Heterogeneity
Gene x Environment Interaction II

Boulder Workshop, June 2022

Hermine H. Maes
with credit to many workshop faculty, Sarah Medland, Brad Verhulst, Marleen de Moor, Conor Dolan

GxE Model & Theory

classic Mx (Neale) 199?- definition variables

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

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.

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.

bivariate twin distribution, v
relates with twin pair su
However, as well as suffering
is sensitive to non-normality
beyond indicating that *some*
ring, it sheds no light on u
both G and E as measured
power for detecting $G \times E$;
very informative also, beginr
ing biology. For example, se
APOE $e4$ allele on cognitive
higher $e4$-associated risk th
Additionally, the $e4$ allele mc
in women on cognitive decli
ciated with less cognitive de
the $e4$ risk allele.

In the present paper we
measured E, which is most
study. For example, additive
symptoms interact with ma

Purcell 2002 Twin Research

Purcell 2002 Twin Research

Terminology

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.

If GxE or rGE ignored

- GxE vs rGE

$G \times E$ 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

- Biased parameter estimates
  - **A** x **C** acts like **A**
  - **A** x **E** acts like **E**
  - **A** * **C** acts like **C**
  - **A** * **E** acts like **A**

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

One or more environmental moderators that may

- i) be continuous or binary
- ii) differ between twins in a pair

> 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. 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.

Biometric model & linear A x M

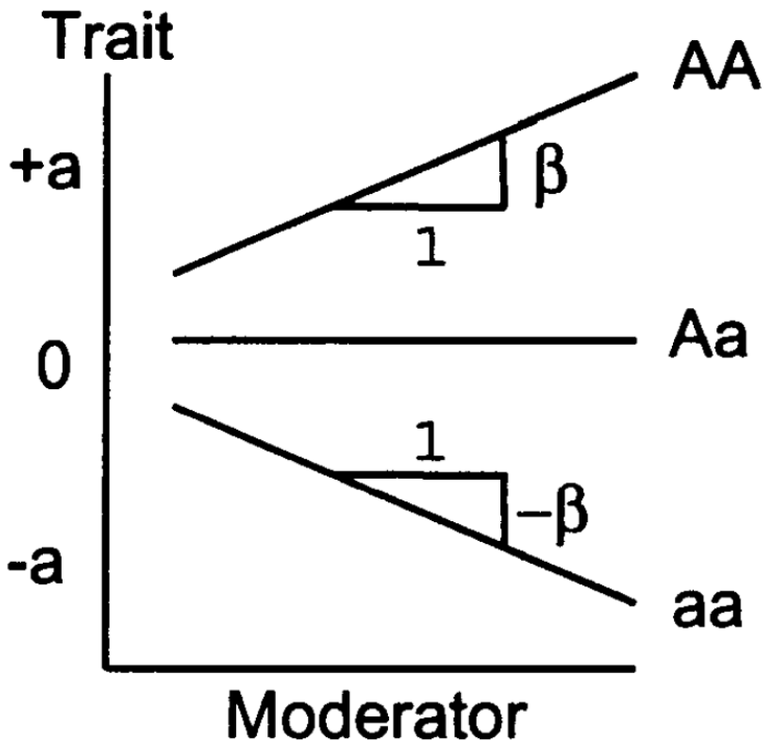


Figure 1

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

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.

One or more environmental moderators that may

iii) interact with residual environmental as well as genetic effects

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

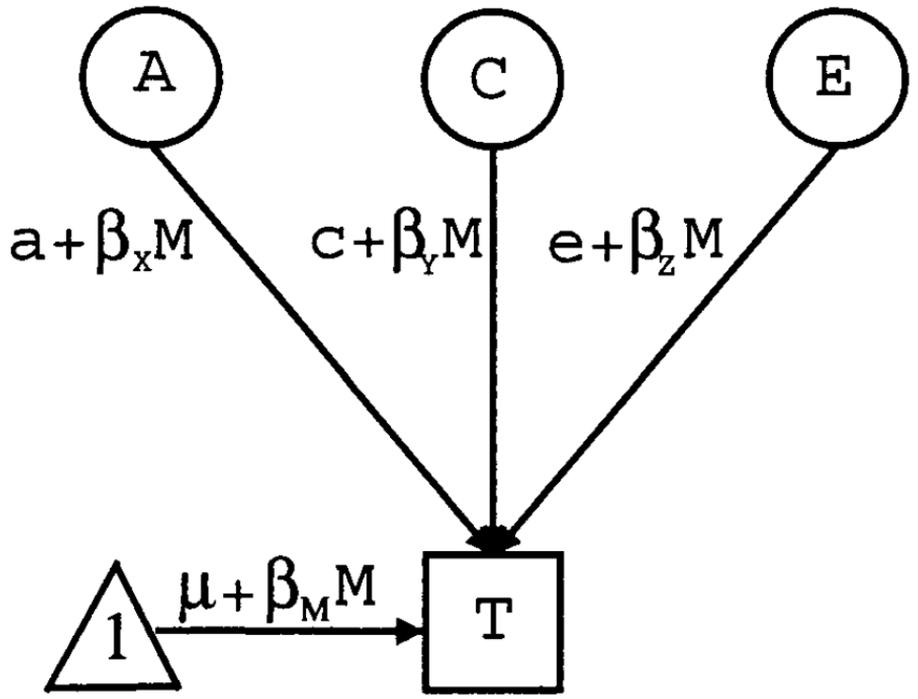


Figure 2

Partial path diagram for the *ACE–XYZ – M* model, shown for one twin only. Latent variables have unit variance.

One or more environmental moderators that may

iv) have non-linear moderators properties

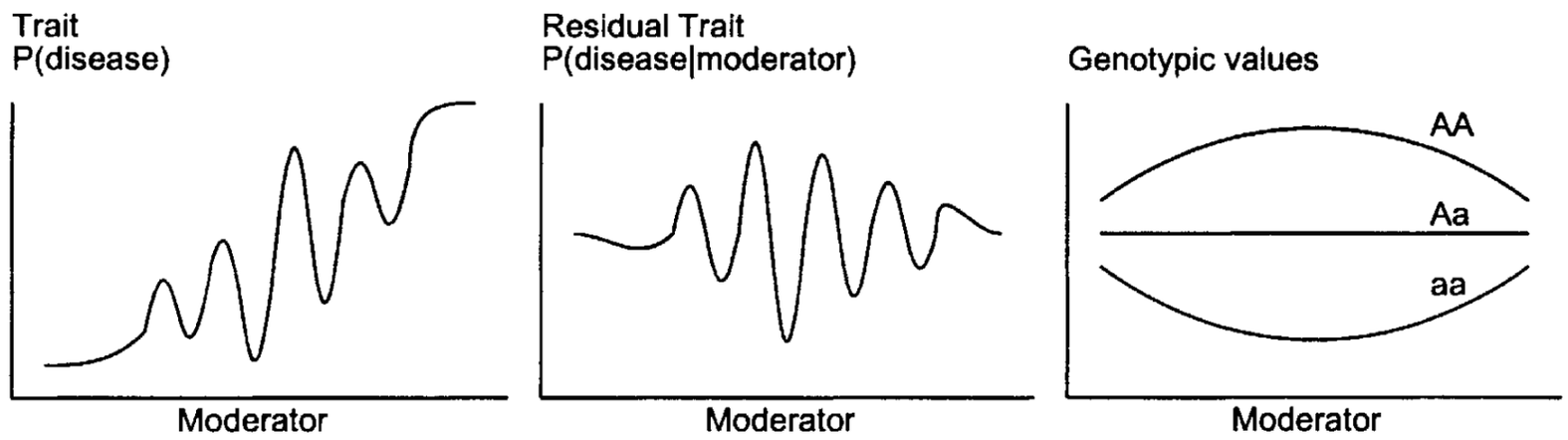


Figure 6

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

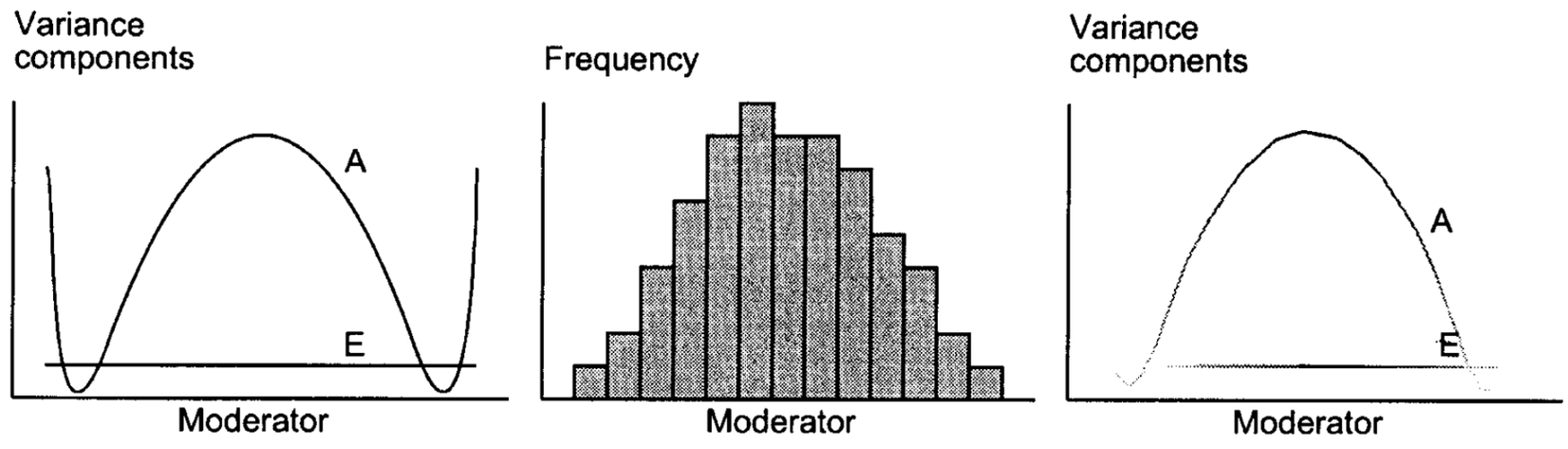


Figure 7

Visualization of variance components for the nonlinear *G* x *E* example.

One or more environmental moderators that may

- v) show scalar (different quantities) or qualitative (different genes) interactions - requires moderator to differ between twins

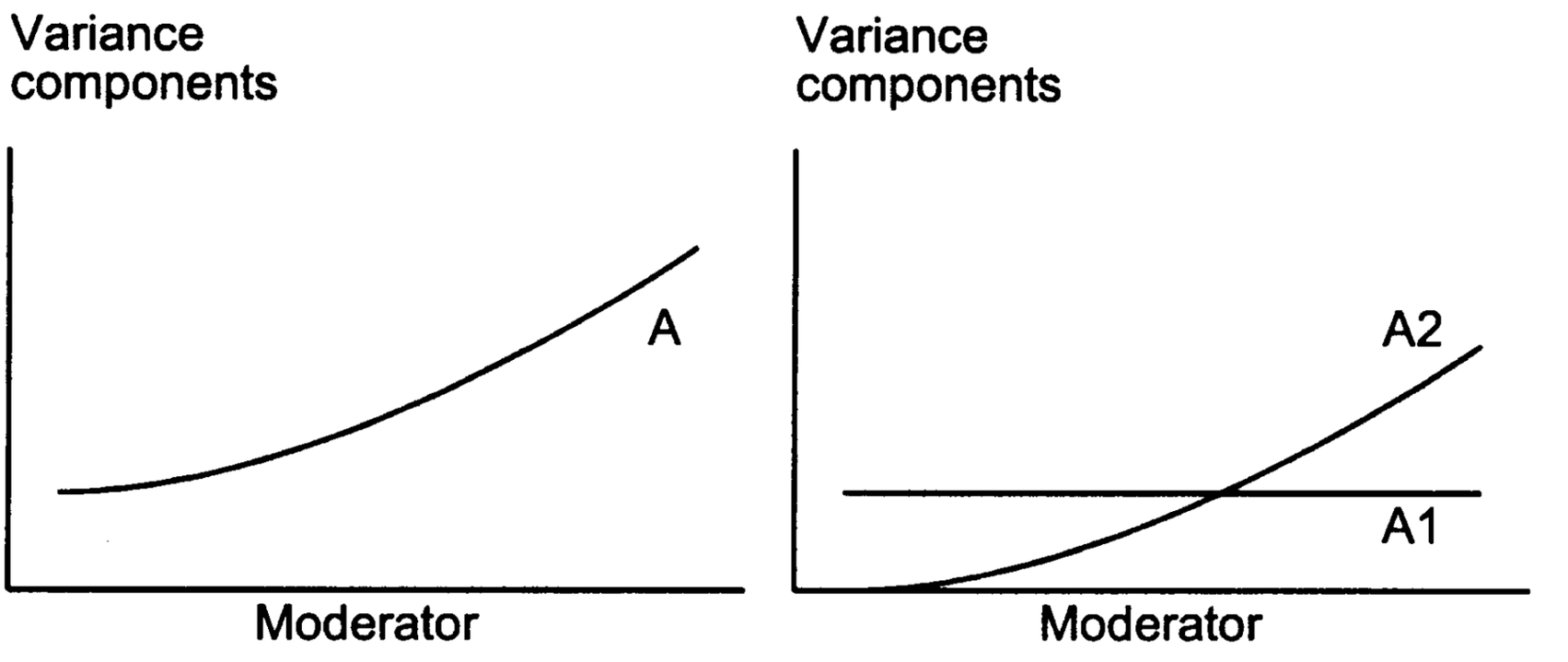


Figure 10

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

One or more environmental moderators that may

- 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

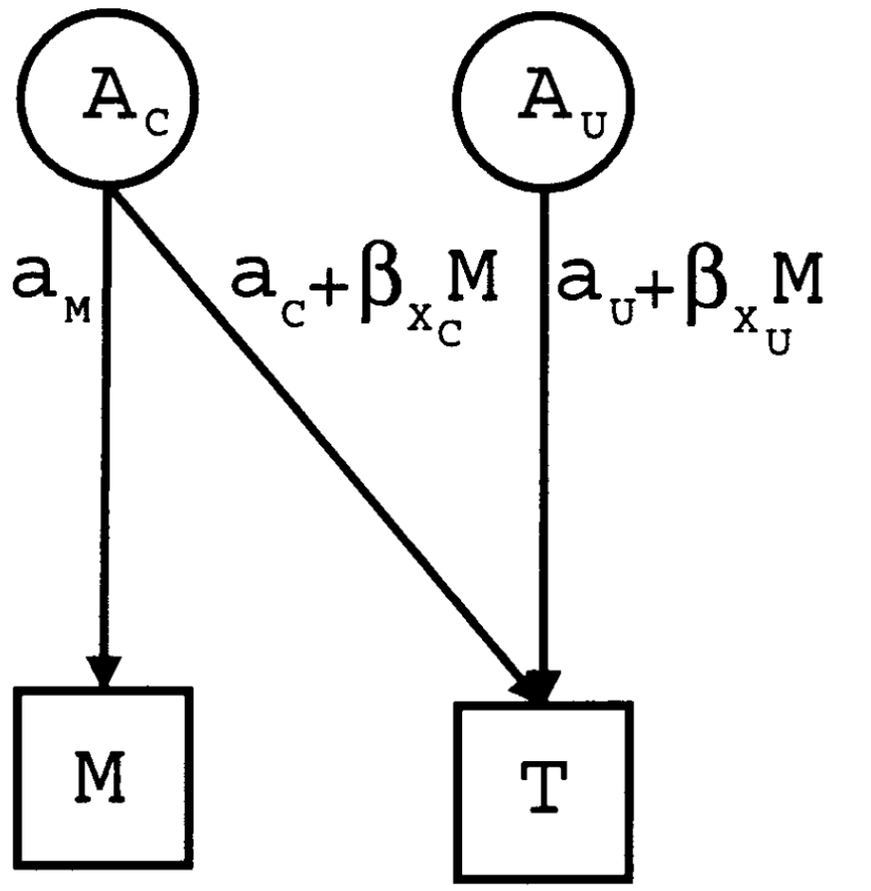


Figure 9

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

Continuous Moderator: Definition Variable Approach

- Is magnitude of genetic/environmental influences on trait the same across all ages? **β*VA** /=0? etc.
- Examples continuous moderators
  - **A** effects moderated by Age = Age **x** **A** interaction
  - Age x **A**,**C**,**E** interaction
  - **A**,**C**,**E** effects on IQ moderated by socio-economic status (SES) (Turkheimer et al 2003)
  - **A** effects on intelligence moderated by SES (Tucker-Drob & Bates 2016)

GxE Application: continuous moderator

PSYCHOLOGICAL SCIENCE

Research Article

SOCIOECONOMIC STATUS MODIFIES HERITABILITY OF IQ IN YOUNG CHILDREN

Eric Turkheimer, Andreana Haley, Mary Waldron, Brian D'Onofrio, and Irving I. Gottesman

University of Virginia

Abstract—*Scores on the Wechsler Intelligence Scale for Children were analyzed in a sample of 7-year-old twins from the National Collaborative Perinatal Project. A substantial proportion of the twins were raised in families living near or below the poverty level. Biometric analyses were conducted using models allowing for components attributable to the additive effects of genotype, shared environment, and nonshared environment to interact with socioeconomic status (SES) measured as a continuous variable. Results demonstrate that the proportions of IQ variance attributable to genes and environment vary nonlinearly with SES. The models suggest that in impoverished families, 60% of the variance in IQ is accounted for by the shared environment, and the contribution of genes is close to zero; in affluent families, the result is almost exactly the reverse.*

PSYCHOLOGICAL SCIENCE

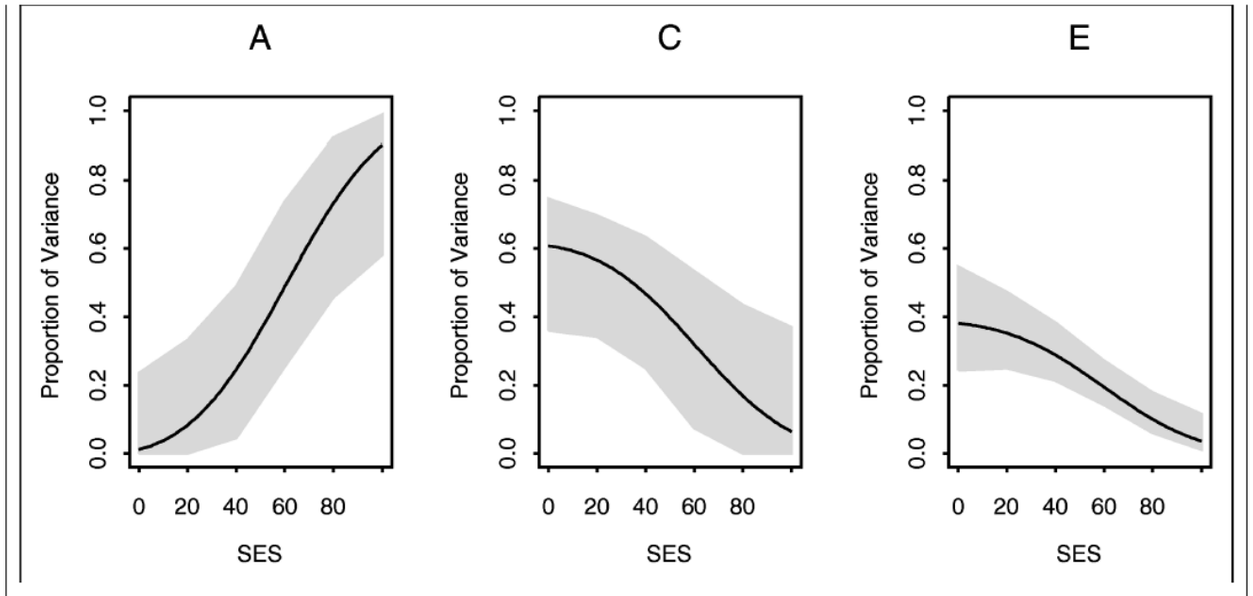


Fig. 3. Proportion of total Full-Scale IQ variance accounted for by A, C, and E plotted as a function of observed socioeconomic status (SES). Shading indicates 95% confidence intervals.

Turkheimer et al. 2003 Psychological Science

GxE Application: continuous moderator

Research Article

Psychological Science
2016, Vol. 27(2) 138–149

DOI: 10.1177/0956797615612727
pss.sagepub.com

Large Cross-National Differences in Gene × Socioeconomic Status Interaction on Intelligence

Elliot M. Tucker-Drob[1,2] and Timothy C. Bates[3]
[1]Department of Psychology, University of Texas at Austin; [2]Population Research Center, University of Texas at Austin; and [3]Department of Psychology, University of Edinburgh

Abstract
A core hypothesis in developmental theory predicts that genetic influences on intelligence and academic achievement are suppressed under conditions of socioeconomic privation and more fully realized under conditions of socioeconomic advantage: a Gene × Childhood Socioeconomic Status (SES) interaction. Tests of this hypothesis have produced apparently inconsistent results. We performed a meta-analysis of tests of Gene × SES interaction on intelligence and academic-achievement test scores, allowing for stratification by nation (United States vs. non–United States), and we conducted rigorous tests for publication bias and between-studies heterogeneity. In U.S. studies, we found clear support for moderately sized Gene × SES effects. In studies from Western Europe and Australia, where social policies ensure more uniform access to high-quality education and health care, Gene × SES effects were zero or reversed.

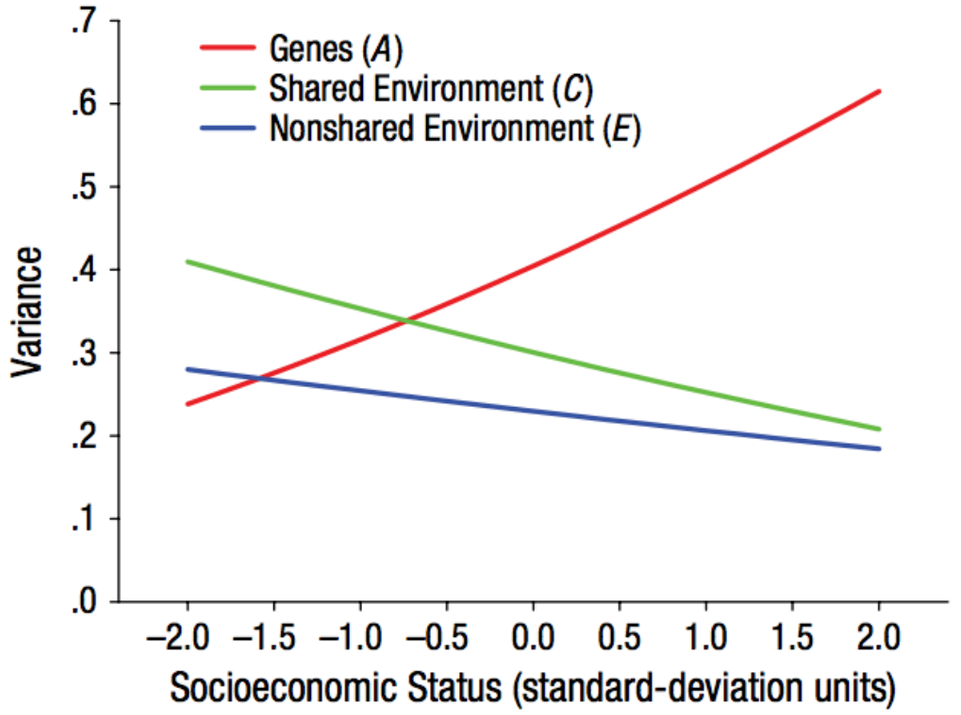


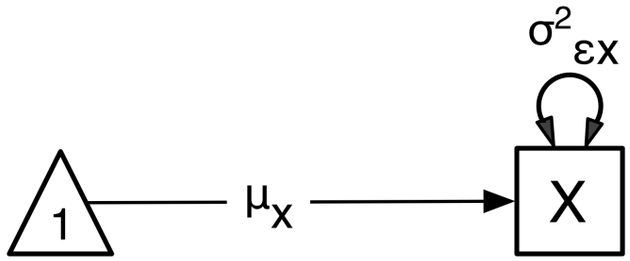
Fig. 1. Variance in cognitive-test performance for the U.S. sample accounted for by genetic and environmental factors, graphed as a function of socioeconomic status (SES). Cognitive test scores were standardized to a *z* scale within each data set prior to model fitting. This plot is very close to (but not identical with) a plot in which the *y*-axis represents the instantaneous proportion of variance for each level of SES.

Tucker-Drob & Bates 2016 Psychological Science

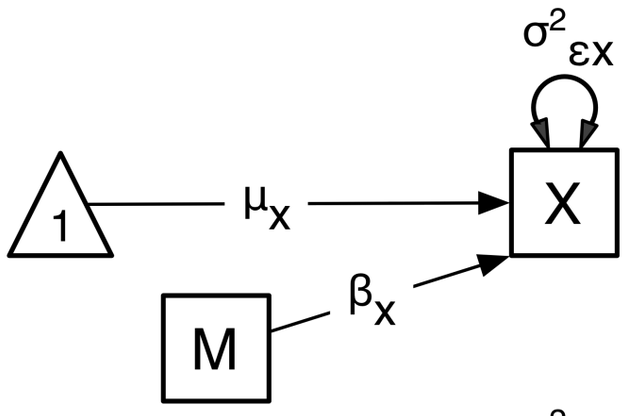
Definition Variables in classic Mx or OpenMx

- definition variables are variables that may vary per subject/pair and are not dependent variables
- Represented by diamond in path diagram
- Model main effects of covariates on means
  - Regression of phenotype on covariate: Moderator (M)
  - Residual variance, subject to ACE modeling
- Model changes in variance components as function of moderator variable (e.g. age, SES)

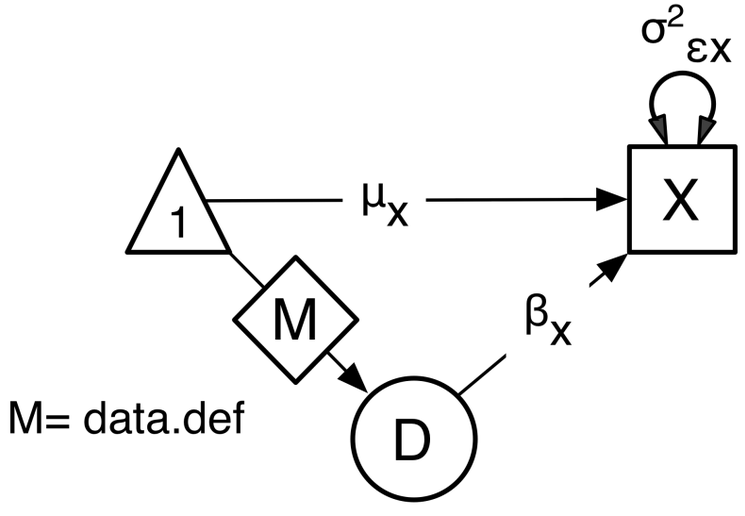
Definition variables in Path Diagrams



Main effect **μx** (intercept)

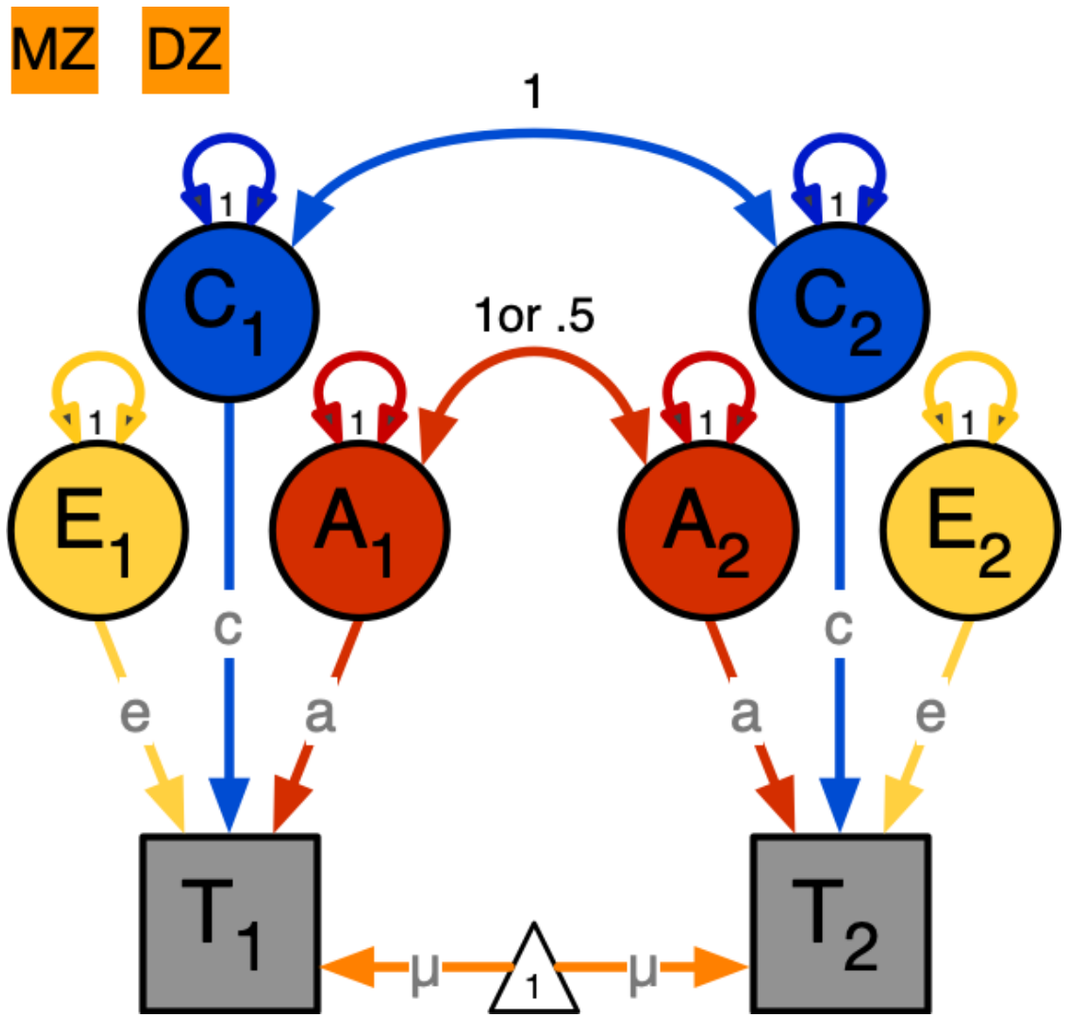


Main effect **μx** +
regression on moderator βx * **M**



Main effect **μx** +
regression on moderator
using definition variable βx * 

Standard ACE Model



r_z = 1 or 0.5

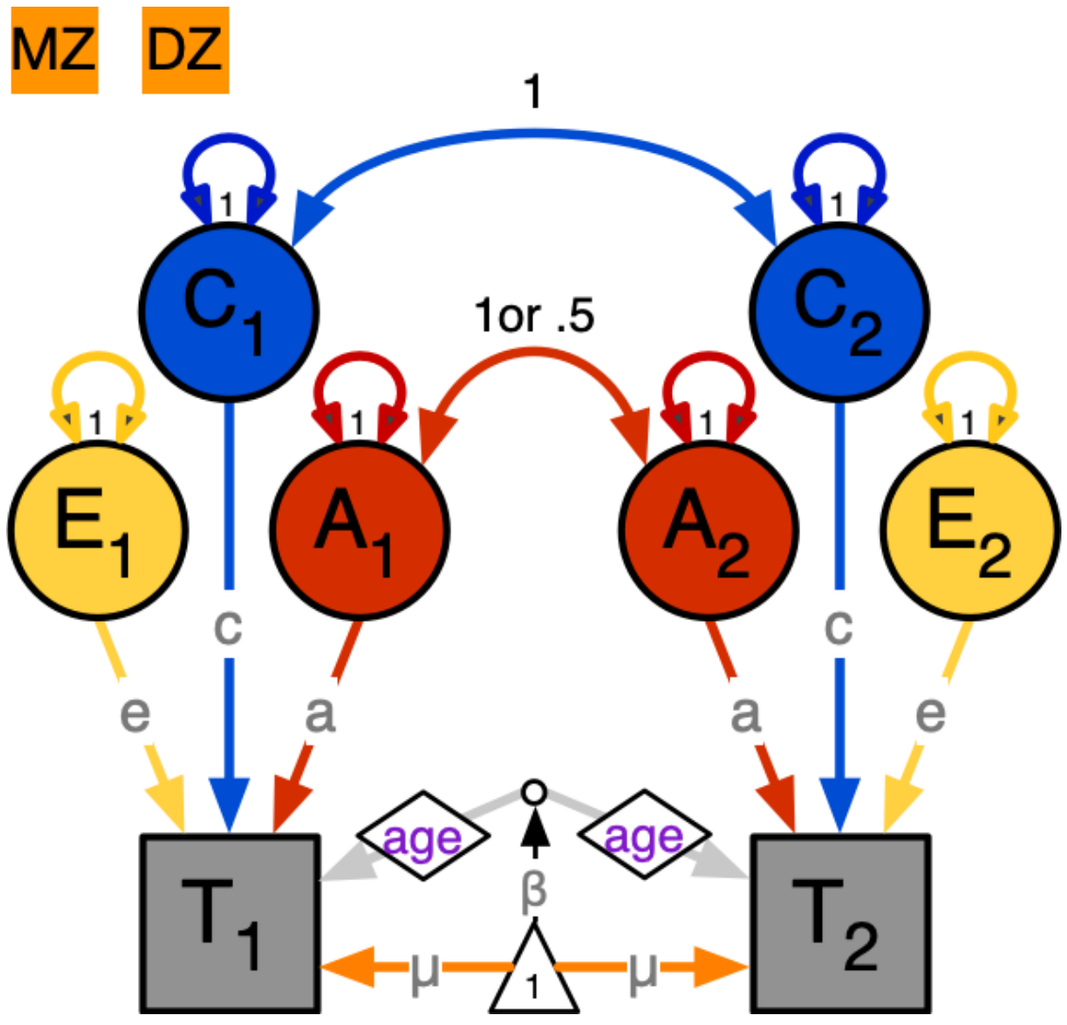
Means

Twin 1	Twin 2
μ	μ

Covariance

	Twin 1	Twin 2
Twin 1	$a^2+c^2+e^2$	$r_z*a^2+c^2$
Twin 2	$r_z*a^2+c^2$	$a^2+c^2+e^2$

Standard ACE Model
Main Effect on Means



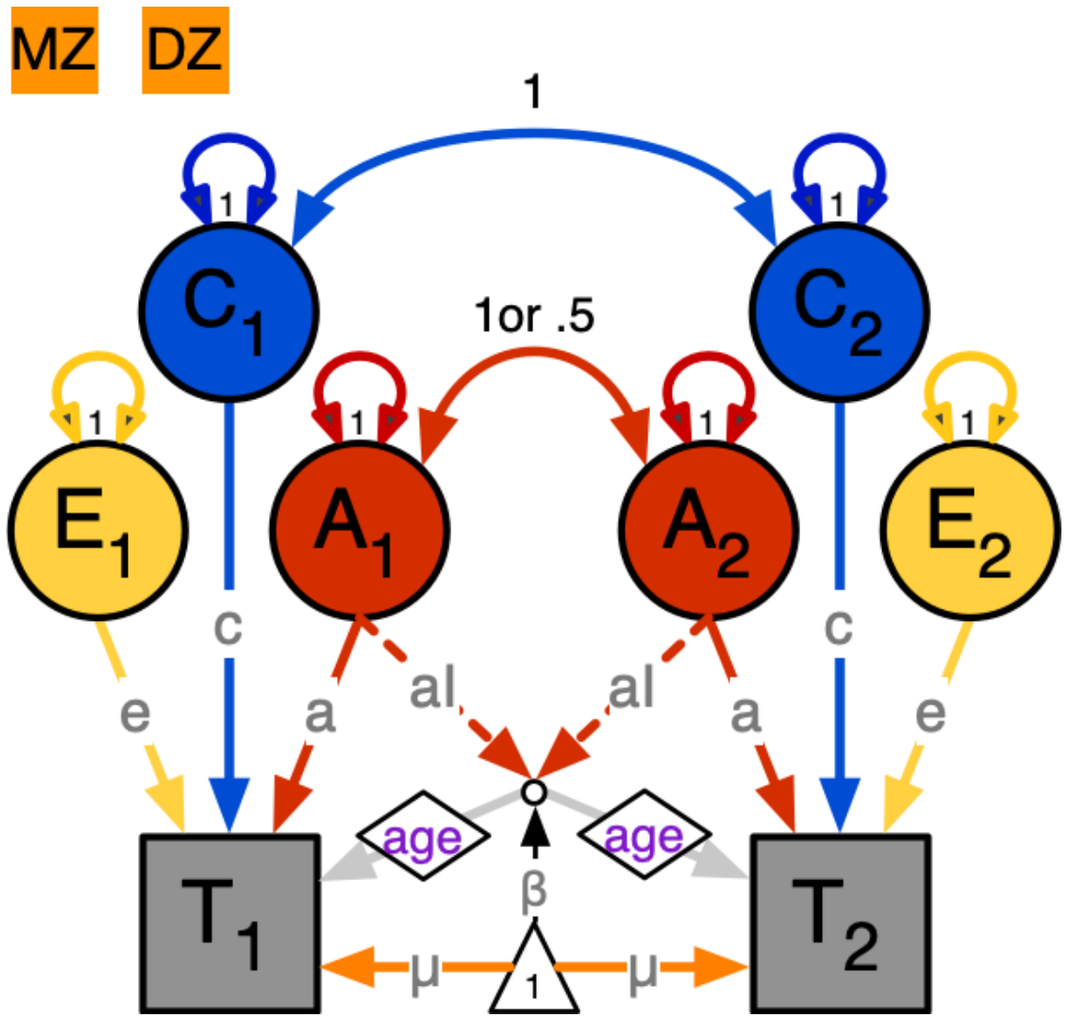
Means

Twin 1	Twin 2
$\mu + \beta(age_i)$	$\mu + \beta(age_i)$

Covariance

	Twin 1	Twin 2
Twin 1	$a^2+c^2+e^2$	$r_z*a^2+c^2$
Twin 2	$r_z*a^2+c^2$	$a^2+c^2+e^2$

Standard ACE Model
Effect on means & on variance component **a**



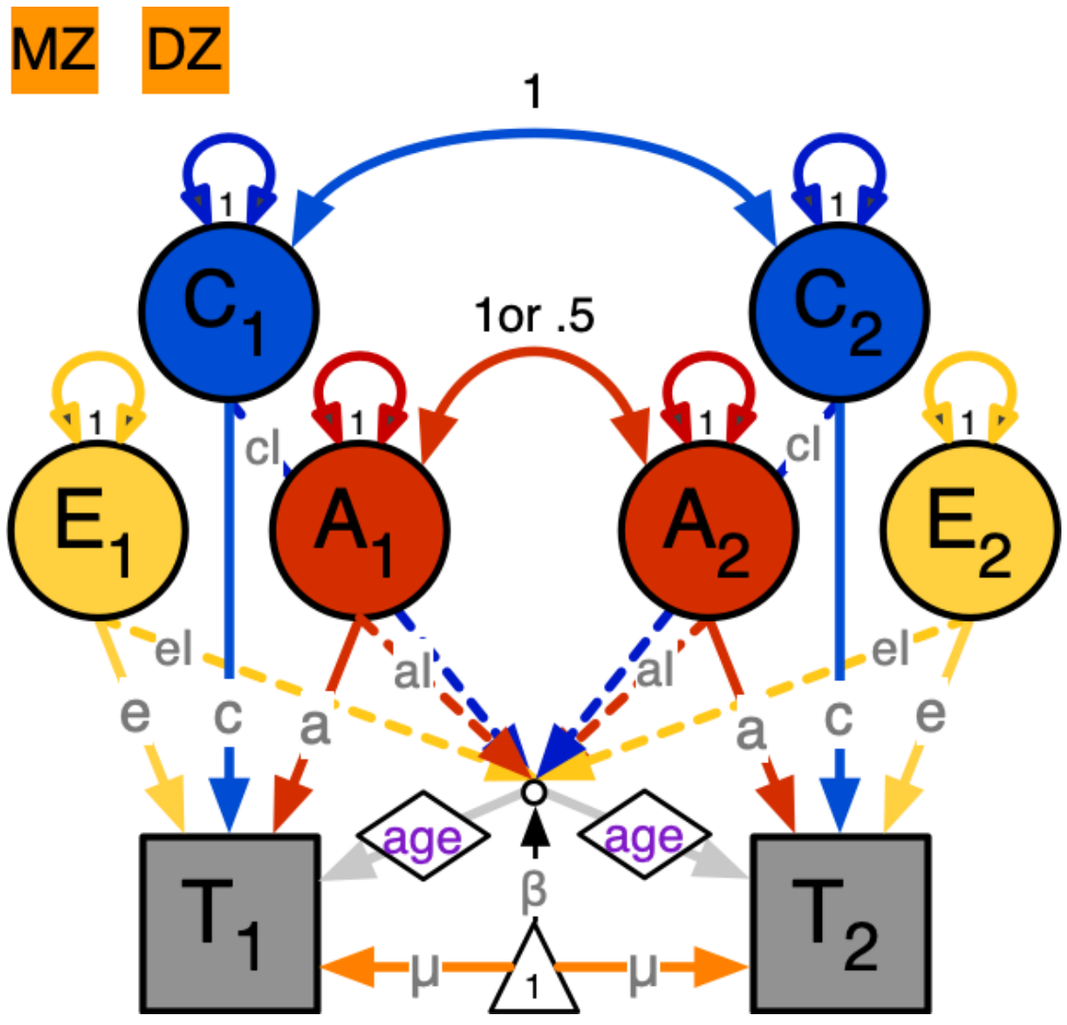
Means

Twin 1	Twin 2
$\mu + \beta(age_i)$	$\mu + \beta(age_i)$

Covariance

	Twin 1	Twin 2
Twin 1	$(a+al*age_i)^2+c^2+e^2$	
Twin 2	$r_z*(a+al*age_i)^2+c^2$	

Standard ACE Model
Effect on means & **a**, **c**, **e**



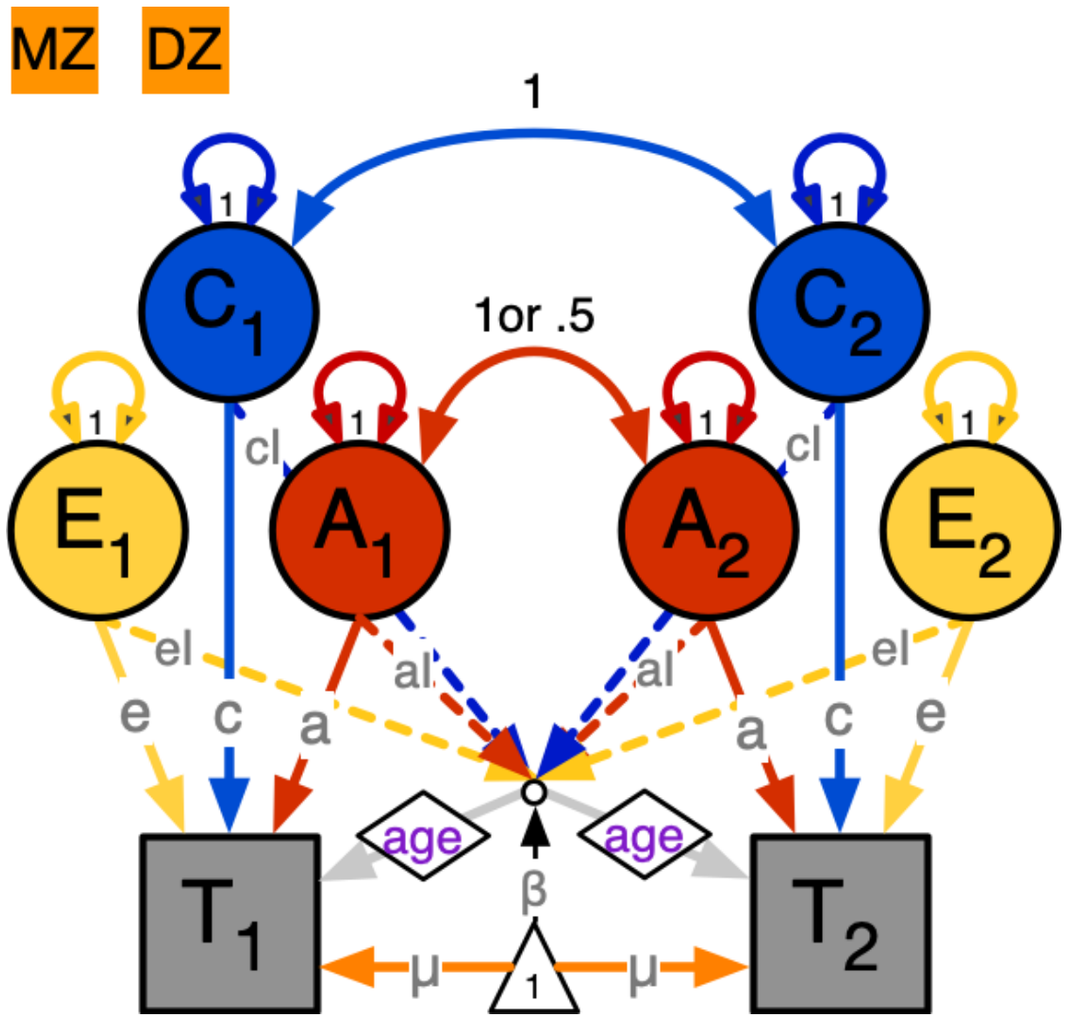
Means

Twin 1	Twin 2
$\mu + \beta(age_i)$	$\mu + \beta(age_i)$

Covariance

	Twin 1	Twin 2
Twin 1	$(a+al*age_i)^2+(c+cl*age_i)^2+(e+el*age_i)^2$	
Twin 2	$r_z*(a+al*age_i)^2+(c+cl*age_i)^2$	

Moderated ACE Model
~ Purcell Path Diagram



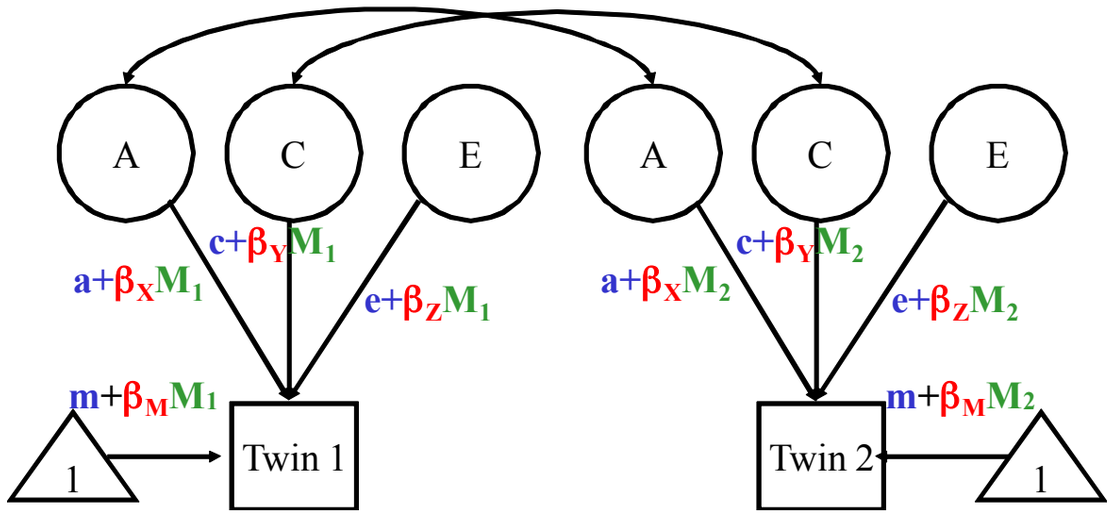
age = M1

$\beta = \beta_M$

$al = \beta_x$

$cl = \beta_y$

$el = \beta_z$



Main Effect on phenotype (β **x M** linear regression)

Moderation effects on path loadings (**A x M**, **C x M**, **E x M** interaction)

Questions:

hermine.maes@vcuhealth.org