Heterogeneity, sex limitation & GxE interaction

Boulder workshop, March 2024

Laura W. Wesseldijk and José J. Morosoli

With credit to Hermine Maes, Sarah Medland and Conor Dolan

Heterogeneity

When there is *variability* or *diversity*; elements are not uniform.

Univariate Analysis:

- What are the contributions of genetic (additive or dominance) and environmental (shared or unique) factors to individual differences in a trait?

Heterogeneity:

- Are the contributions of genetic and environmental factors equal for different groups (such as sex, or country) or for different levels of a covariate (such as SES, urbanicity, environmental exposure, etc.) ?

Heterogeneity

Binary variables: multi-group approach

Continuous variables: moderation model

Terminology depends on research question

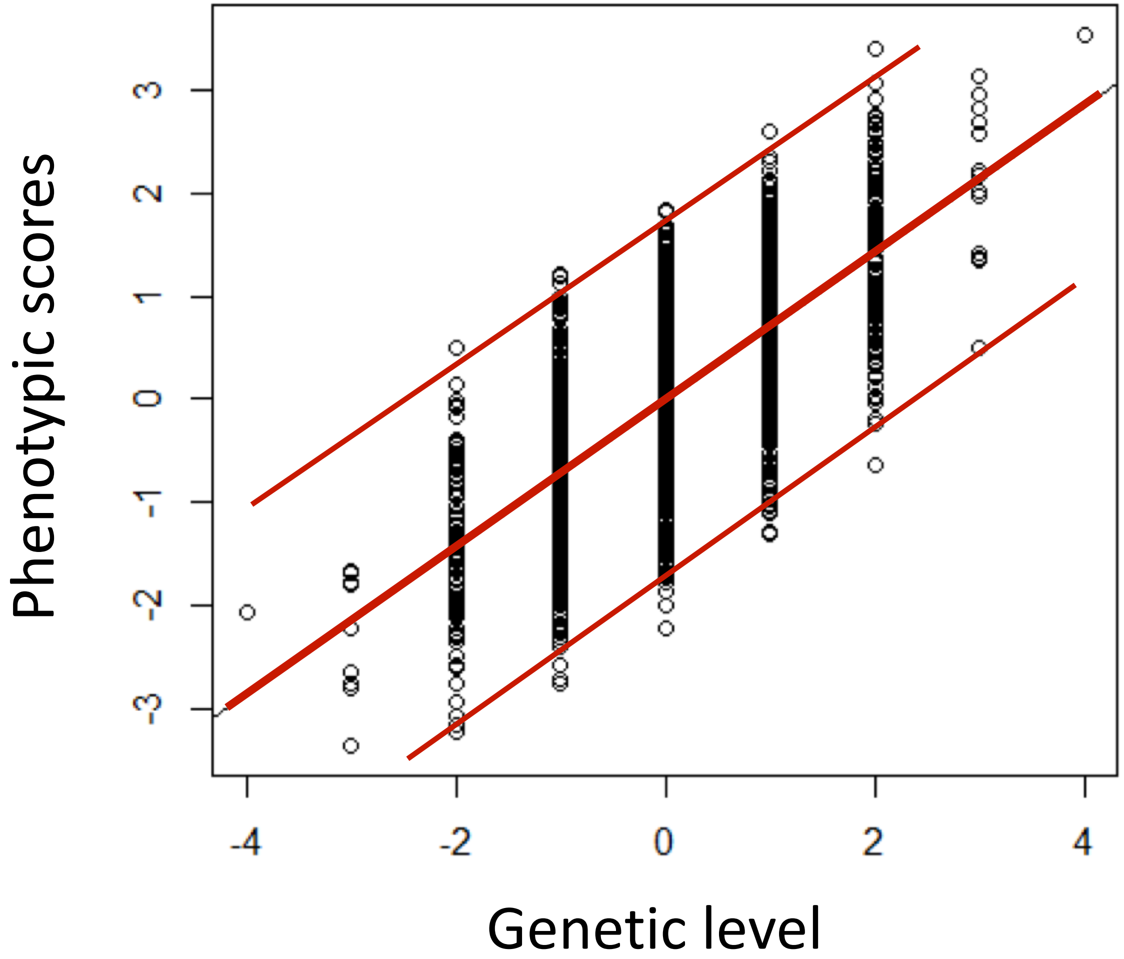
- Moderation, confounding, gene-environment interaction (GxE)

Heterogeneity

Comparison	Concordant group membership	Discordant group membership
sex	MZ & DZ: MM & FF pairs	DZ: opposite sex pairs
country	MZ & DZ: NL & USA pairs	
environment	MZ & DZ: urban & rural pairs	MZ & DZ: 1 urban/ 1 rural
age	MZ & DZ: young & old pairs	

- Differences in magnitude of effects (quantitative)
  - Are the contributions of genetic and environmental factors different in two groups?
- Differences in nature of effects (qualitative)
  - Are different genetic or environmental factors influencing the trait in two groups?

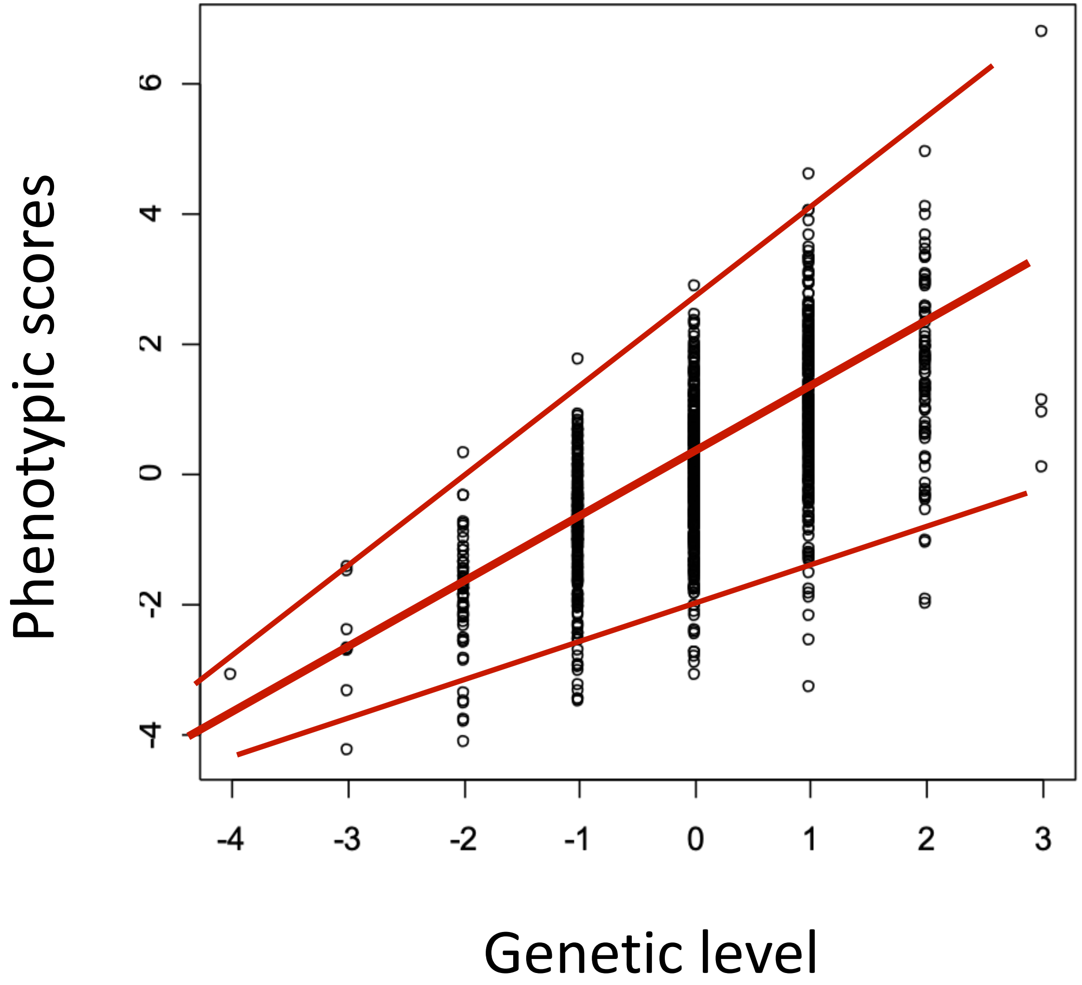
GxE – homoskedastic model



Environmental effects
Variance of **E** given the **genetic level**

Environmental effects are the same given any **genetic level**

GxE – heteroskedastic model

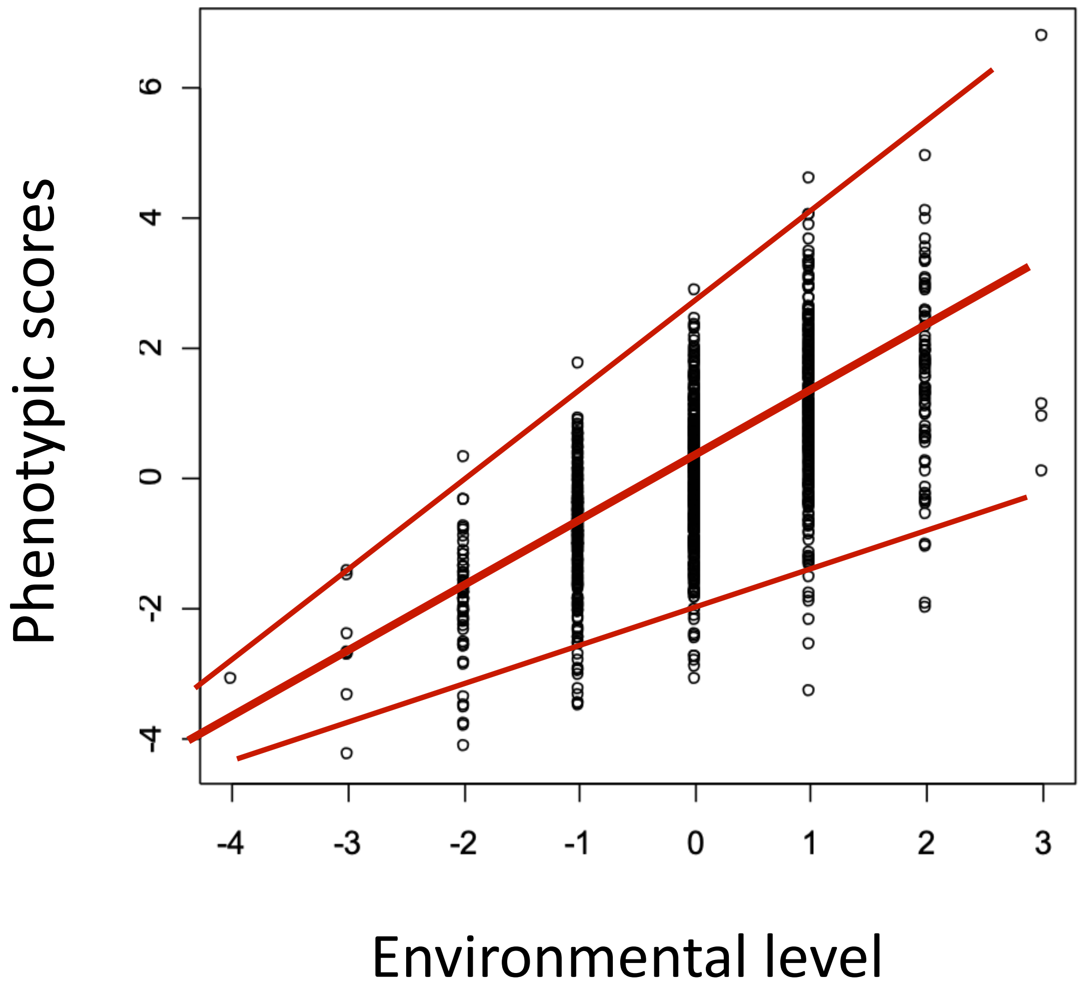


Environmental effects

Variance of **E** given the **genetic level**

Environmental effects systematically vary with the **genetic level**

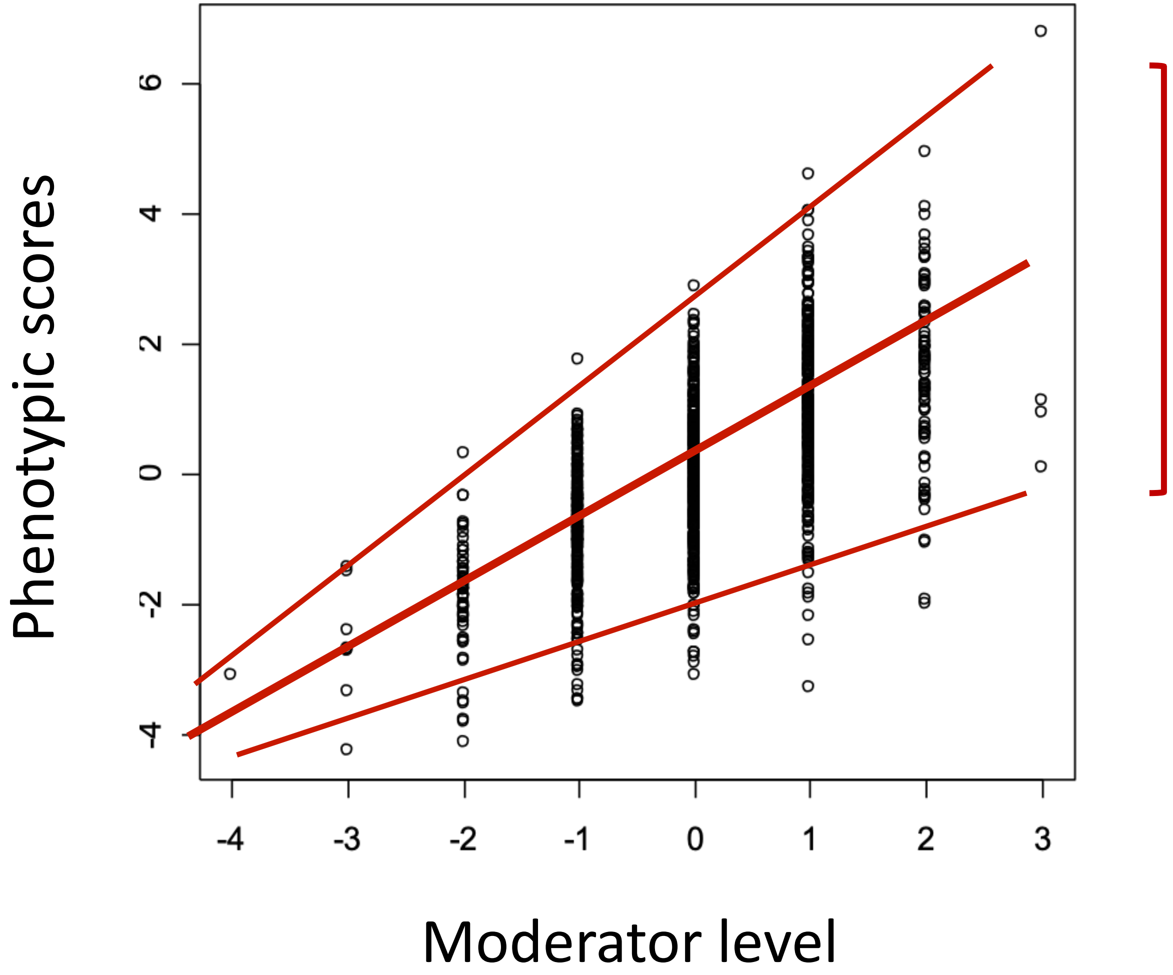
GxE – heteroskedastic model



Genetic effects
Variance of **A** given the **environmental level**

Genetic effects systematically vary with the **environmental level**

GxE – moderation



Variance of **A** or/and **E** given the level of **the moderator**

Genetic effects or environmental effects systematically vary with the level of **M**

Assumptions about moderator

- Moderator is a measured environmental variable
- Environmental measures display genetic variance

BEHAVIORAL AND BRAIN SCIENCES (1991) 14, 373–427
Printed in the United States of America

The nature of nurture: Genetic influence on "environmental" measures

Robert Plomin[a] and C. S. Bergeman[b]

Psychological Medicine, 2007, **37**, 615–626. © 2006 Cambridge University Press
doi:10.1017/S0033291706009524 First published online 19 December 2006 Printed in the United Kingdom

REVIEW ARTICLE

Genetic influences on measures of the environment: a systematic review

KENNETH S. KENDLER[1,2*] AND JESSICA H. BAKER[1,3]

Gene-environment interaction (GxE)

- Gene-environment interaction
  - Genetic control of sensitivity to the environment
  - Environmental control of gene expression

Classic examples:

- The heritability of depression in females depends on marital status (**binary**)
- Socio-economic status modifies the heritability of IQ (**continuous**)

Twin Research (1998) 1, 119–122

http://www.stockton-press.co.uk/tr

Interaction of marital status and genetic risk for symptoms of depression

Andrew C Heath[1], Lindon J Eaves[2] and Nicholas G Martin[3]

[1]Department of Psychiatry, Washington University, St Louis
[2]Virginia Institute of Psychiatric and Behavioral Genetics, Richmond, USA
[3]Queensland Institute of Medical Research, Brisbane, Australia

Depression scores (DSSI) were available for 1232 MZ and 751 DZ female twin pairs who completed a mailed questionnaire. Pairs were divided into those concordant for being in a marriage-like state, concordant for having no partners, and those discordant. The pattern of twin correlations differed according to marital status. Our results suggest that having a marriage-like relationship acts as a protective factor in reducing the impact of inherited liability to symptoms of depression in the general population.

Research Article

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

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

DOI: 10.1177/0956797615612727
pss.sagepub.com
SAGE

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.

Gene-environment correlation (rGE)

- Gene-environment correlation
  - Genetic control of exposure to the environment
  - Environmental control of gene frequency

Types:

- Passive: inherit both genes and childhood environment from parents
- Evocative/reactive: environment reacts to genotype of the child
- Active: genotype causes active behavior to search for an environment

If GxE or rGE ignored

If there is heterogeneity and it is ignored, estimates are biased

Biased parameter estimates

- **A x C** acts like **A**
- **A x E** acts like **E**
- **Correlation A & C** acts like **C**
- **Correlation A & E** acts like **A**

References to read:

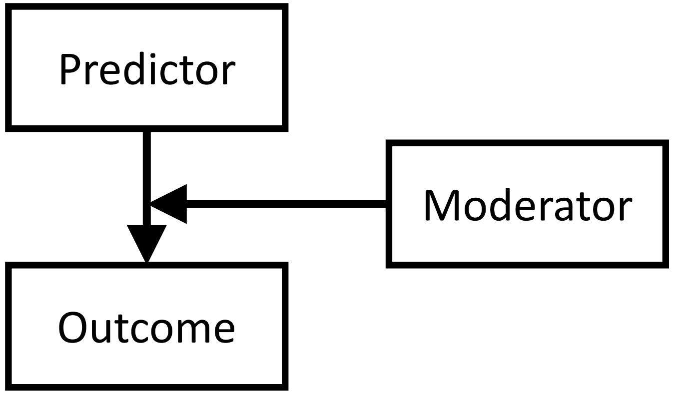
Purcell S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin research : the official journal of the International Society for Twin Studies*, *5*(6), 554–571.

Keller, M. C., & Coventry, W. L. (2005). Quantifying and addressing parameter indeterminacy in the classical twin design. *Twin Research and Human Genetics*, 8(3), 201-213.

Moderation vs. mediation

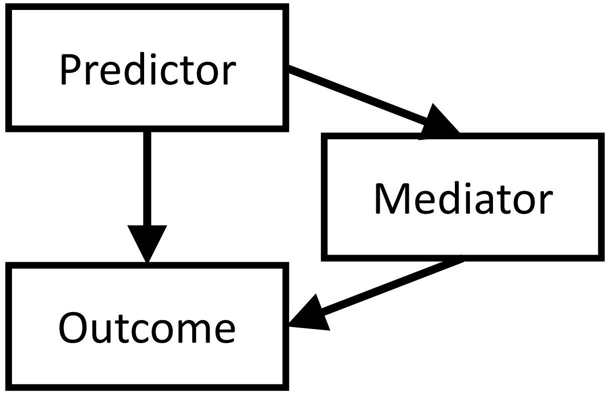
Moderation

- Moderation occurs when the strength or direction of the relationship between two variables depends on a third variable.



Mediation

- Mediation occurs when the relationship between two variables is explained by a third variable.



Heterogeneity

Two in-depth examples:

1. Multigroup approach with a binary moderator: the sex-limitation model with 5 groups
2. Moderation model with a continuous moderator

G x E: multi-group approach, **the sex-limitation model**

- Differences in magnitude of effects (quantitative sex differences)
  - Are the contributions of genetic and environmental factors different in males and females?
- Differences in nature of effects (qualitative sex differences)
  - Are different genetic or environmental factors influencing the trait in males and females?

1948

ON SEX LIMITATION IN HUMAN GENETICS*

By H. HARRIS, M.B., B.Chir.(Camb.)

IT is well known that in many instances of hereditary disease the condition is observed to occur more frequently in one

cases, the sons never inherit the peculiarity directly from their fathers, but the daughters, and the daughters alone, transmit the latent tendency, so that the sons of the daughters

Art. III.

On Sexual Limitation in Hereditary Disease. By William Sedgwick.

(Concluded from our last.)

From hereditary diseases of the organ of vision, the transition is easy to those affecting the organ of hearing, for there are some defects which these organs seem, as it were, to share in common. This connexion has been already referred to by some writers, amongst whom Mr. White Cooper* states that imperfection of the two senses (of sight and hearing) not infrequently co-exist, especially in the curious class of cases we have just been considering, where the inability to distinguish colours is often associated with a corresponding inability to distinguish musical sounds. Dr. Earle relates, in his case of colour-blindness, that "the whole family, of which the chart has been exhibited, is probably no less generally characterized by a defective musical ear than an imperfect apprecia… are utterly incapable of distin…

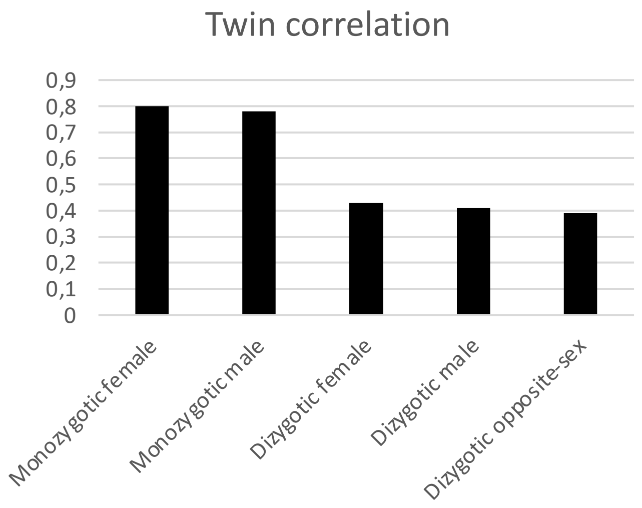
1840

L'HÉRÉDITÉ DANS LES MALADIES,

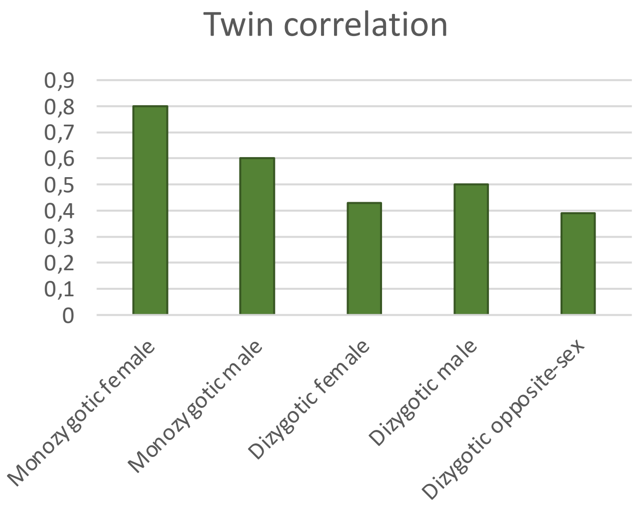
Par P. A. PIORRY,

Docteur en médecine, Chevalier de la Légion-d'Honneur, Médecin de l'Hôpital de la Pitié, Agrégé à la Faculté de Médecine de Paris, Professeur de Clinique et de Pathologie interne, Membre de l'Académie Royale de Médecine, des Sociétés médicales de Tours, de Boulogne, de Gœttingue, de l'Académie Royale de Médecine de Madrid, etc.

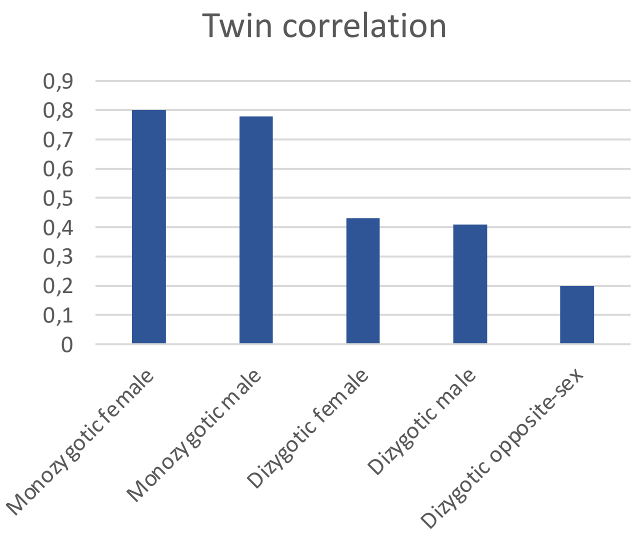
Twin correlations



No sex differences



Quantitative sex differences

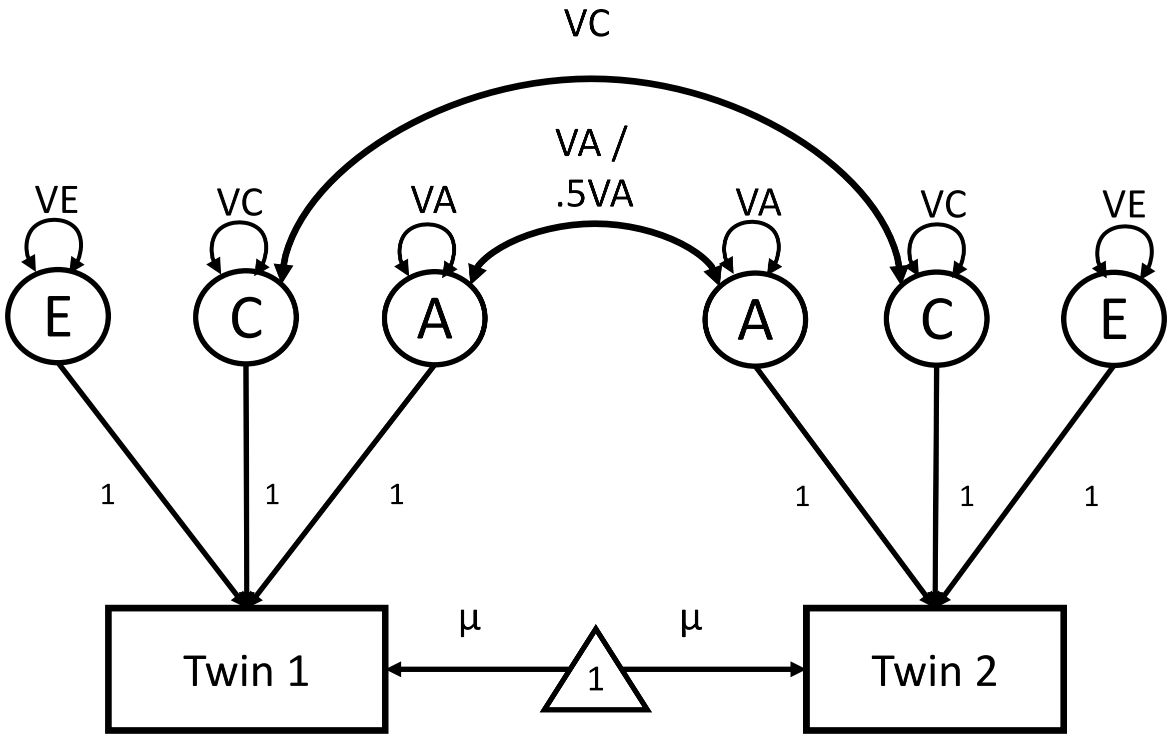


Qualitative sex differences

No Heterogeneity

The classical twin model

No Heterogeneity



rz = 1 for MZ twins and .5 for DZ twins

	Twin 1	Twin 2
Twin 1	**VA + VC + VE**	**rz*VA + VC**
Twin 2	**rz*VA + VC**	**VA + VC + VE**

	Twin 1	Twin 2
Means	**μ**	**μ**

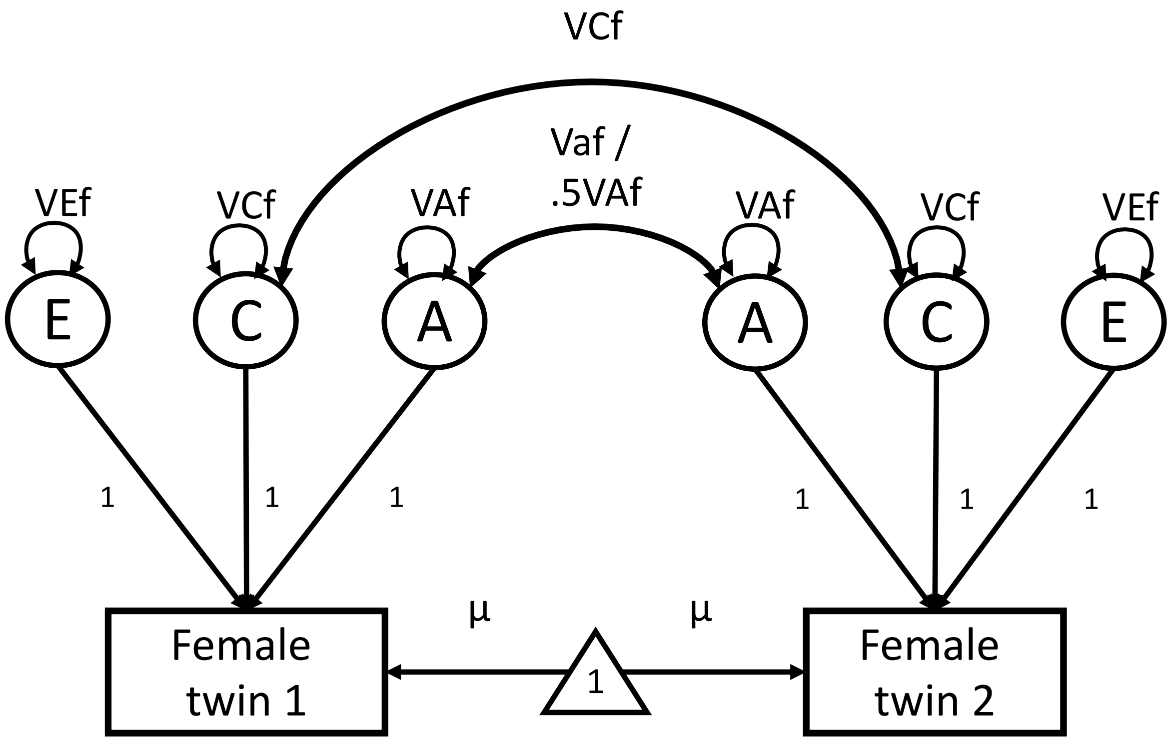
Quantitative

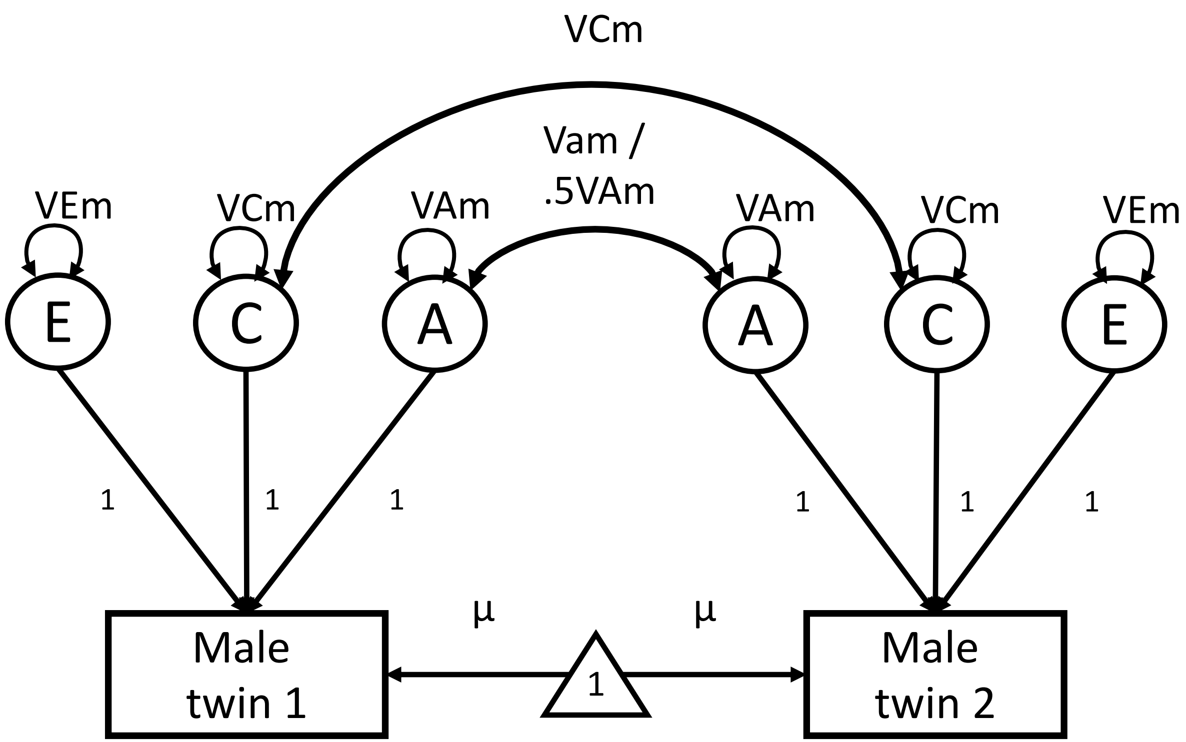
Non-scalar limitation model

- Multigroup approach: 4 groups
  - Monozygotic females
  - Monozygotic males
  - Dizygotic females
  - Dizygotic males
- No opposite-sex twins

- Differences in magnitude of effects (quantitative)
  - Are the contributions of genetic and environmental factors different in two groups?

Quantitative





	Female twin 1	Female twin 2
Female twin 1	**VAf + VCf + VEf**	**rz*VAf + VCf**
Female twin 2	**rz*VAf + VCf**	**VAf + VCf + VEf**

rz = 1 for MZ and .5 for DZ

	Male twin 1	Male twin 2
Male twin 1	**VAm + VCm + VEm**	**rz*VAm + VCm**
Male twin 2	**rz*VAm + VCm**	**VAm + VCm + VEm**

General sex limitation model – option 1

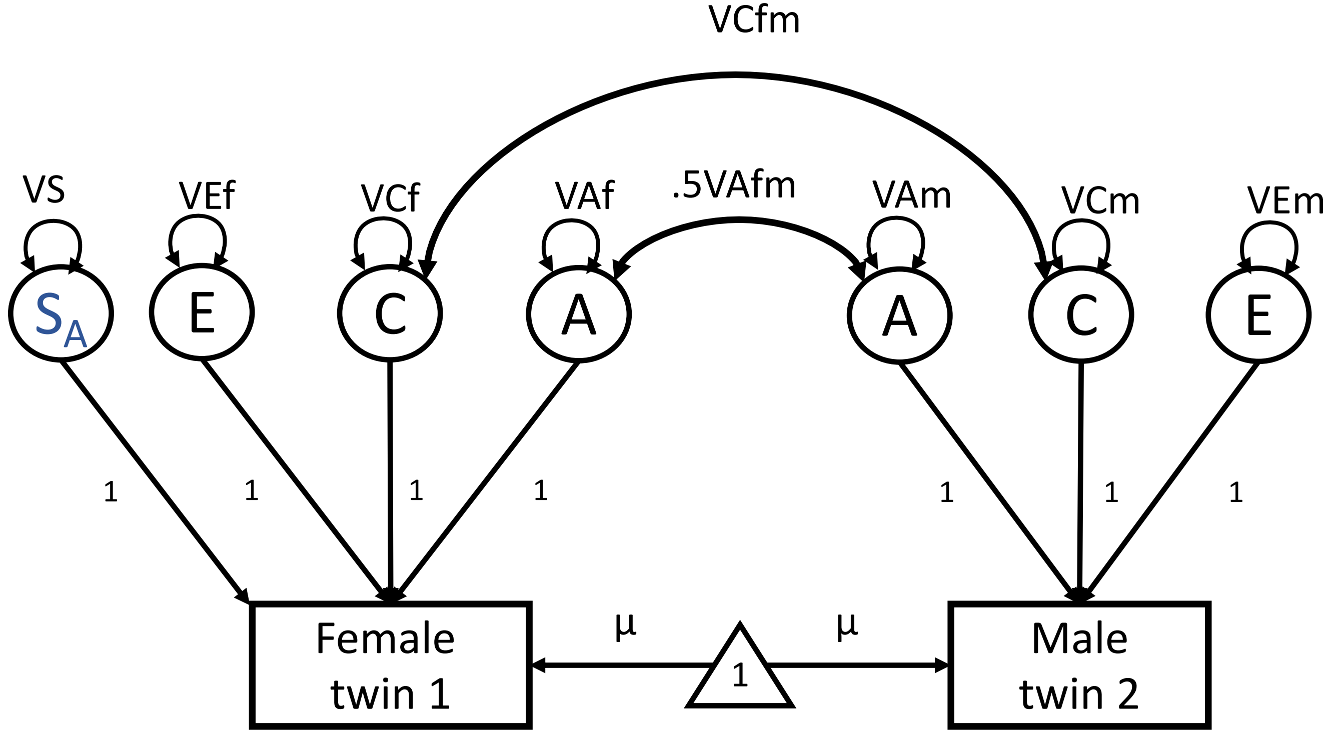
Quantitative / Qualitative

To not only test for quantitative, but also qualitative sex differences

- Multigroup approach: 5 groups
- Method 1: With VA specific for one of the sexes, rg on .5

- Differences in magnitude of effects (quantitative)
  - Are the contributions of genetic and environmental factors different in two groups?
- Differences in nature of effects (qualitative)
  - Are different genetic or environmental factors influencing the trait in two groups?

Quantitative / Qualitative



	Female Twin 1	Male Twin 2
Female Twin 1	**VAf + VCf + Vef +VSaf**	**VAf*.5*VAm + VCf*VCm**
Male Twin 2	**Vaf*.5*VAm + VCf*VCm**	**VAm + VCm + VEm**

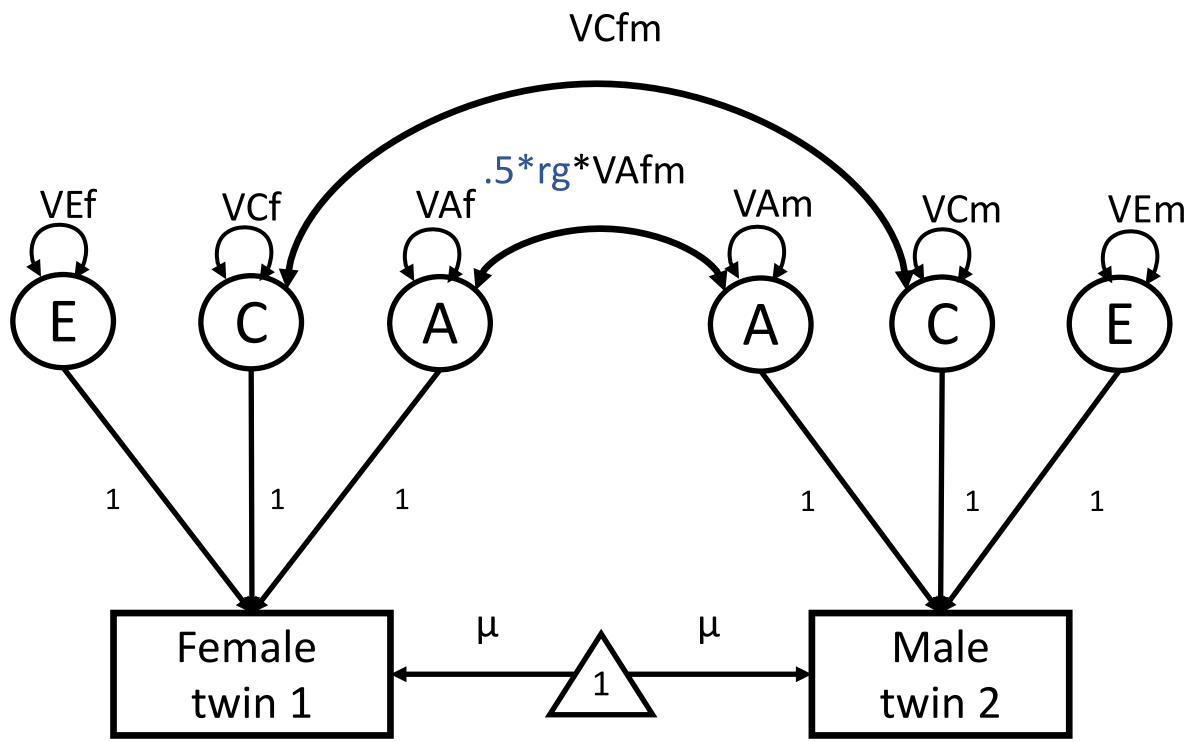
Quantitative / Qualitative

General sex limitation model – option 2

To not only test for quantitative, but also qualitative sex differences

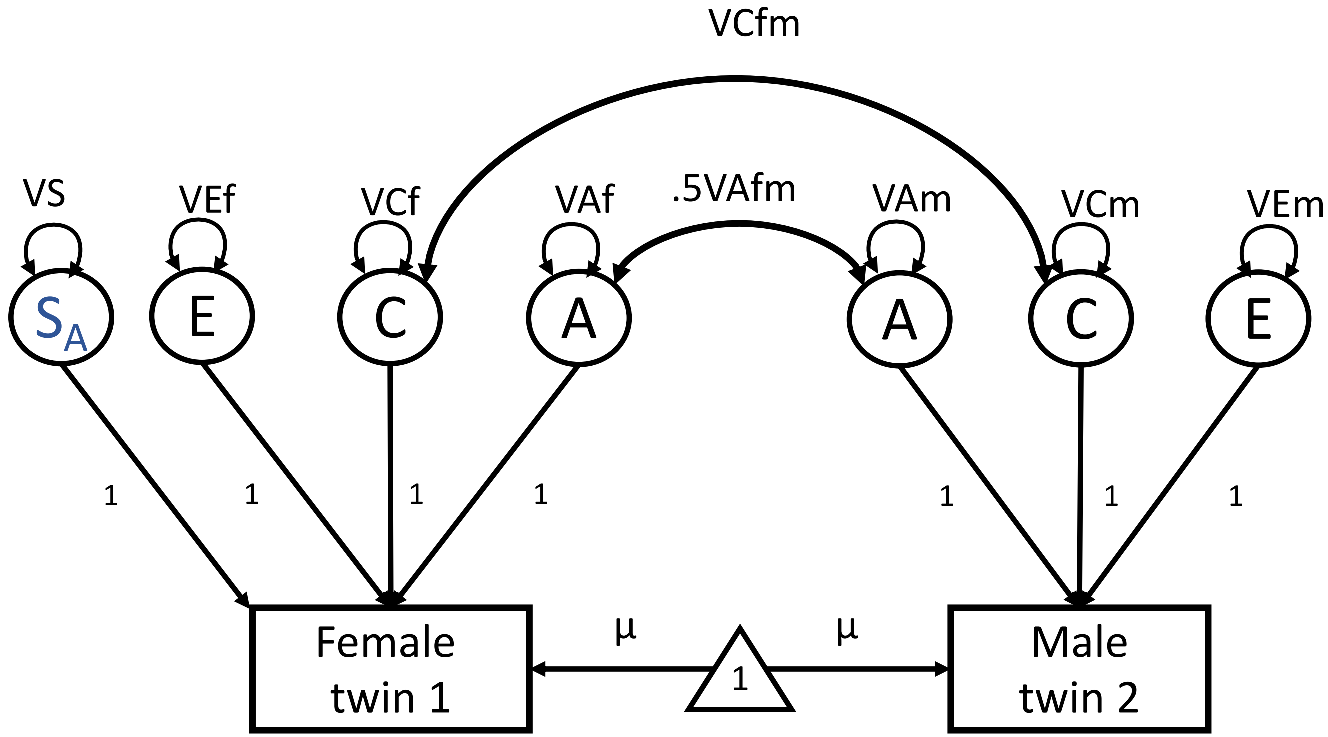
- Multigroup approach: 5 groups
- Method 2: with rg freely estimated

Quantitative / Qualitative



	Female Twin 1	Male Twin 2
Female Twin 1	**VAf + VCf + VEf**	**Vaf*.5*rg*VAm + VCf*VCm**
Male Twin 2	**Vaf*.5*rg*VAm + VCf*VCm**	**VAm + VCm + VEm**

Quantitative / Qualitative



	Female Twin 1	Male Twin 2
Female Twin 1	**VAf + VCf + Vef +VSaf**	**VAf*.5*VAm + VCf*VCm**
Male Twin 2	**Vaf*.5*VAm + VCf*VCm**	**VAm + VCm + VEm**

G x E: continuous variable, **the moderation model**

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.

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 e*4 allele on cognitive decline, where women show higher *e*4-associated risk than men (Yaffe et al., 2000). Additionally, the *e*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 *e*4 risk allele.

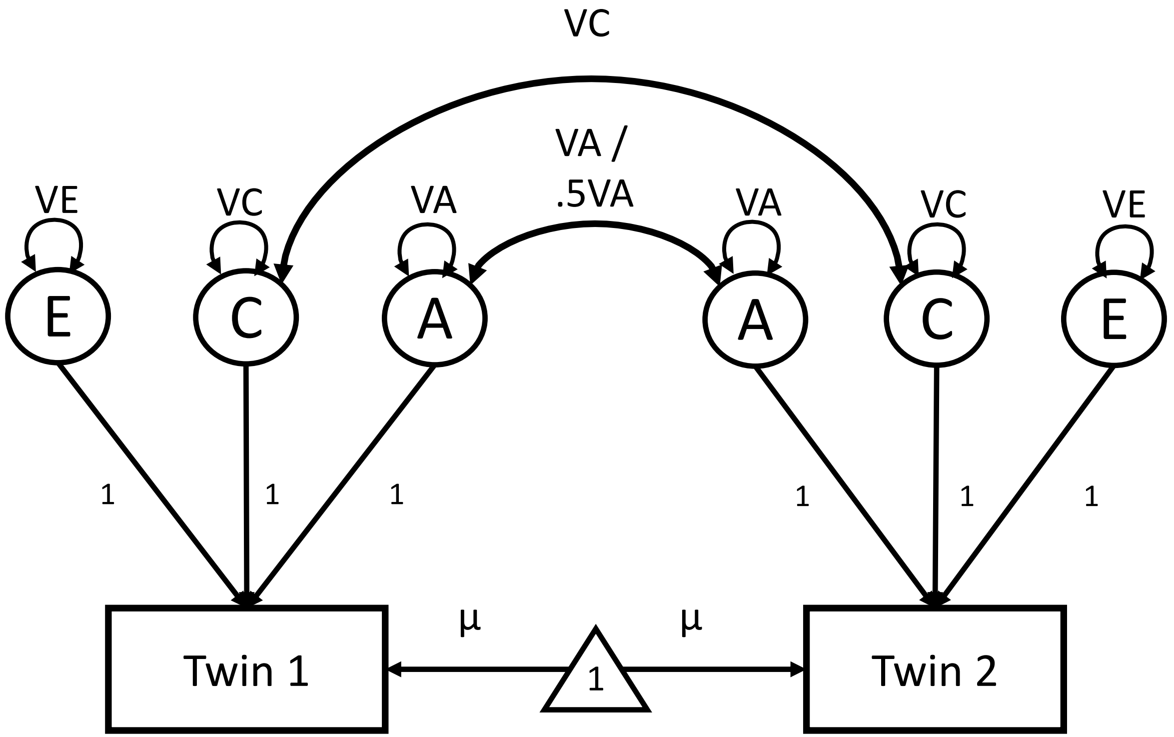
In the present paper we consider the case of latent $G \times$ measured E, which is most relevant to the classical twin study. For example, additive genetic effects on depression symptoms interact with marital status in women, where

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.

No Heterogeneity

The classical twin model

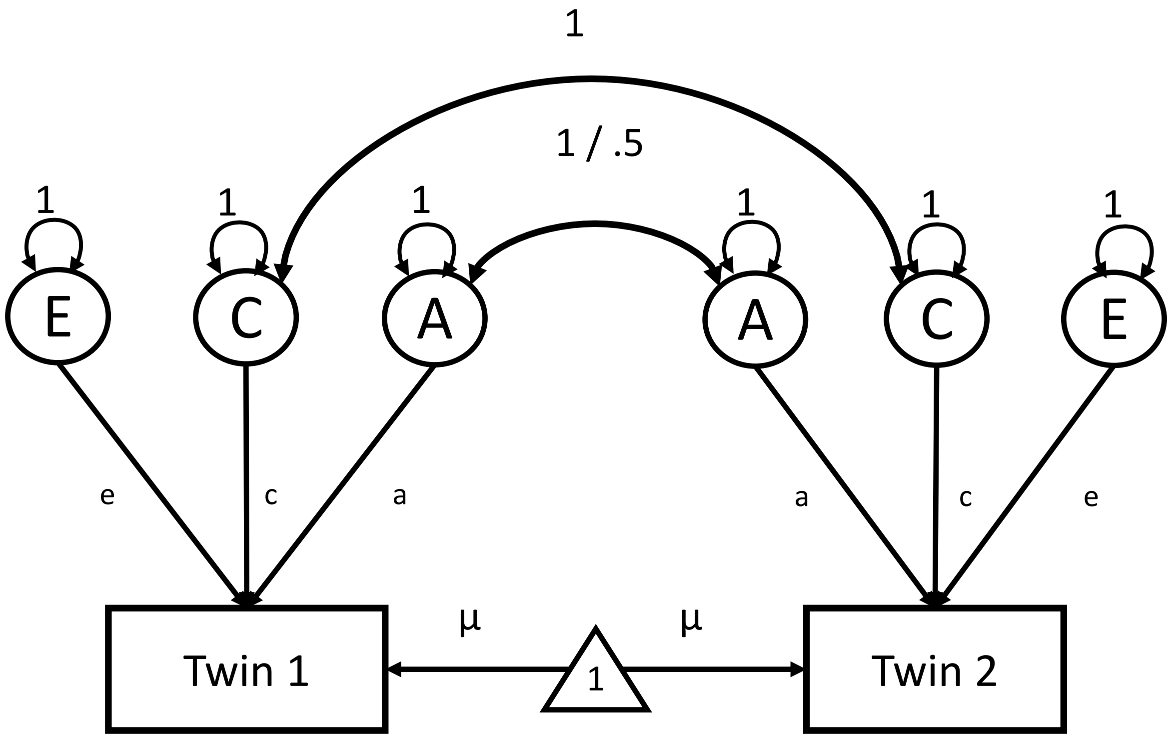
No Heterogeneity



rz = 1 for MZ twins and
.5 for DZ twins

	Twin 1	Twin 2
Twin 1	**VA + VC + VE**	**rz*VA + VC**
Twin 2	**rz*VA + VC**	**VA + VC + VE**

No Heterogeneity



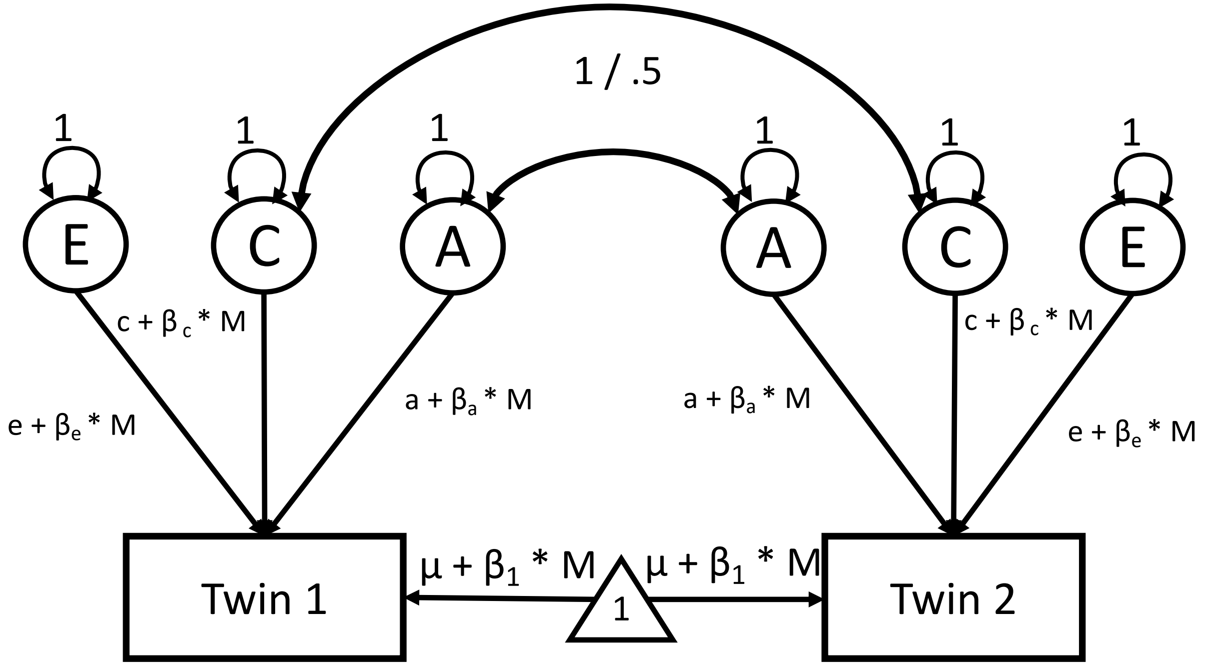
rz = 1 for MZ twins and .5 for DZ twins

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

Quantitative

G x E: the moderation model

Purcell, 2002



	Twin 1	Twin 2
Twin 1	$(a+\beta_a*M)^2+(c+\beta_c*M)^2+(e+\beta_e*M)^2$	
Twin 2	$rz*(a+\beta_a*M)^2+(c+\beta_c*M)^2$	

rZ = 1 for MZ twins and

.5 for DZ twins

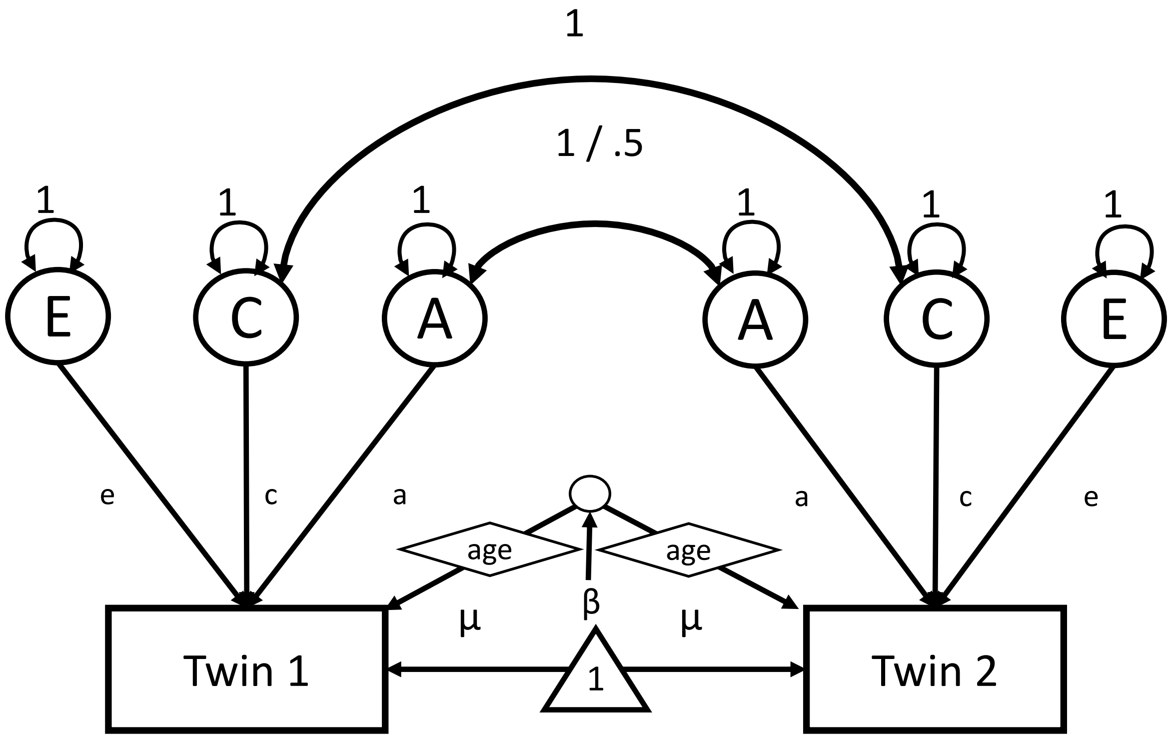
The moderation model in OpenMx

How do we test for GxE with a continuous moderator in OpenMx?

With the use of definition variables

Definition variables

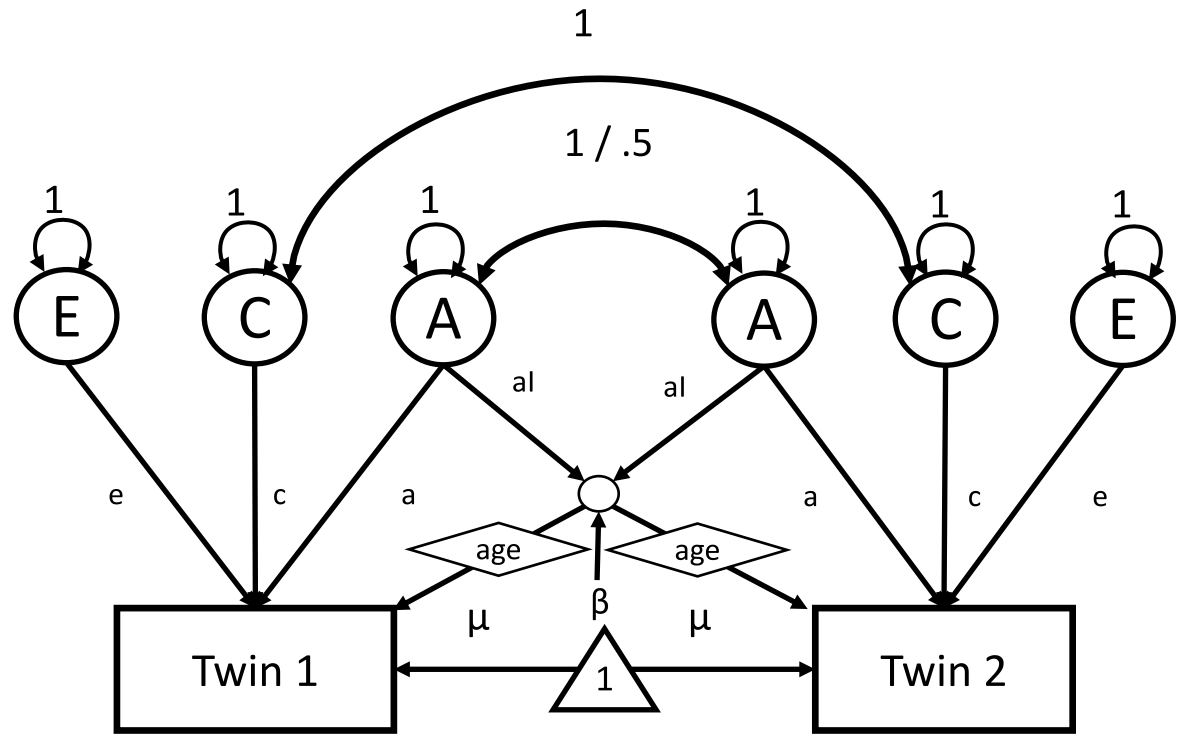
- Variables that may vary per subject/pair and are not dependent variables (diamond in path diagrams)
  - Model main effects of covariates on means (moderator)
  - Residual variance can be partioned into ACE
  - But also ACE as a function of the moderator



r_Z = 1 for MZ twins and .5 for DZ twins

	Twin 1	Twin 2
Twin 1	$\mathbf{a^2+c^2+e^2}$	$\mathbf{rz*a^2+c^2}$
Twin 2	$\mathbf{rz*a^2+c^2}$	$\mathbf{a^2+c^2+e^2}$

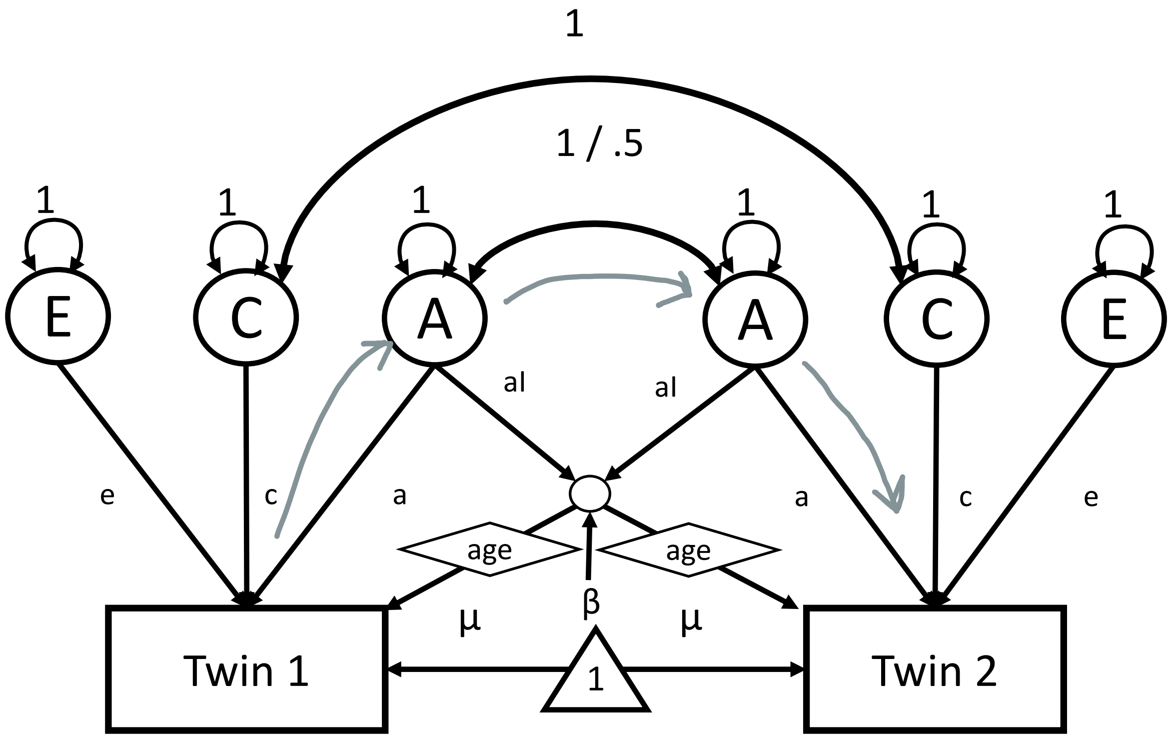
	Twin 1	Twin 2
Means	**μ + $\beta(age_i)$**	**μ + $\beta(age_i)$**



r_Z = 1 for MZ twins and .5 for DZ twins

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

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



$rz*(a+al*age_i)\ (a+al*age_i)$

=

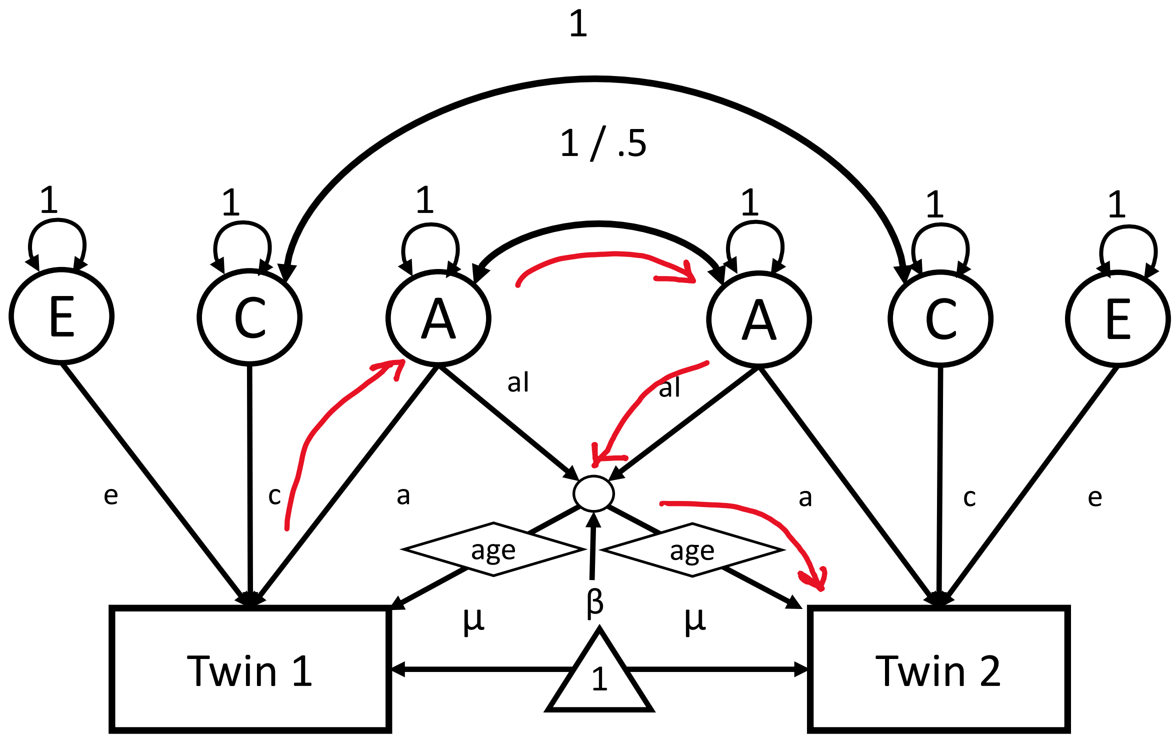
$rz*a*a$

r_Z = 1 for MZ twins and

.5 for DZ twins

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

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



rz*(a+al*agei) (a+al*agei)

=

rz*a*a +

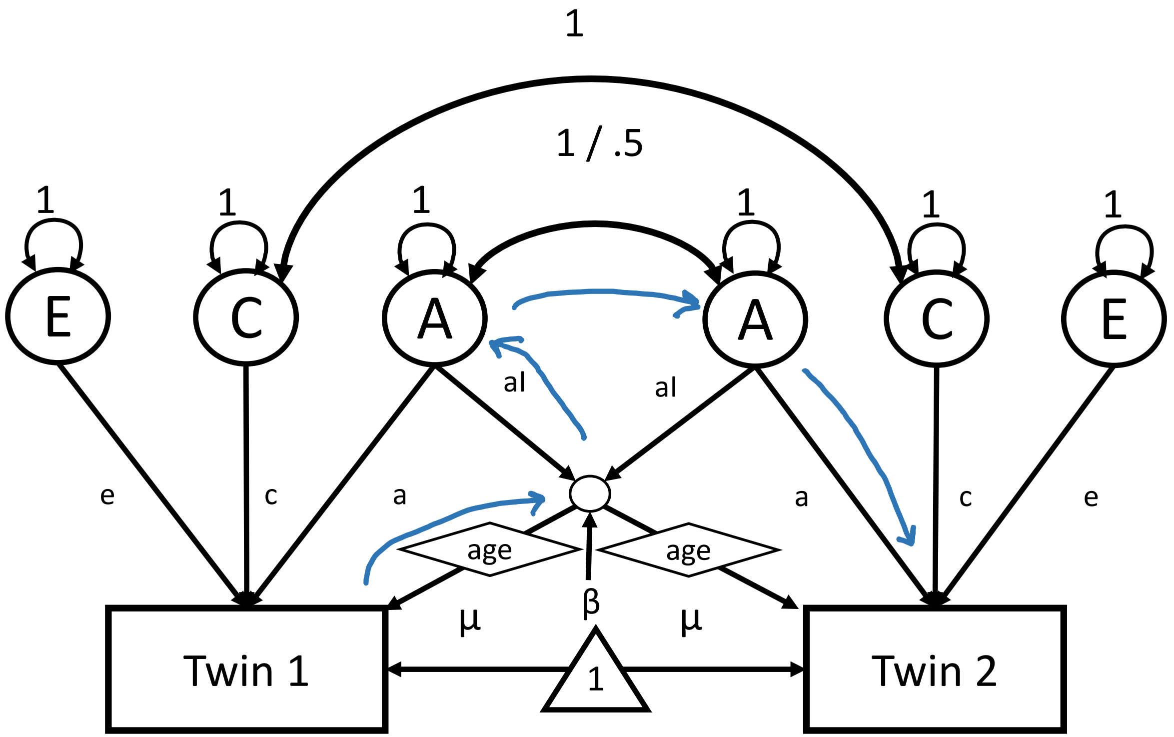
rz*a*al *age

rz = 1 for MZ twins and

.5 for DZ twins

	Twin 1	Twin 2
Twin 1	**(a+al*agei)²+c²+e²**	**rz*(a+al*agei)²+c²**
Twin 2	**rz*(a+al*agei)²+c²**	**(a+al*agei)²+c²+e²**

	Twin 1	Twin 2
Means	**μ + β(age$_i$)**	**μ + β(age$_i$)**

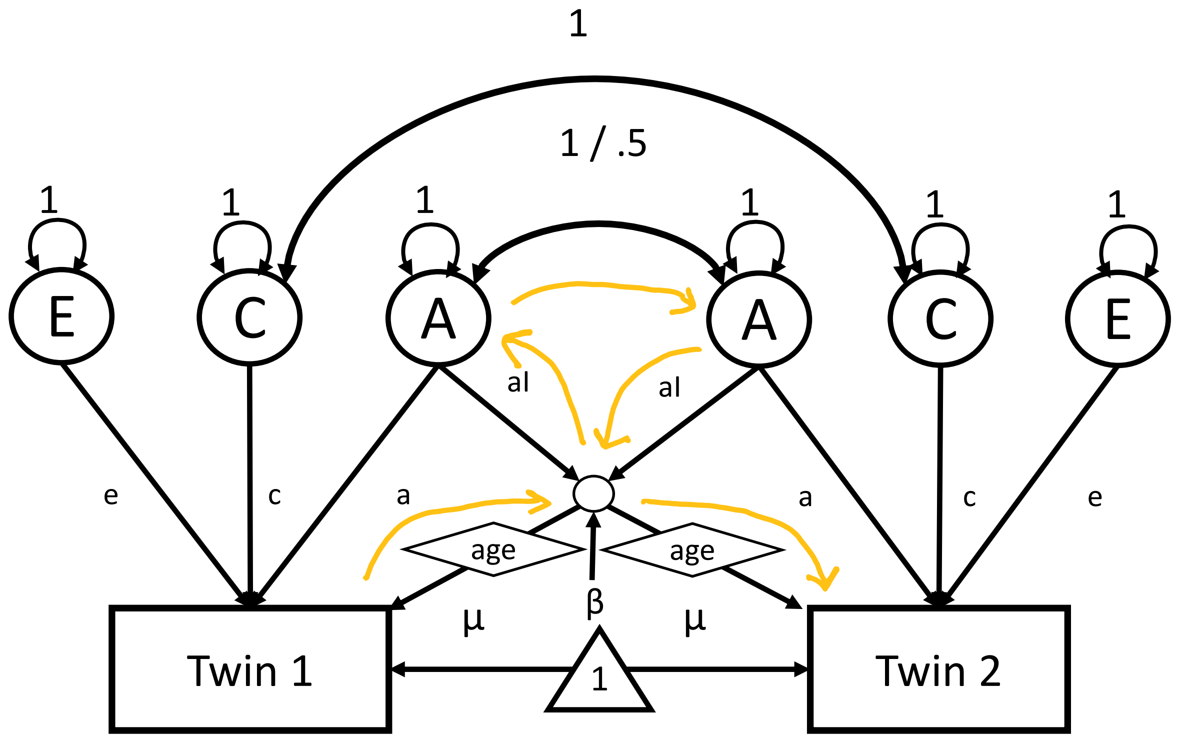


$$rz*(a+al*age_i)\,(a+al*age_i)$$
$$=$$
$$rz*a*a +$$
$$rz*a*al*age +$$
$$rz*al*age*a$$

rz = 1 for MZ twins and .5 for DZ twins

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

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



$rz*(a+al*age_i)\ (a+al*age_i)$

$=$

$rz*a*a\ +$

$rz*a*al*age\ +$

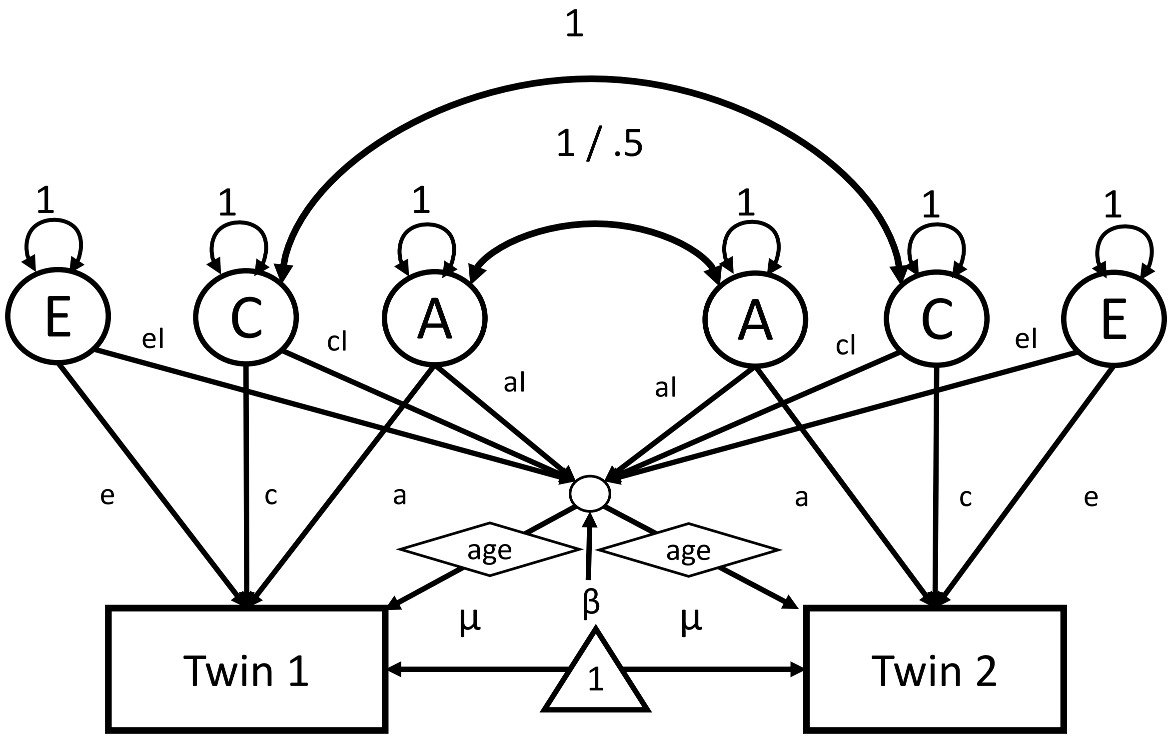
$rz*al*age*a\ +$

$rz*al*age*al*age$

r_Z = 1 for MZ twins and .5 for DZ twins

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

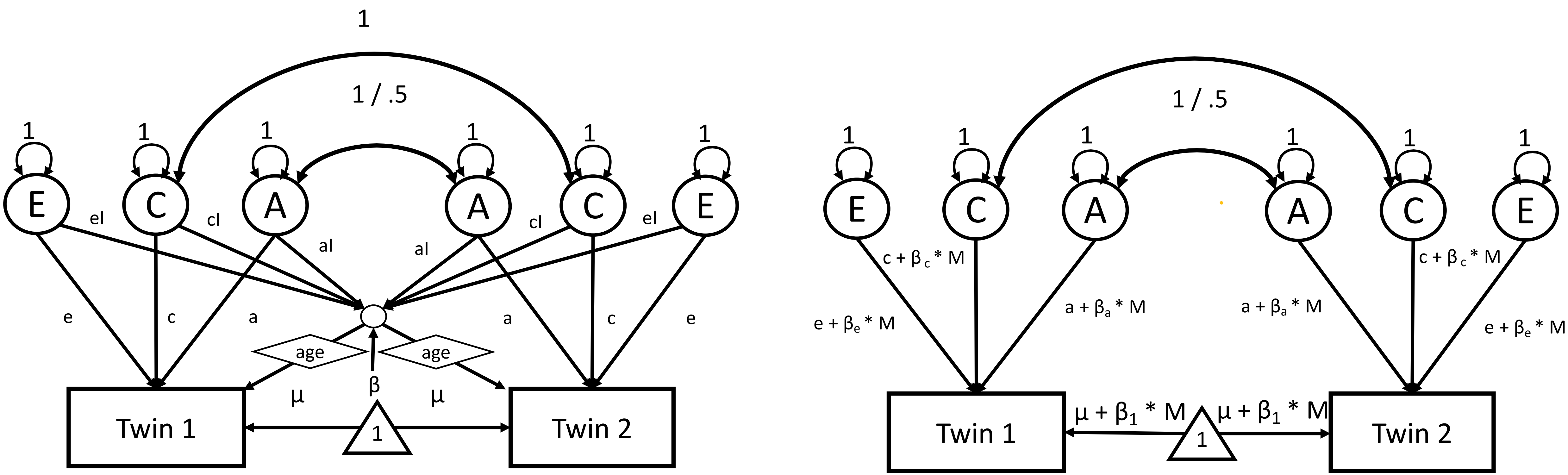
	Twin 1	Twin 2
Means	$\mu + \beta(age_i)$	$\mu + \beta(age_i)$



r_Z = 1 for MZ twins and .5 for DZ twins

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

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



age = M

$\beta = \beta_1$

al = β_a

cl = β_c

el = β_e

G x E: moderation model - extensions

- The moderator can be the same or different between twins of a pair
- The moderator can also interact with C and E
- Moderators may also have non-linear properties
- May show quantitative interactions, but can also test qualitative differences (only if the moderator differs between twin member)
- A moderator could be correlated with genetic effects on the trait, which allows to test for GxE in the presence of rGE (also in bivariate setting)

References to read:

Purcell S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin research : the official journal of the International Society for Twin Studies*, *5*(6), 554–571.

van der Sluis, S., Posthuma, D. & Dolan, C.V. A Note on False Positives and Power in G × E Modelling of Twin Data. *Behav Genet* **42**, 170–186 (2012).

Practical

- Sex limitation model: 5 group ADE (multigroup approach for binary variable)
- Moderation model, age-moderated ADE estimates (for continuous variable)
- In your Rstudio copy my files into your own directory
  - `system("cp -R /faculty/laura/2024/Day_2/practical_sex_limitation_GxE/* ./ ")`