Heterogeneity

Sarah Medland

& Nathan Gillespie

Types of Heterogeneity

- Terminology depends on research question
  - Moderation, confounding, GxE
- Systematic differences
  - Measured or Manifest moderator/confounder
    - Discrete traits
    - Ordinal & Continuous traits (Thursday)
  - Unmeasured or latent moderator/confounder
    - Moderation and GxE

Heterogeneity Questions

- Univariate Analysis:
  - What are the contributions of additive genetic, dominance/shared environmental and unique environmental factors to the variance?
- Heterogeneity:
  - Are the contributions of genetic and environmental factors equal for different groups,
  - sex, race, ethnicity, SES, environmental exposure, etc.?

The language of heterogeneity

- Are these differences due to differences in the magnitude of the effects (quantitative)?
  - e.g. Is the contribution of genetic/environmental factors greater/smaller in males than in females?
- Are the differences due to differences in the source/nature of the effects (qualitative)?
  - e.g. Are there different genetic/environmental factors influencing the trait in males and females?

The language of heterogeneity

- Sex differences = Sex limitation

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

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 Goëttingue, de l'Académie Royale de Médecine de Madrid, etc.

1861

THE

BRITISH AND FOREIGN

MEDICO-CHIRURGICAL

REVIEW

OR

QUARTERLY JOURNAL

OF

PRACTICAL MEDICINE AND SURGERY.

VOL. XXVII.

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 appreciation of colours. Several of the individuals comprised in it are utterly incapable of distinguishing one tune from another."†

* Cyclopædia of Anatomy and Physiology, art. "Vision," p. 1453.
† American Journal of Medical Science, vol. xxxv. p. 847. 1845.

The language of heterogeneity

Quantitative

- differences in the magnitude of the effects

Models

- Scalar
- Non-scalar with OS twins

Qualitative

- differences in the source/nature of the effects

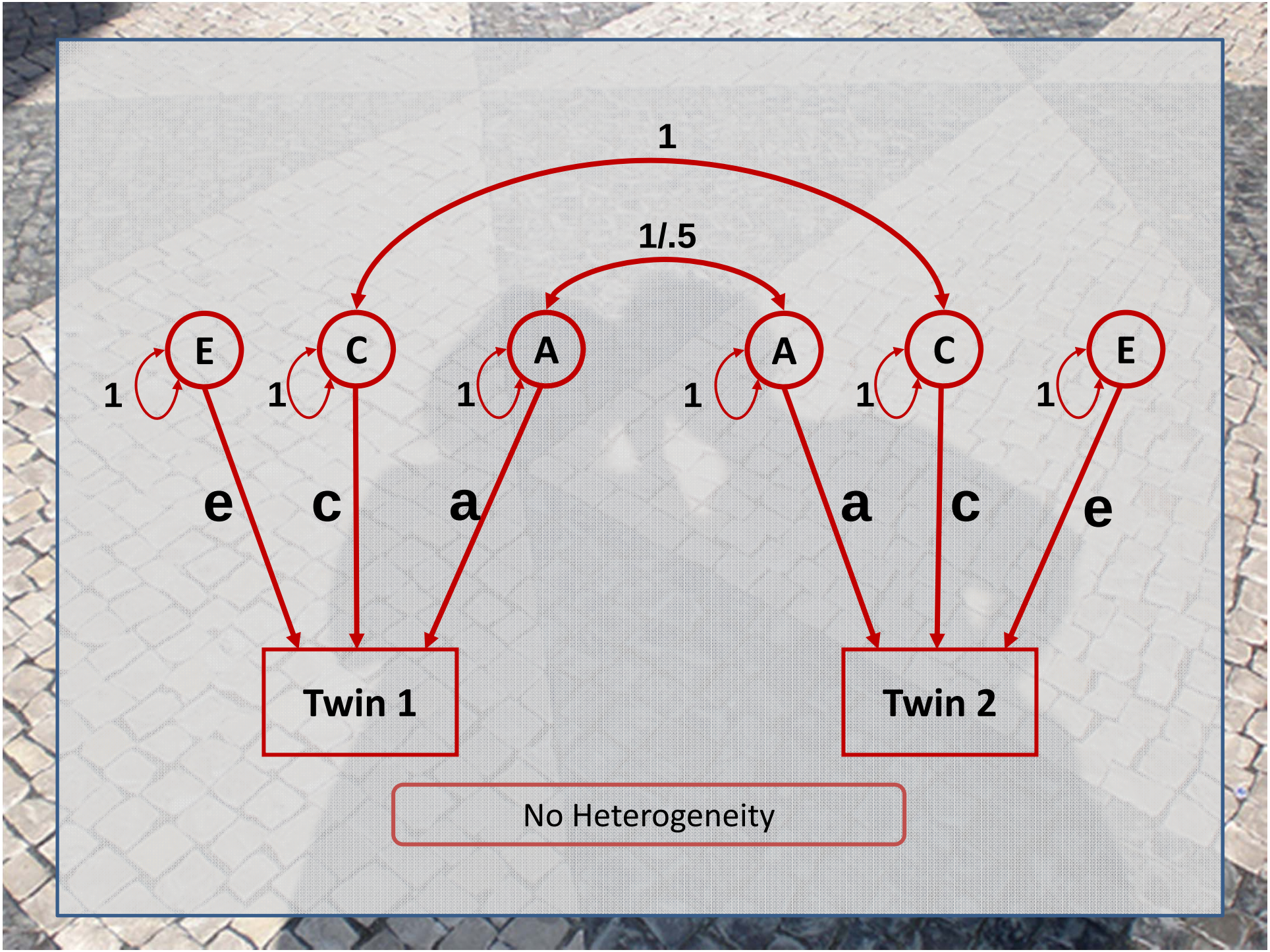
Models

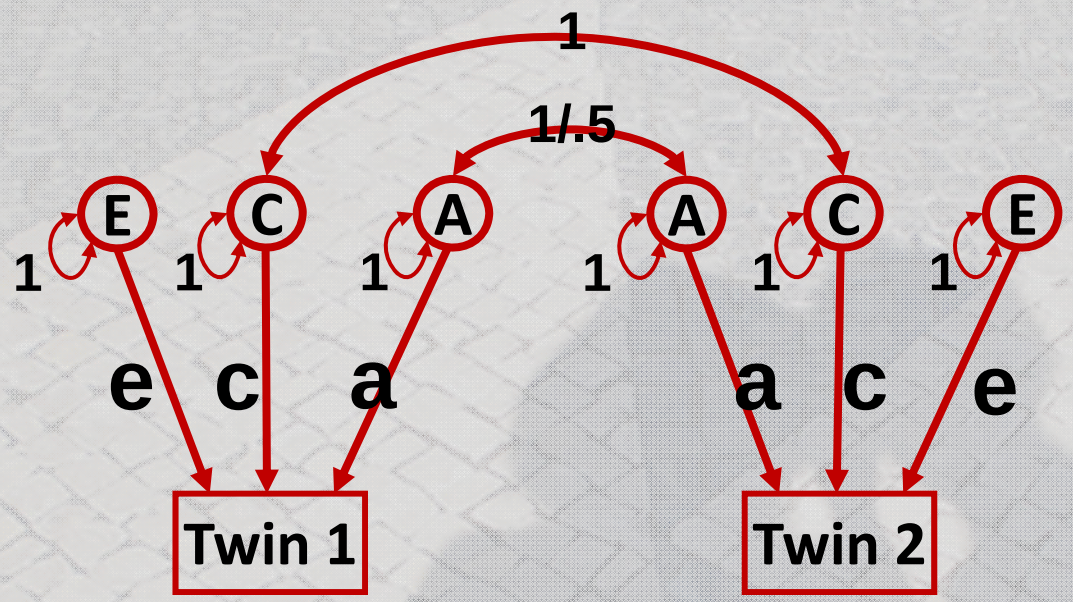
- Non-scalar without OS twins
- General Non-scalar

The language of heterogeneity

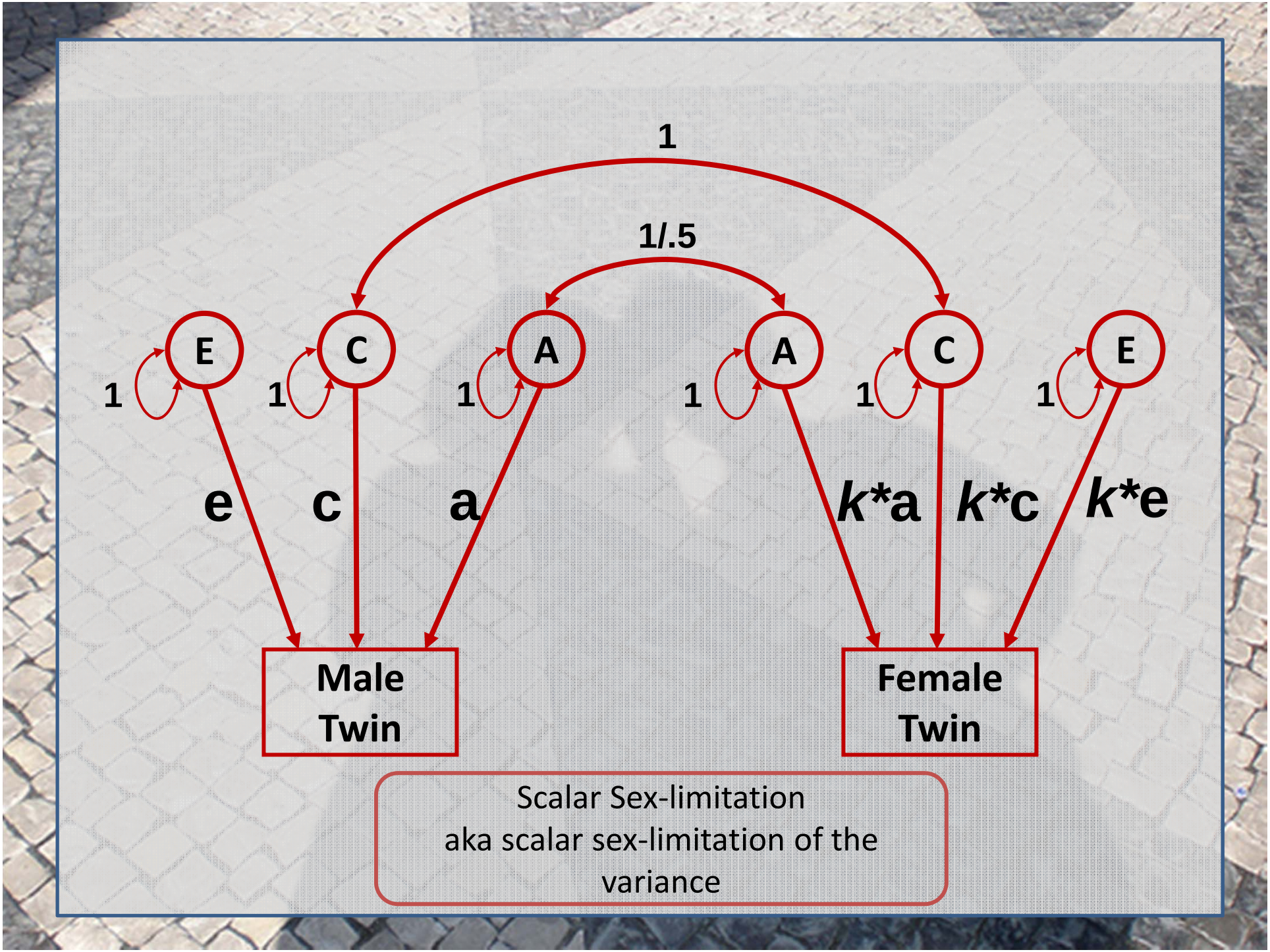
- Scalar limitation (Quantitative)
  - % of variance due to A,C,E are the same between groups
  - The total variance is not ie:
    - $var_{Female} = k*var_{Male}$
    - $A_{Female} = k*A_{Male}$
    - $C_{Female} = k*C_{Male}$
    - $E_{Female} = k*E_{Male}$

k here is the *scalar*


1
1/.5
E
C
A
A
C
E
1
1
1
1
1
1
e
c
a
a
c
e
Twin 1
Twin 2
No Heterogeneity



MZ		DZ	
$a^2+c^2+e^2$	a^2+c^2	$a^2+c^2+e^2$	$.5a^2+c^2$
a^2+c^2	$a^2+c^2+e^2$	$.5a^2+c^2$	$a^2+c^2+e^2$


1
1/.5
E
C
A
A
C
E
1
1
1
1
1
1
e
c
a
k*a
k*c
k*e
Male Twin
Female Twin
Scalar Sex-limitation
aka scalar sex-limitation of the variance

The language of heterogeneity

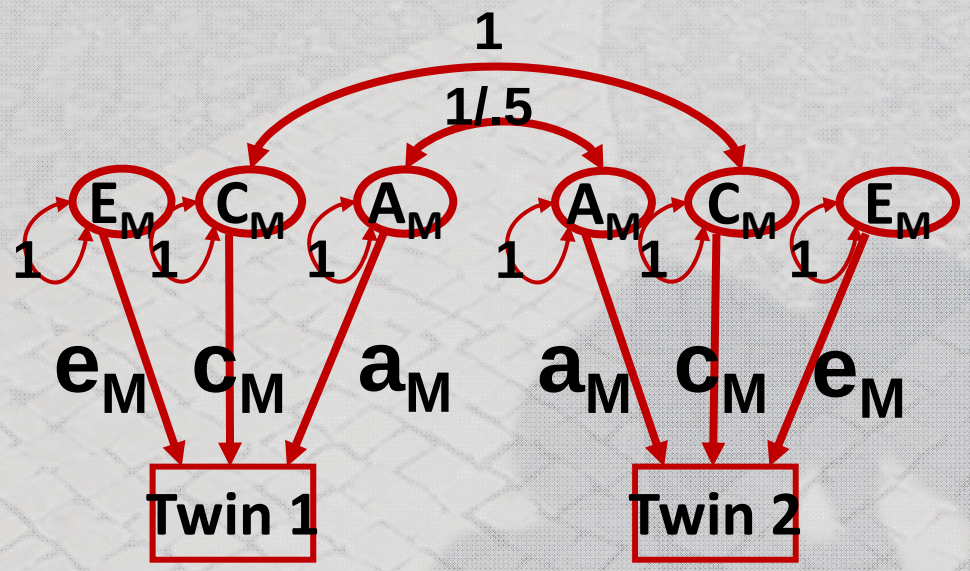
- Non-Scalar limitation
  - Without opposite sex twin pairs (Qualitative)
    - $var_{Female} \neq var_{Male}$
    - $A_{Female} \neq A_{Male}$
    - $C_{Female} \neq C_{Male}$
    - $E_{Female} \neq E_{Male}$

The language of heterogeneity

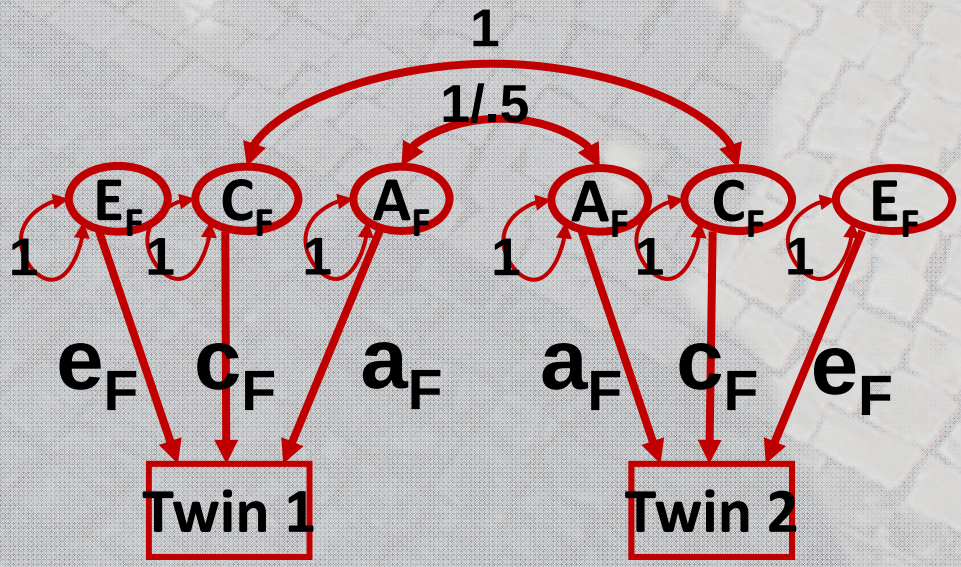
- Non-Scalar limitation
  - Without opposite sex twin pairs (Qualitative)
- Male Parameters
  - $means_M$
  - $A_M\ C_M$ *and* E_M
- Female Parameters
  - $mean_F$
  - $A_F\ C_F$ *and* E_F

Parameters are estimated separately

Male ACE model

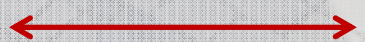


Female ACE model



The language of heterogeneity

- Non-Scalar limitation
  - With opposite sex twin pairs (Quantitative)
- Male Parameters
  - $means_M$
  - $A_M\ C_M$ *and* E_M



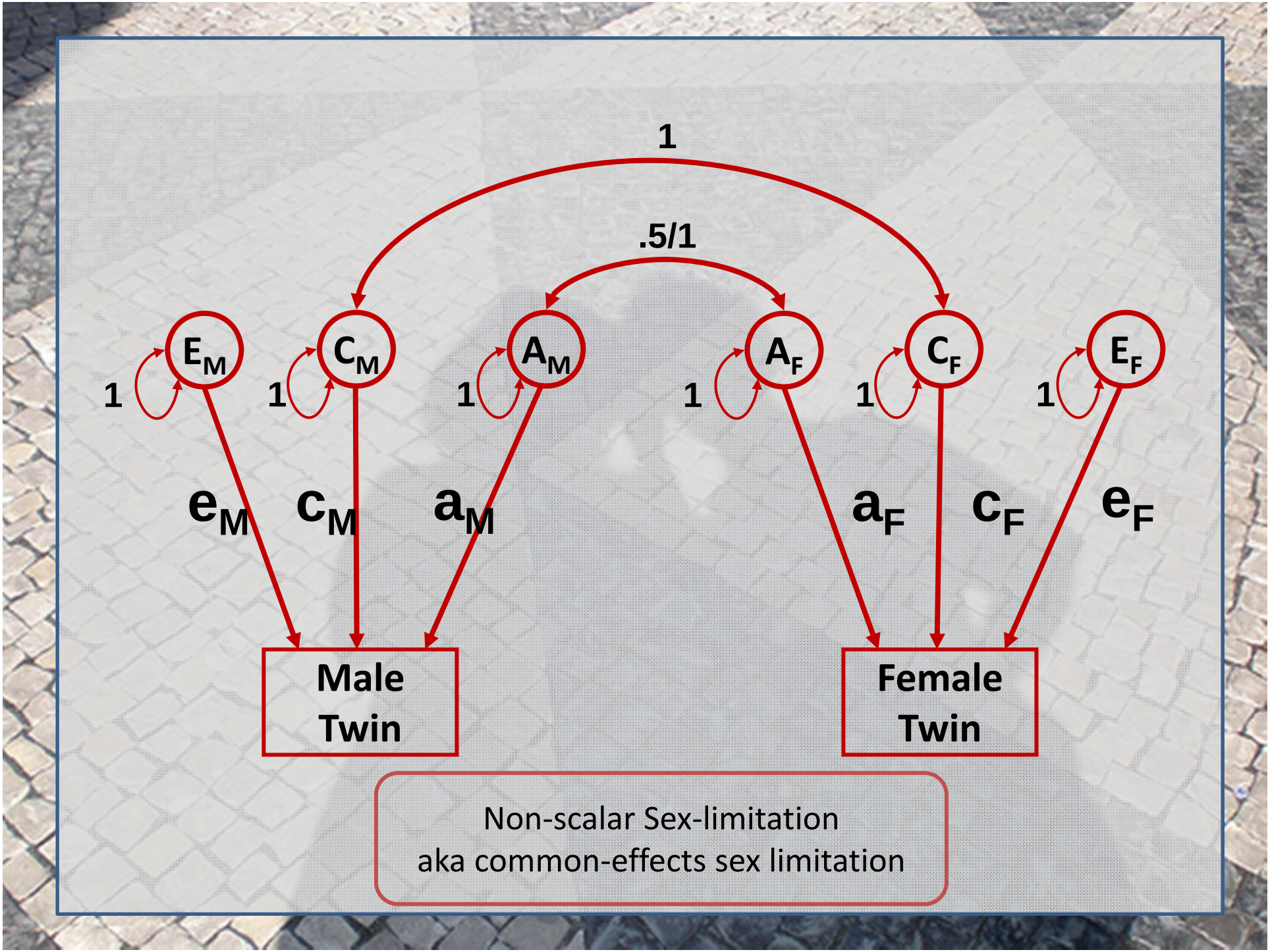
- Female Parameters
  - $mean_F$
  - $A_F\ C_F$ *and* E_F

Parameters are estimated jointly – linked via the opposite sex correlations

$r(A_{Female}, A_{male}) = .5$

$r(C_{Female} \neq C_{Male}) = 1$

$r(E_{Female} \neq E_{Male}) = 0$

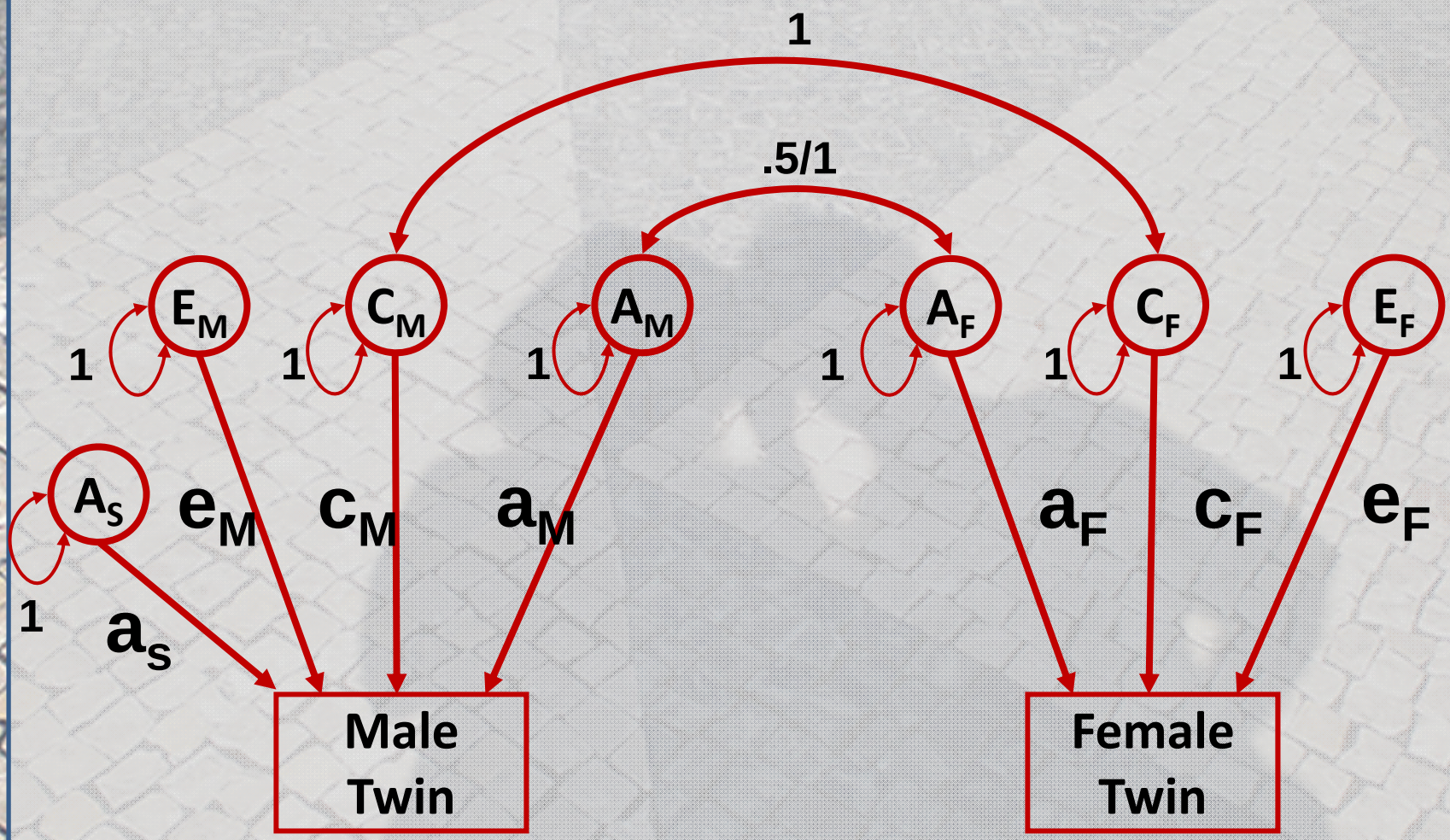

1
.5/1
E_M
C_M
A_M
A_F
C_F
E_F
1
1
1
1
1
1
e_M
c_M
a_M
a_F
c_F
e_F
Male Twin
Female Twin
Non-scalar Sex-limitation
aka common-effects sex limitation

The language of heterogeneity

- General Non-Scalar limitation
  - With opposite sex twin pairs (semi-Qualitative)
- Male Parameters
  - $means_M$
  - $A_M\ C_M\ E_M$ *and* $\boldsymbol{A}_{\mathbf{Specific}}$
  - Extra genetic/ environmental effects

- Female Parameters
  - $mean_F$
  - $A_F\ C_F$ *and* E_F

Parameters are estimated jointly – linked via the opposite sex correlations



General Non-scalar Sex-limitation
aka general sex limitation

The language of heterogeneity

- General Non-Scalar limitation via r_G
  - With opposite sex twin pairs (semi-Qualitative)

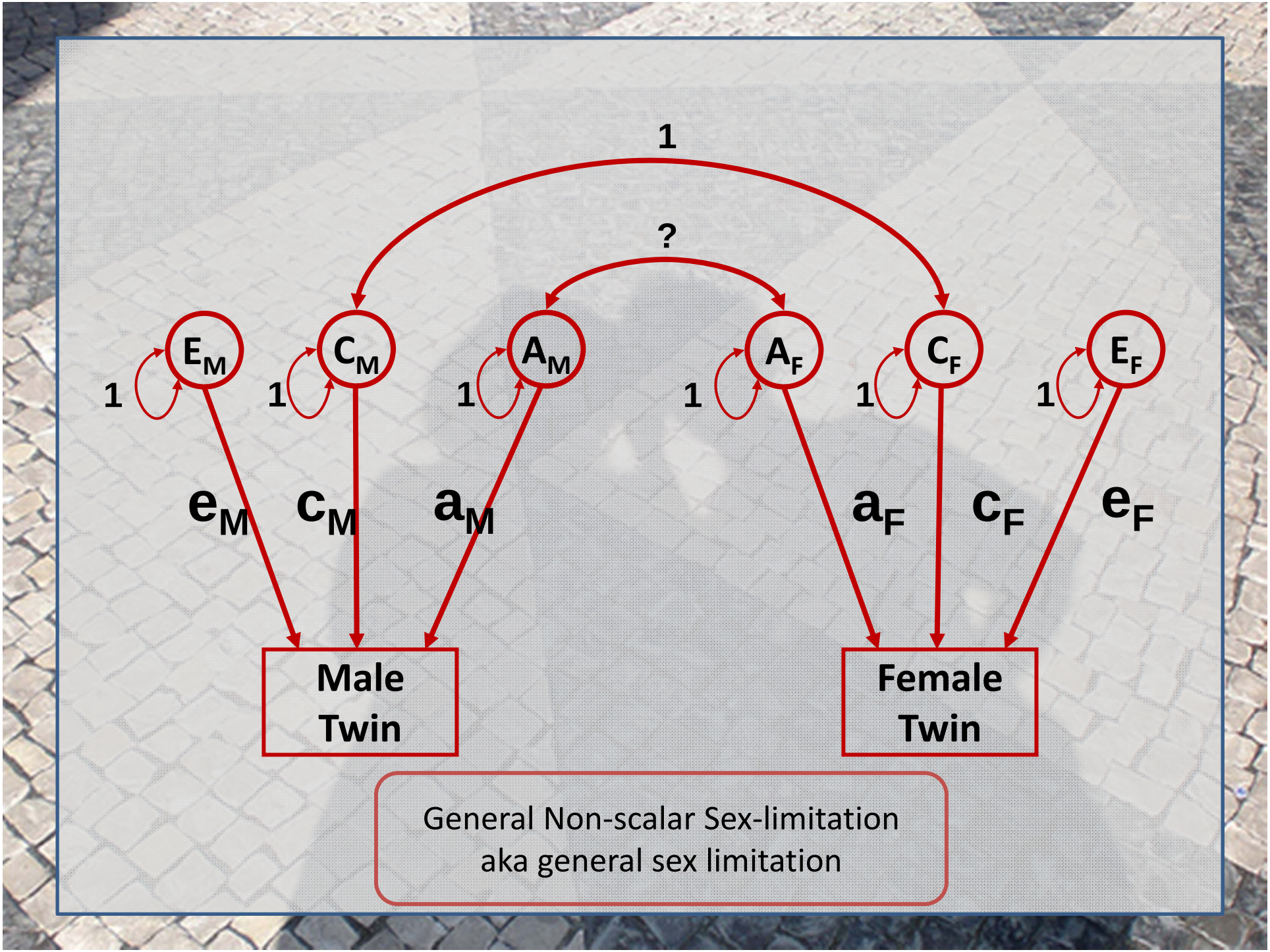
• Male Parameters		• Female Parameters
– $means_M$	⟷	– $mean_F$
– $A_M\ C_M\ E_M$		– $A_F\ C_F$ *and* E_F

Parameters are estimated jointly – linked via the opposite sex correlations

$$\boldsymbol{r(A_{Female}, A_{male}) = ?}\ \textbf{(estimated)}$$

$$r(C_{Female} \neq C_{Male}) = 1$$

$$r(E_{Female} \neq E_{Male}) = 0$$


1
?
E_M
C_M
A_M
A_F
C_F
E_F
1
1
1
1
1
1
e_M
c_M
a_M
a_F
c_F
e_F
Male Twin
Female Twin
General Non-scalar Sex-limitation
aka general sex limitation

How important is sex-limitation?

- Let have a look
  - Height data example using older twins
  - Zygosity coding
    - 6 & 8 are MZF & DZF
    - 7 & 9 are MZM & DZM
    - 10 is DZ FM
  - Scripts ACEf.R ACEm.R ACE.R
  - Left side of the room ACEm.R
  - Right side of the room ACE.R
  - Record the answers from the estACE* function

How important is sex-limitation?

- Female parameters

```
> estACEf
             [,1]         [,2]         [,3]         [,4]
[1,] 0.003385837 0.0001737119 0.000573992 0.004133541
[2,] 0.819112984 0.0420249614 0.138862055 1.000000000
```

- Male parameters

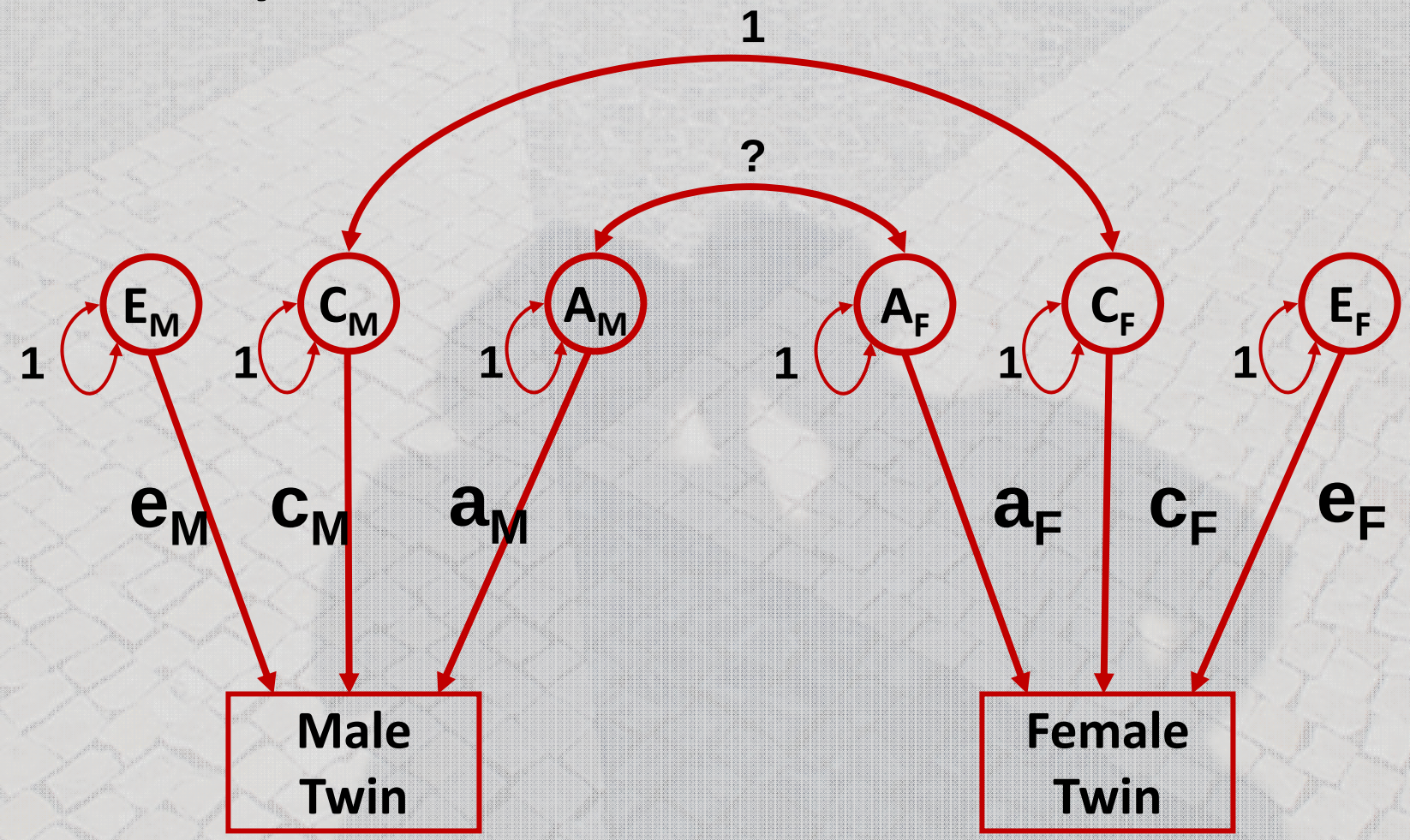
```
             [,1]         [,2]         [,3]         [,4]
[1,] 0.003183846 0.001179251 0.0004724089 0.004835506
[2,] 0.658430752 0.243873385 0.0976958621 1.000000000
```

- Combined parameters

```
             [,1]         [,2]         [,3]         [,4]
[1,] 0.009280147 3.922705e-17 0.0004994819 0.009779629
[2,] 0.948926289 4.011098e-15 0.0510737106 1.000000000
```

- Conclusions?

Lets try this model



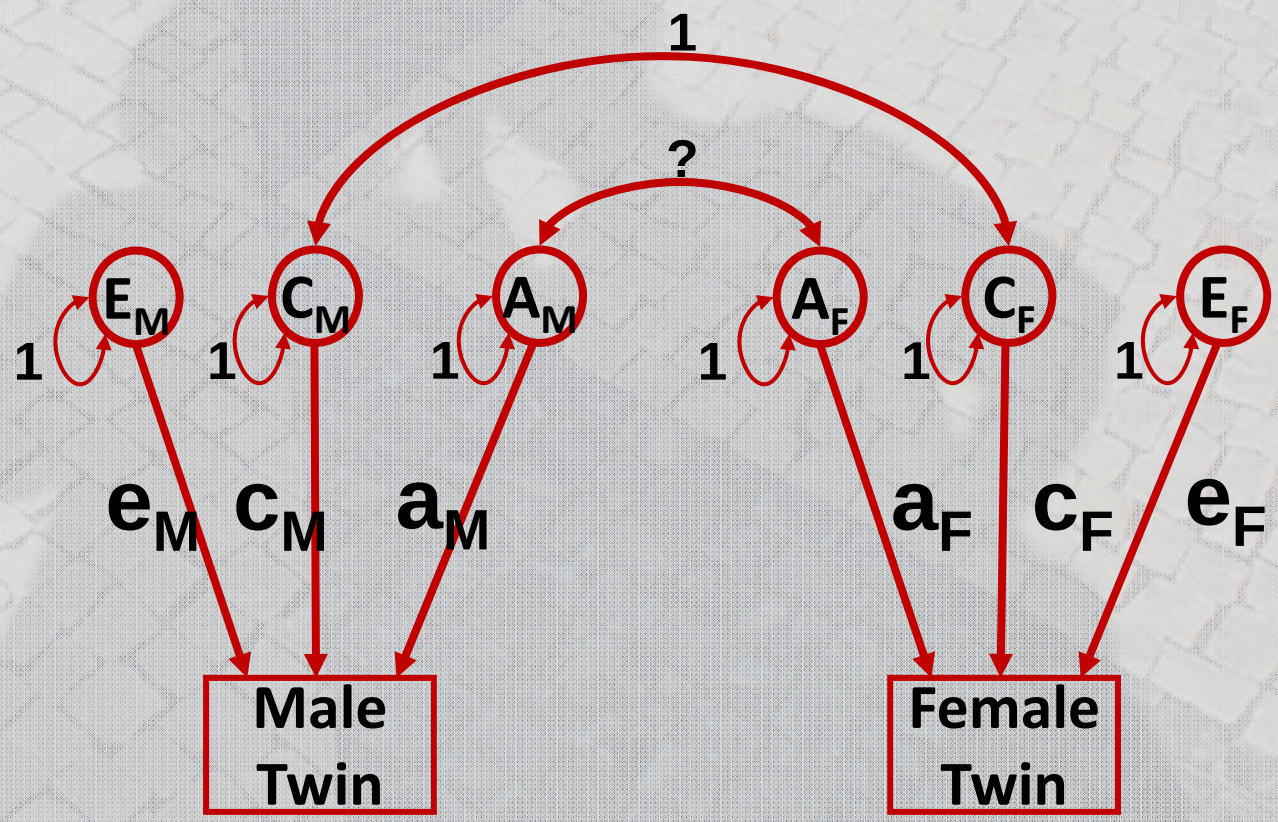
General Non-scalar Sex-limitation
aka general sex limitation

twinHet5AceCon.R

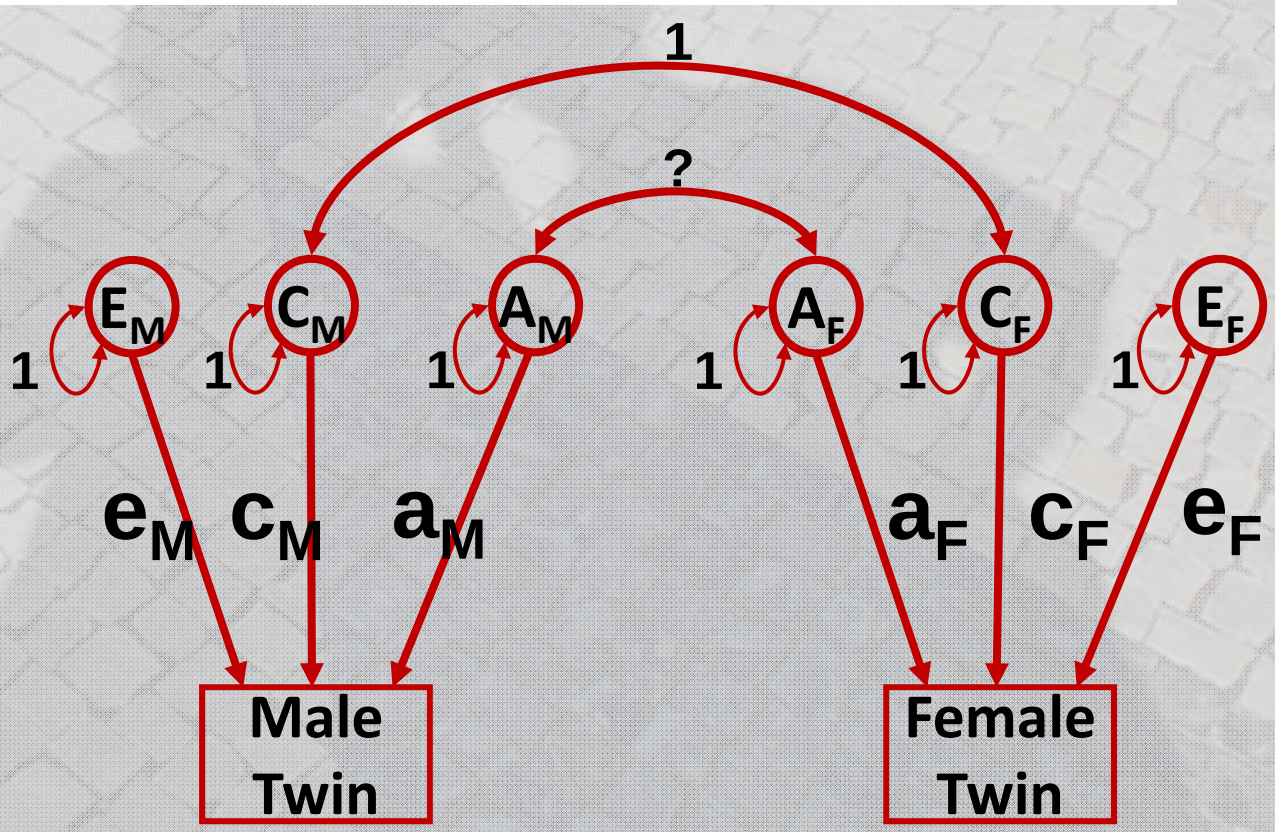
- Use data from all zygosity groups

```
# Select Data for Analysis
mzfData     <- subset(twinData, zyg==6, selVars)
dzfData     <- subset(twinData, zyg==7, selVars)
mzmData     <- subset(twinData, zyg==8, selVars)
dzmData     <- subset(twinData, zyg==9, selVars)
dzoData     <- subset(twinData, zyg==10, selVars) #fm
```

```
# ACE Model
# Matrices declared to store a, c, and e Path Coefficients
pathAf     <- mxMatrix( "Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, label="af11", name="af" )
pathCf     <- mxMatrix( "Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, label="cf11", name="cf" )
pathEf     <- mxMatrix( "Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, label="ef11", name="ef" )

pathAm     <- mxMatrix( "Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, label="am11", name="am" )
pathCm     <- mxMatrix( "Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, label="cm11", name="cm" )
pathEm     <- mxMatrix( "Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, label="em11", name="em" )

pathRg     <- mxMatrix( "Lower", nrow=1, ncol=1, free=TRUE, values=1, label="rg11", name="rg" )
```



```
# Matrices generated to hold A, C, and E computed Variance Components
covAf      <- mxAlgebra( af %*% t(af), name="Af" )
covCf      <- mxAlgebra( cf %*% t(cf), name="Cf" )
covEf      <- mxAlgebra( ef %*% t(ef), name="Ef" )

covAm      <- mxAlgebra( am %*% t(am), name="Am" )
covCm      <- mxAlgebra( cm %*% t(cm), name="Cm" )
covEm      <- mxAlgebra( em %*% t(em), name="Em" )

# Algebra to compute total variances and standard deviations (diagonal only)
covPf      <- mxAlgebra( Af+Cf+Ef, name="Vf" )
covPm      <- mxAlgebra( Am+Cm+Em, name="Vm" )
```



Means

```
# Algebra for expected Mean and Variance/Covariance Matrices in MZ & DZ twins
meanGf    <- mxMatrix( "Full", nrow=1, ncol=ntv, free=TRUE,
                       values= 20, label="meanf", name="expMeanGf" )
meanGm    <- mxMatrix( "Full", nrow=1, ncol=ntv, free=TRUE,
                       values= 20, label="meanm", name="expMeanGm" )
meanGfm   <- mxMatrix( "Full", nrow=1, ncol=ntv, free=TRUE,
                       values= 20, label=c("meanf","meanm"), name="expMeanGfm" )
```

- Have a think about this as we go through
  - is this the best way to set this up?

Covariances

```
covMZf    <- mxAlgebra( rbind( cbind(Af+Cf+Ef , Af+Cf),
                               cbind(Af+Cf    , Af+Cf+Ef)), name="expCovMZf" )

covDZf    <- mxAlgebra(  rbind( cbind(Af+Cf+Ef    , 0.5%x%Af+Cf),
                               cbind(0.5%x%Af+Cf , Af+Cf+Ef)), name="expCovDZf" )

covMZm    <- mxAlgebra(  rbind( cbind(Am+Cm+Em , Am+Cm),
                               cbind(Am+Cm    , Am+Cm+Em)), name="expCovMZm" )

covDZm    <- mxAlgebra(  rbind( cbind(Am+Cm+Em    , 0.5%x%Am+Cm),
                               cbind(0.5%x%Am+Cm , Am+Cm+Em)), name="expCovDZm" )

covDZo    <- mxAlgebra(  rbind( cbind(Af+Cf+Ef     , 0.5%*%rg%x%(af%*%t(am))+cf%*%t(cm)),
                               cbind(0.5%*%rg%x%(am%*%t(af))+cm%*%t(cf) , Am+Cm+Em)),  name="expCovDZo" )
```

```
# Data objects for Multiple Groups
dataMZf   <- mxData( observed=mzfData, type="raw" )
dataDZf   <- mxData( observed=dzfData, type="raw" )
dataMZm   <- mxData( observed=mzmData, type="raw" )
dataDZm   <- mxData( observed=dzmData, type="raw" )
dataDZo   <- mxData( observed=dzoData, type="raw" )

# Objective objects for Multiple Groups
objMZf    <- mxFIMLObjective( covariance="expCovMZf", means="expMeanGf", dimnames=selVars )
objDZf    <- mxFIMLObjective( covariance="expCovDZf", means="expMeanGf", dimnames=selVars )
objMZm    <- mxFIMLObjective( covariance="expCovMZm", means="expMeanGm", dimnames=selVars )
objDZm    <- mxFIMLObjective( covariance="expCovDZm", means="expMeanGm", dimnames=selVars )
objDZo    <- mxFIMLObjective( covariance="expCovDZo", means="expMeanGfm", dimnames=selVars )
```

```r
# Combine Groups
parsZf    <- list( pathAf, pathCf, pathEf, covAf, covCf, covEf, covPf, estVarsZf )
parsZm    <- list( pathAm, pathCm, pathEm, covAm, covCm, covEm, covPm, estVarsZm )

modelMZf  <- mxModel( parsZf, meanGf, covMZf, dataMZf, objMZf, name="MZf" )
modelDZf  <- mxModel( parsZf, meanGf, covDZf, dataDZf, objDZf, name="DZf" )
modelMZm  <- mxModel( parsZm, meanGm, covMZm, dataMZm, objMZm, name="MZm" )
modelDZm  <- mxModel( parsZm, meanGm, covDZm, dataDZm, objDZm, name="DZm" )
modelDZo  <- mxModel( parsZf, pathRg, parsZm, meanGfm, covDZo, dataDZo, objDZo, name="DZo" )
```

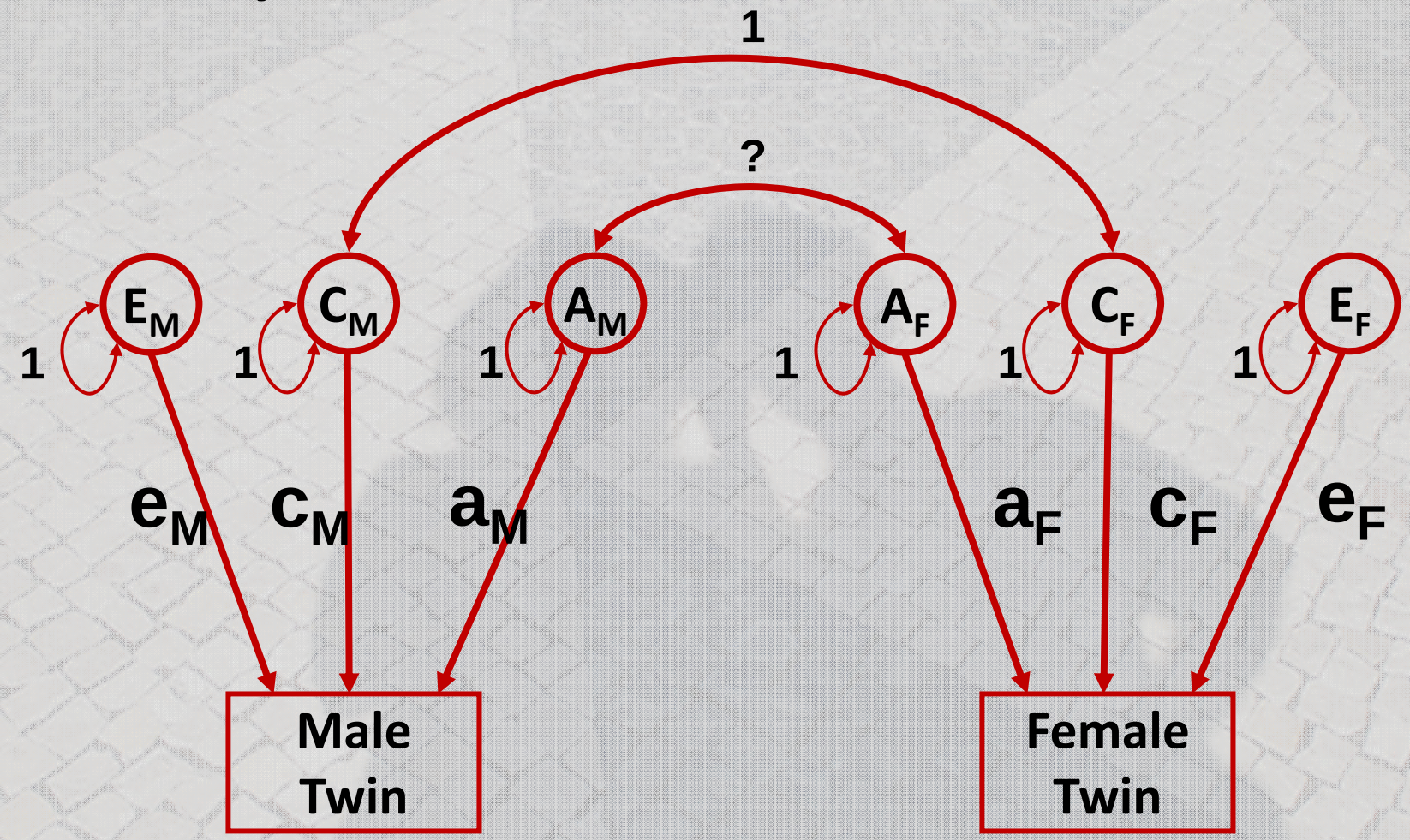
```r
minus2ll  <- mxAlgebra( MZf.objective+ DZf.objective+ MZm.objective+ DZm.objective+ DZo.objective, name="m2LL" )
obj       <- mxAlgebraObjective( "m2LL" )
QualAceModel  <- mxModel( "QualACE", parsZf, parsZm, modelMZf, modelDZf, modelMZm, modelDZm, modelDZo, minus2ll, obj )
```

```r
# Run Qualitative Sex Differences ACE model
QualAceFit    <- mxRun(QualAceModel)
QualAceSumm   <- summary(QualAceFit)
QualAceSumm
round(QualAceFit@output$estimate,4)
round(cbind(QualAceFit$VarsZf@result,QualAceFit$VarsZm@result),4)
```

Run it

- What would we conclude?
- Do we believe it?
- Checking the alternate parameterisation…

Lets try this model



General Non-scalar Sex-limitation
aka general sex limitation

- af = -.06
- am = -.06
- rg = -.9

Dzr = -.06*(.5*-.9) .06

=.45

NATURE | LETTER

previous article next article

Hundreds of variants clustered in genomic loci and biological pathways affect human height

Hana Lango Allen, Karol Estrada, Guillaume Lettre, Sonja I. Berndt, Michael N. Weedon, Fernando Rivadeneira, Cristen J. Willer, Anne U. Jackson, Sailaja Vedantam, Soumya Raychaudhuri, Teresa Ferreira, Andrew R. Wood, Robert J. Weyant, Ayellet V. Segrè, Elizabeth K. Speliotes, Eleanor Wheeler, Nicole Soranzo, Ju-Hyun Park, Jian Yang, Daniel Gudbjartsson, Nancy L. Heard-Costa, Joshua C. Randall, Lu Qi, Albert Vernon Smith, Reedik Mägi + et al.

Affiliations | Contributions | Corresponding authors

Nature **467**, 832–838 (14 October 2010) | doi:10.1038/nature09410

Received 23 April 2010 | Accepted 28 July 2010 | Published online 29 September 2010

Means

```
# Algebra for expected Mean and Variance/Covariance Matrices in MZ & DZ twins
meanGf    <- mxMatrix( "Full", nrow=1, ncol=ntv, free=TRUE,
                       values= 20, label="meanf", name="expMeanGf" )
meanGm    <- mxMatrix( "Full", nrow=1, ncol=ntv, free=TRUE,
                       values= 20, label="meanm", name="expMeanGm" )
meanGfm   <- mxMatrix( "Full", nrow=1, ncol=ntv, free=TRUE,
                       values= 20, label=c("meanf","meanm"), name="expMeanGfm" )
```

- Add a correction using a regression model
- expectedMean = maleMean + β*sex

β is the female deviation from the male mean

Sex is coded 0/1