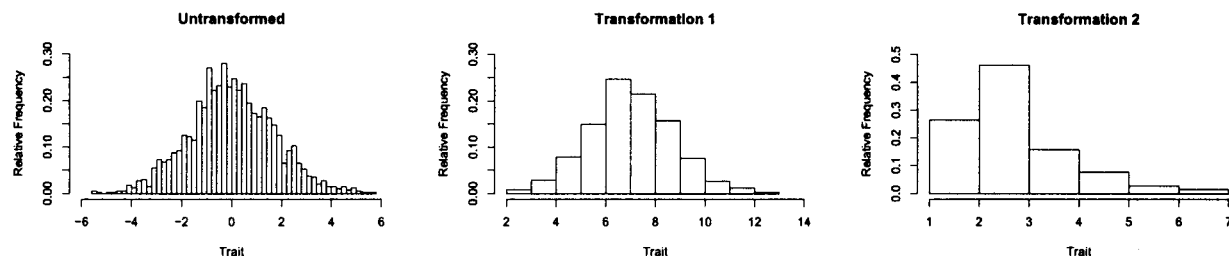
Table 8

Tests of Moderation under Skewed Trait Distributions. The LRT Column is the 3 Degree of Freedom Likelihood Ratio Test Statistic for $\beta_X = \beta_Y = \beta_Z = 0$

Simulated		Estimated				
β_M	β_x	β_M	β_x	β_y	β_z	LRT
Untransformed						
—	—	0.00	0.02	−0.02	−0.02	2.88
0.5	—	0.51	−0.02	0.01	0.01	3.67
0.5	0.2	0.50	0.16	0.03	0.01	19.16
Transformation 1						
—	—	0.00	0.03	−0.02	−0.02	2.74
0.5	—	0.50	−0.03	0.02	0.00	3.65
0.5	0.2	0.50	0.16	0.03	0.02	18.32
Transformation 2						
—	—	0.00	0.00	−0.01	0.00	3.60
0.5	—	0.28	0.06	0.09	0.08	60.50
0.5	0.2	0.30	0.17	0.10	0.09	126.69


Figure 12

Binary moderators and continuous approximations, for data simulated with a binary moderating effect (10% threshold): (a) ACE model (b) ACE – X model with continuous moderator (c) ACE – X – X² model with continuous moderator (d) ACE – X model with binary moderator (10%) (e) ACE – X model with binary moderator (5%) (f) ACE – X model with binary moderator (25%).


Figure 13

Example of data simulated and then transformed to investigate tests of moderation in skewed distributions. The first transformation bins the data-points into 15 bins; the second transformation is more severe and introduces a skew in the data.

simulated under each of the three models. In analysis, two models were fit to the data: *ACE-XYZ-M* and *ACE-M*, the difference in fit between which provides a 3 degree of freedom test of *any* moderating effect. Each replicate dataset was subjected to two transformation schemes, illustrated in Figure 13. Transformation 1 simply bins the continuous trait score into a less informative 15-point scale; transformation 2 bins the scores more severely and introduces a skew into the distribution.

Table 8 gives the results for this set of simulations. We would not expect the likelihood ratio test of *ACE-XYZ-M* against *ACE-M* to be significant for scenario (1) or (2), whereas it should be significant for (3) due to the simulated interaction. This holds for the untransformed data and under the first transformation (the critical value for a χ^2 statistic with 3 *df* at the $\beta_{\chi^2} = 0.05$ level is 7.815), but not under the second transformation scheme: the difference in fit is 60.5 which has a *p*-value of 4.6×10^{-13} for condition (2) where no interaction is actually simulated. In other words, the transformation scheme has induced evidence for some kind of moderating effect.

This effect may seem to be a cause of concern, given high prevalence of such measurement scales. Inspection of the moderating coefficients should reveal a predictable signature however, $\beta_x \approx \beta_y \approx \beta_z$ when $a \approx c \approx e$. Plotting the expected variance components will reveal only a gentle trend for all variance components to be attenuated similarly at low levels of the moderator. Whilst possible as a real model, researchers should be cautious in their conclusions, especially in the presence of heteroscedasticity. A scenario when β_x , β_y and β_z are all significant in a similar direction is also consistent with what might be called “phenotypic interaction” between the trait and the moderator, or “ $P \times E$ ”. In this case, the moderator doesn’t interact with any component of variance specifically; rather, it increases variation in the entire trait, at what can be thought of as a “later stage” in the trait’s etiology.

Summary

The distributions of the moderator (binary versus continuous) and of the trait (normal versus non-normal) were investigated in this section. It appears that, under a nonlinear model, using a continuously measured moderator works well even if the actual moderation operates as a binary threshold effect. It was also shown how certain types of skewed trait distributions might generate spurious evidence for interaction.

Discussion

As long as an individual’s genetic makeup is represented by a single, latent “*A*”, then the possibilities of gene–environment interaction will approach an unavoidable limit. In the future, the analysis of multiple measured genotypes interacting with multiple measured environmental factors will be necessary, in order to refine the broad brushstrokes we currently use to characterize the quantitative genetics of complex human traits. Nonetheless, twin analysis of gene–environment interaction using continuous moderator variables should still offer some interesting insights into the etiology of many complex traits, although several issues not

yet covered may emerge in the application of these models to real data.

The simulations presented in this paper generate data that is “cleaner” than we might expect in practice. Although this is typically the case with all simulation studies, the present models rely on relatively subtle phenomena and so the extent to which systematic and stochastic biases generate misleading results has not been fully addressed. Most simulations were conducted using a moderately large sample of 500 MZ twins and 500 DZ twins: the behavior of the models in smaller and larger samples is of interest, also.

In addition to its cleanliness, a simulated dataset comes with the knowledge of the true model, which inevitably guides analysis. In practice, for a specific dataset it might not be obvious how best to approach the various inter-related questions that can be asked: binary versus continuous moderation, linear versus nonlinear effects, interactions versus main effects versus correlations, scalar versus qualitative interactions, multiple moderators, etc. It might therefore be useful to develop a “protocol”, by which different models are sensibly and systematically evaluated and compared.

Although standard bivariate models explain the relationship between any two traits in terms of shared or direct causation, the kind of relationship involved in $G \times E$ might also be plausible. In other words, it is not necessary that the *E* component of $G \times E$ actually be “environmental” in any traditional sense of the word. What constitutes an environment from the gene’s point of view is quite different from an individual’s point of view. For example, the internal biochemical state of the body in which a gene finds itself can sensibly be called its environment. For appropriate traits, it might therefore be worth considering the above interaction models along with the standard bivariate ones. Consider a fictitious example involving anorexia and neurotic symptoms. Say being anorexic has various consequences including chronic low body weight. Low body weight may in turn lead to genes being switched on or off, some of which might operate to increase or decrease the chance of neurotic symptoms. Therefore, there will be an increase in the genetic variance of neuroticism, as a consequence of an anorexia-related state switching on genes. This scenario is distinct from having a set of genes that operate jointly on anorexia and neuroticism (i.e., a genetic correlation); it is distinct from direct causation between anorexia and neuroticism, in that although being anorexic leads to an increased risk of being neurotic, this is only expressed in genetically-predisposed individuals. As such, this dynamic fits within the same analytic framework as the $G \times E$ models considered so far: in this example, a $G_{\text{Neuroticism}} \times E_{\text{Anorexia} \rightarrow \text{Bodyweight}}$ interaction. Such an effect might be called a “Gene-for-trait 1 \times trait 2”, or “ $G \times T$ ”, interaction.

Software

Scripts to perform the above analyses using Mx (Neale, 1997) and a Windows program for plotting variance components can be found at <http://statgen.iop.kcl.ac.uk/gxe/>.

Acknowledgments

S.P. would like to acknowledge the helpful comments of Pak Sham, Karestan C. Koenen, Rick Viken, Brian D'Onofrio and Avshalom Caspi. This work was supported by Medical Research Council grant G9901258, and National Eye Institute grant EY-12562.

References

- Wichers, M., Purcell, S., Danckaerts, M., Derom, R., Vlietinck, R., & van Os, J. (2002). Obstetric complications and child pathology: Evidence for gene-birth weight interaction. *Psychological Medicine* (in press).
- Boomsma, D., de Geus, E., van Baal, G., & Koopmans, J. (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.
- Eaves, L., & Erkanli, A. (2002). Markov Chain Monte Carlo approaches to analysis of genetic and environmental components of human development change and GxE interaction. *Behavior Genetics*, (in press).
- Eaves, L., Last, K., Martin, N., & Jinks, J. (1977). A progressive approach to non-additivity and geno-type-environmental covariance in the analysis of human differences. *British Journal of Mathematical and Statistical Psychology*, 30, 1–42.
- Heath, A., Eaves, L., & Martin, N. (1998). Interaction of marital status and genetic risk for symptoms of depression. *Twin Research*, 1, 119–122.
- Heath, A., Todorov, A., Nelson, E., Madden, P., Bucholz, K., & Martin, N. (2002). Gene–environment interaction effects on behavioral variation and risk of complex disorders: The example of alcoholism and other psychiatric disorders. *Twin Research*, 5, 30–37.
- Jang, K., Vernon, P., Livesley, W., Stein, M., & Wolf, H. (2001). Intra- and extra-familial influences on alcohol and drug misuse: A twin study of gene–environment correlation. *Addiction*, 96, 1307–1318.
- Jinks, J., & Fulker, D. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of behavior. *Psychological Bulletin*, 73, 311–349.
- Kendler, K., & Eaves, L. (1986). Models for the joint effect of genotype and environment on liability to psychiatric illness. *American Journal of Psychiatry*, 143, 279–289.
- Martin, N., Eaves, L., & Heath, A. (1987). Prospects for detecting genotype x environment interactions in twins with breast cancer. *Acta Geneticae Medicae et Gemellologiae*, 36, 5–20.
- Mustanski, B., Viken, R., Kaprio, J., Pulkkinen, L., & Rose, R. (2002). Exploring the association between pubertal development and depression. *Paper presented at the 32nd Annual Meeting of the Behavior Genetics Association, Keystone, CO.*
- Neale, M. (1997). *Mx: Statistical modeling*. Box 980126 VCU, Richmond VA 23298.
- Neale, M., & Cardon, L. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Plomin, R., DeFries, J. C., McClearn, G., & McGuffin, P. (2001). *Behavioral genetics* (4th ed.). London: Worth.
- Purcell, S., & Sham, P. (2002). Variance components models for gene–environment interaction in quantitative trait locus analysis. *Twin Research*, 5, 572–576.
- Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child development*, 63, 1–19.
- Yaffe, K., Haan, M., Byers, A., Tangen, C., & Kuller, L. (2000). Estrogen use, APOE, and cognitive decline: Evidence of gene–environment interaction. *Neurology*, 54, 1949–1954.