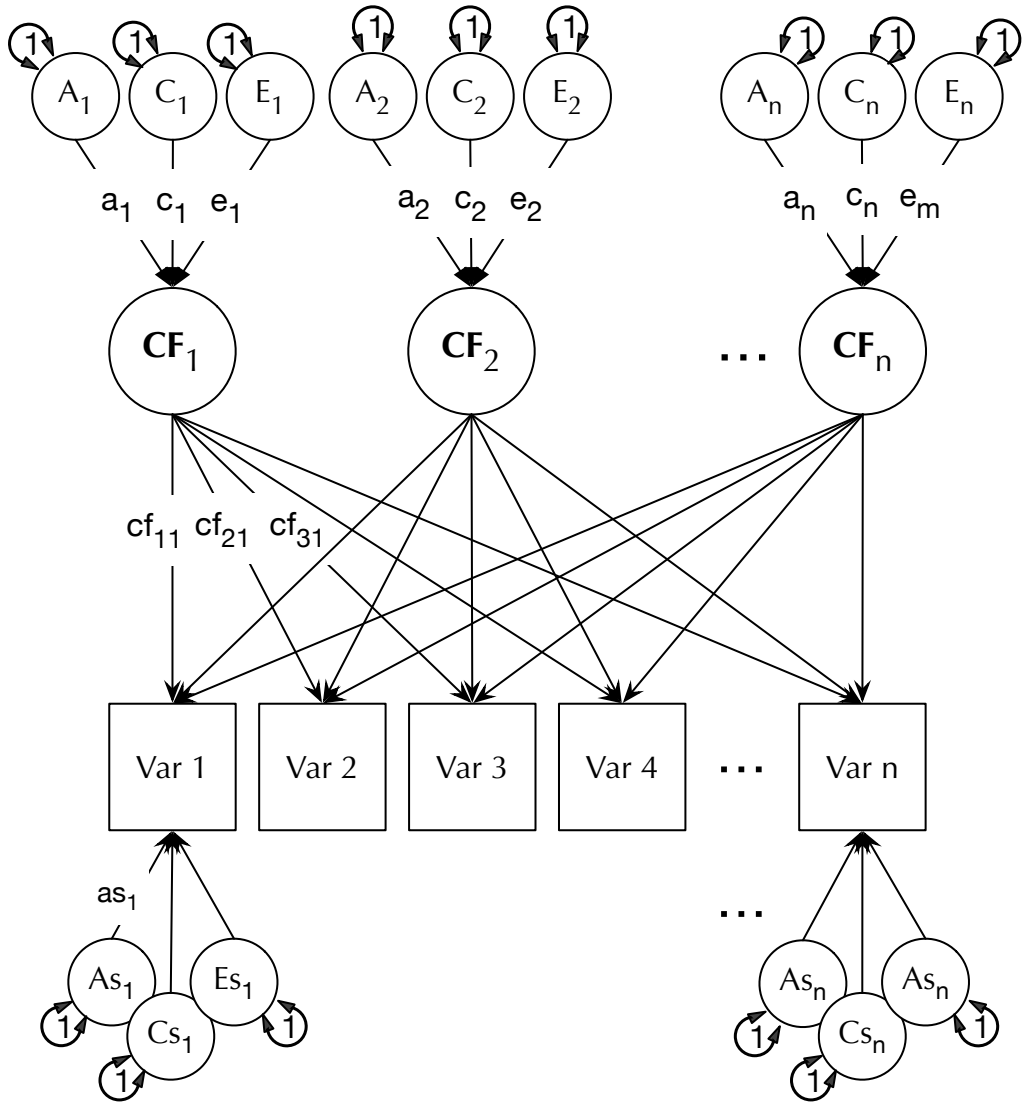
Multivariate Models

- ACE
  - umxACEv; umxACE
- Factor models:
  - Independent Pathway: umxIP
  - Common Pathway: umxCP



Baseline “model” vs. models that expose theory to risky tests of their predictions

Are my phenotypes influenced by the same genes?

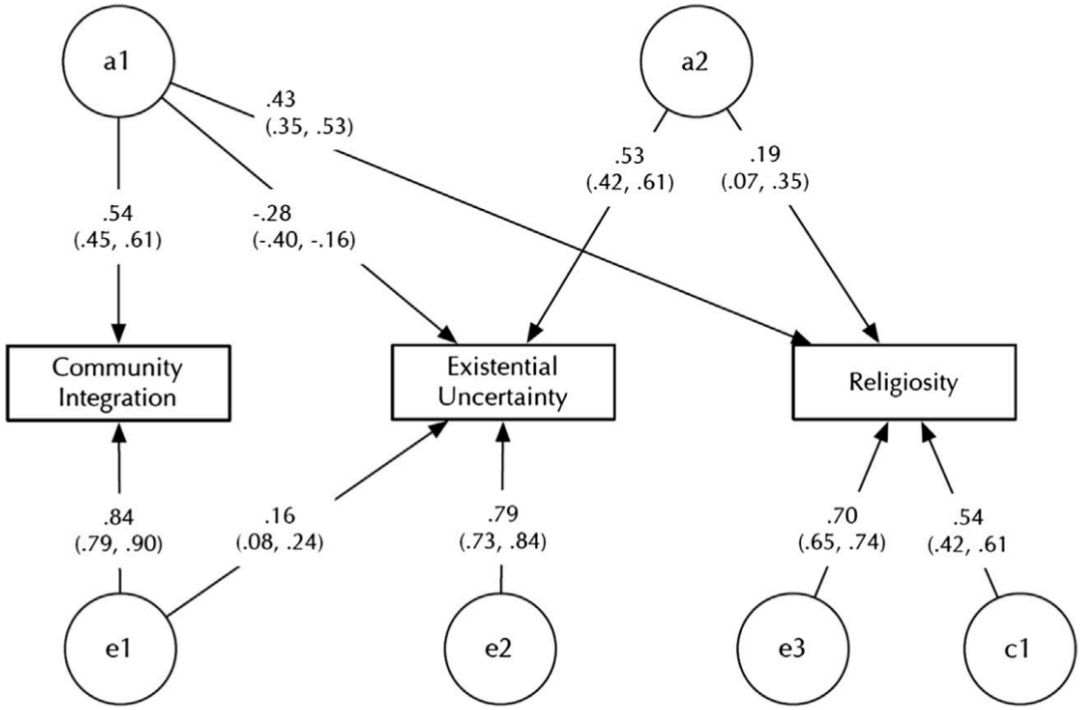
Is there more than one ”set” of gene effects?

Decompose depression into components?

Do A, C, E factors contribute in the same ways to the covariance between traits?

One trait or more?

Is community the same (genetically) as religiosity?



Journal of Research in Personality 47 (2013) 398–405

Contents lists available at SciVerse ScienceDirect

Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

Common genetic influences underpin religiosity, community integration, and existential uncertainty

Gary J. Lewis [a,b,*], Timothy C. Bates [b,*]

[a] Division of Psychology, School of Natural Sciences, University of Stirling, Stirling FK9 4LA, UK
[b] Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK

What is the human affiliation system(s)?

Psychological Science OnlineFirst, published on October 25, 2010 as doi:10.1177/0956797610387439

Research Report

Genetic Evidence for Multiple Biological Mechanisms Underlying In-Group Favoritism

Gary J. Lewis and Timothy C. Bates
University of Edinburgh

Psychological Science
XX(X) 1–6
© The Author(s) 2010
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797610387439
http://pss.sagepub.com

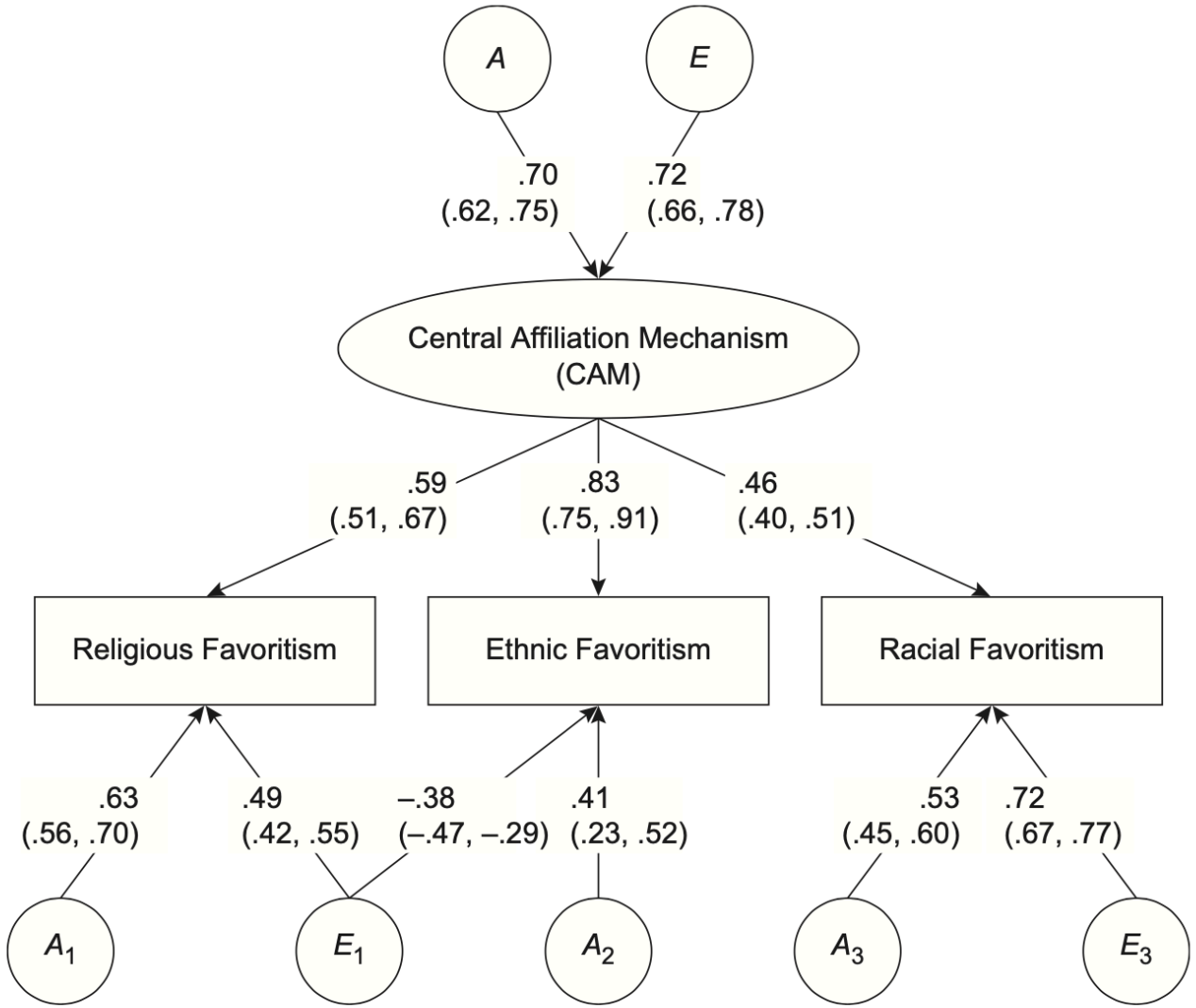


Fig. 3. Final model (Model 3d) of the genetic and environmental influences on in-group favoritism in three domains. Subscripts are used to differentiate influences on religious (1), ethnic (2), and racial (3) favoritism. All paths shown here are significant, and their coefficients are indicated. Variables with no connecting pathway were modeled as independent. The 95% confidence intervals are given in parentheses. *A* = genetic influences; *E* = unique-environment influences.

Do genes operate via personality domains? Or only via facet-facet associations?

Journal of Research in Personality 51 (2014) 9–17

Contents lists available at ScienceDirect

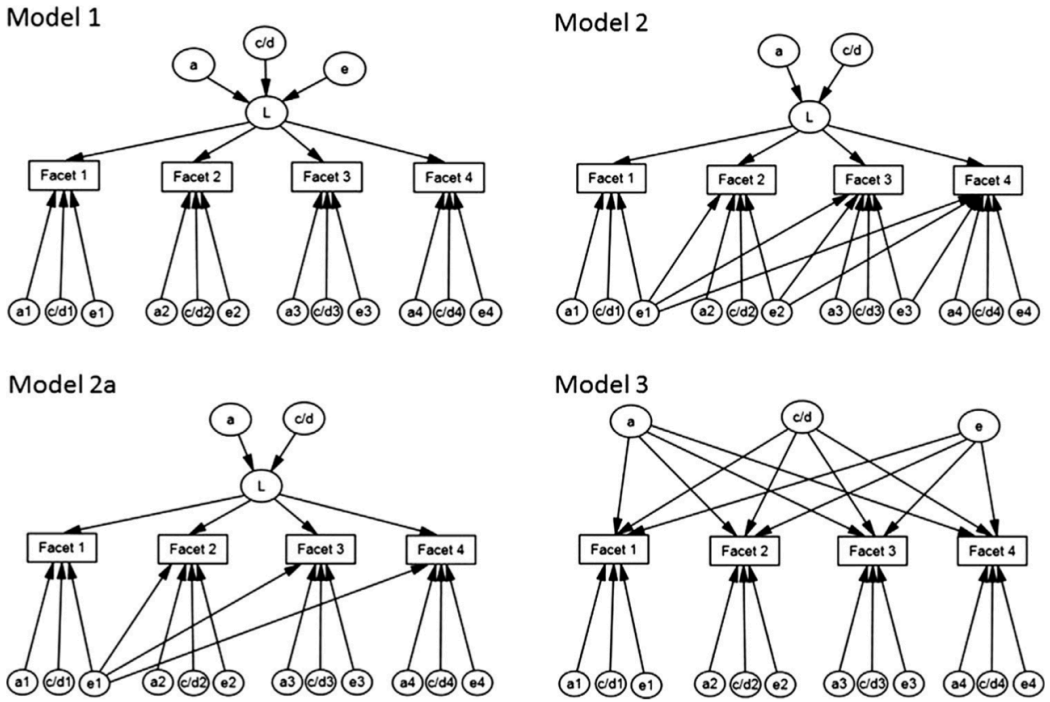
Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

How genes influence personality: Evidence from multi-facet twin analyses of the HEXACO dimensions

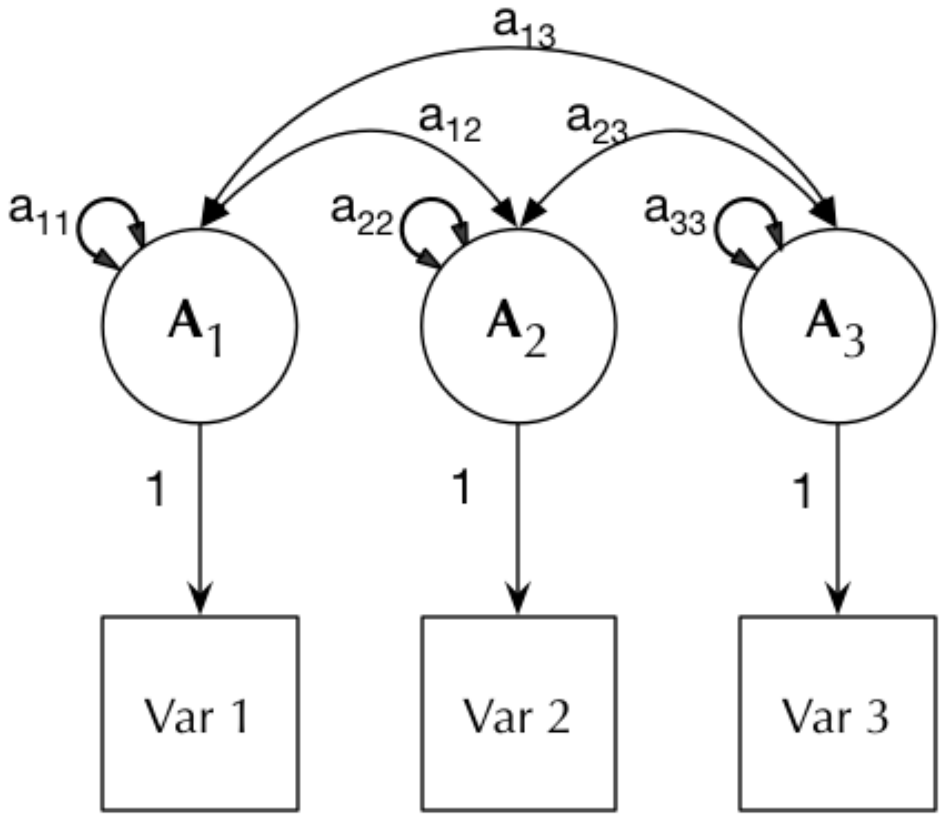
Gary J. Lewis [a,b,*], Timothy C. Bates [c]

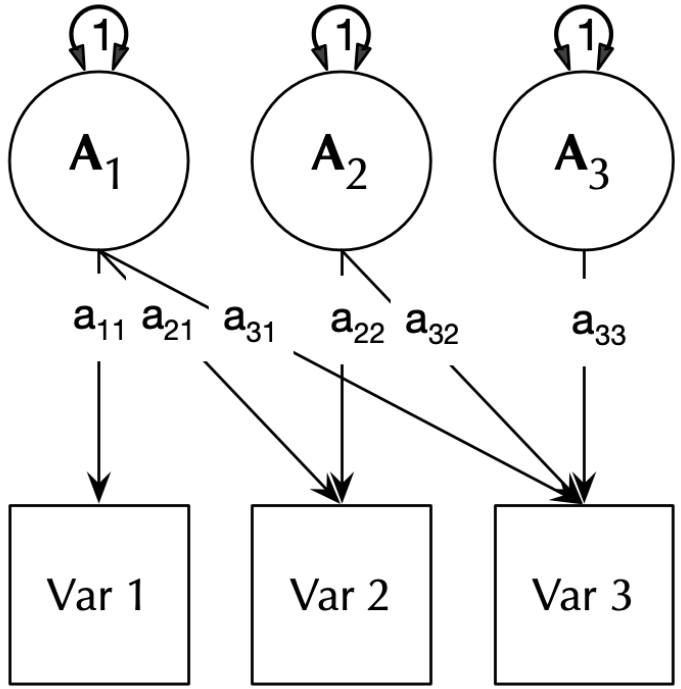
[a] Department of Psychology, University of York, Heslington, York YO10 5DD, UK
[b] Department of Twin Research and Genetic Epidemiology, King's College, St. Thomas Hospital, Westminster Bridge Road, London SE1 7EH, UK
[c] Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK



We've been introducing the ACEv model.

- And the ACE Cholesky model

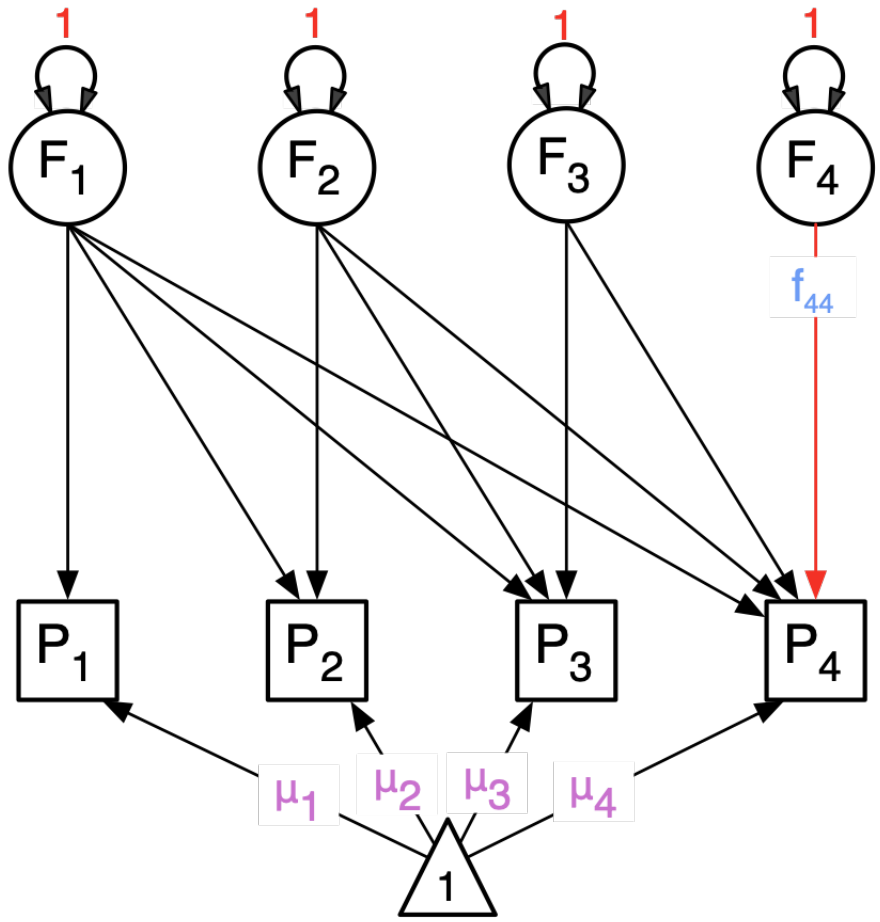




'a' matrix: 3x3 lower

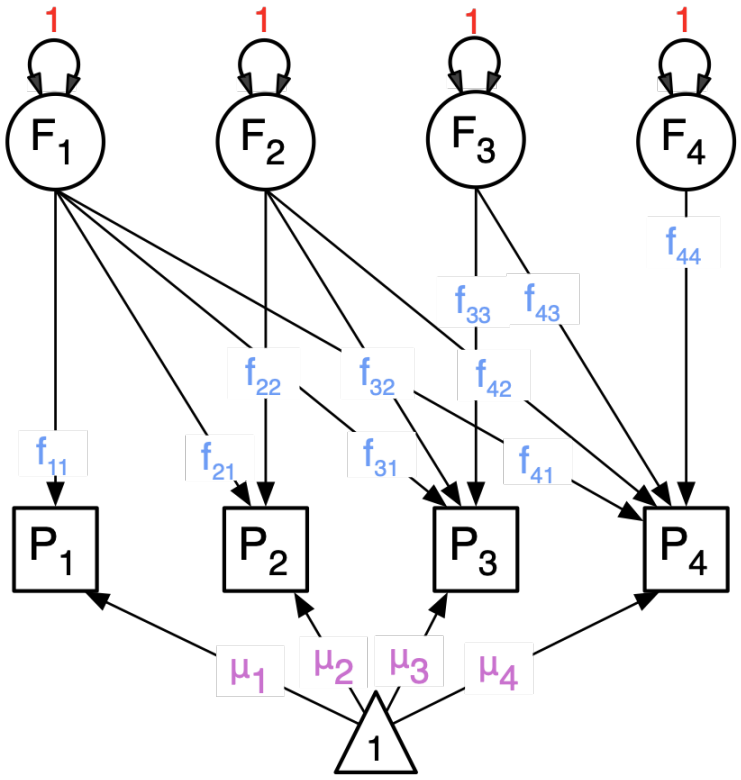
	1	2	3
1	a_r**1**c1		
2	a_r**2**c1	a_r2c2	
3	a_r**3**c1	a_r3c2	a_r3c3

The Cholesky decomposition: Factors load on all Variables but those previous



$$
\begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_4 \end{array}
\begin{array}{c}
\begin{array}{cccc} F_1 & F_2 & F_3 & F_4 \end{array} \\
\begin{bmatrix}
f_{11} & 0 & 0 & 0 \\
f_{21} & f_{22} & 0 & 0 \\
f_{31} & f_{32} & f_{33} & 0 \\
f_{41} & f_{42} & f_{43} & f_{44}
\end{bmatrix}
\end{array}
$$

Can we re-arrange this model? (Think factor analysis and factor rotation)



Published: January 1996

The Cholesky approach: A cautionary note

John C. Loehlin

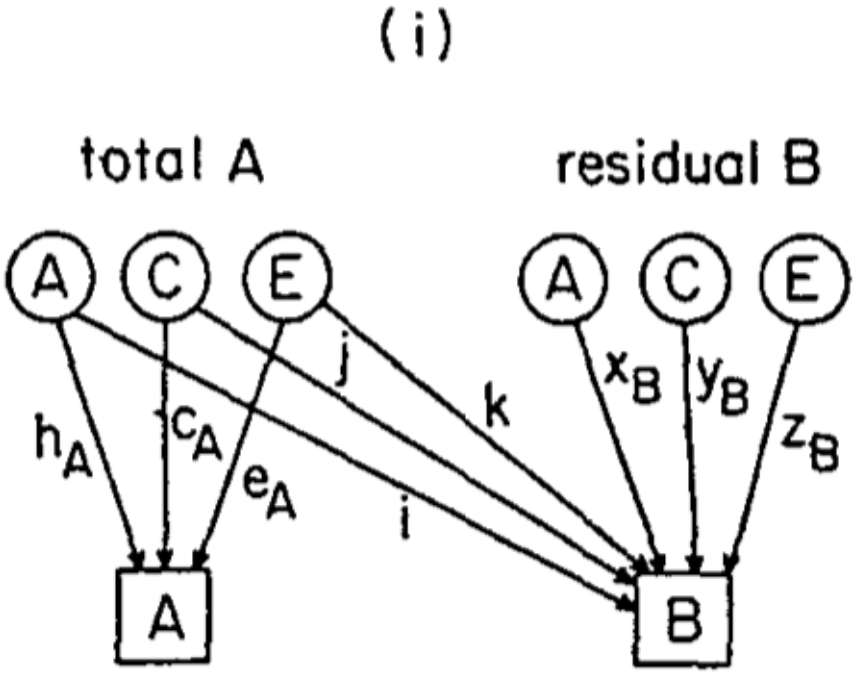
Behavior Genetics **26**, 65–69(1996) | Cite this article

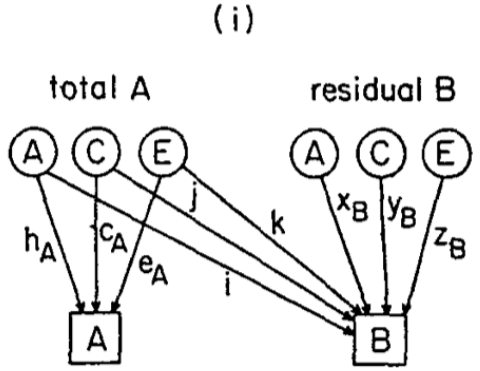
716 Accesses | **239** Citations | Metrics

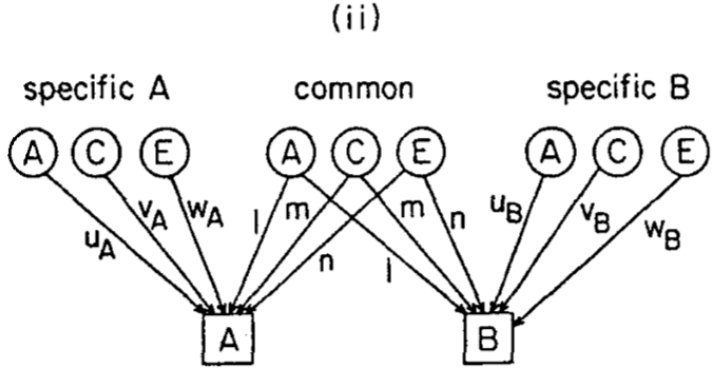
Abstract

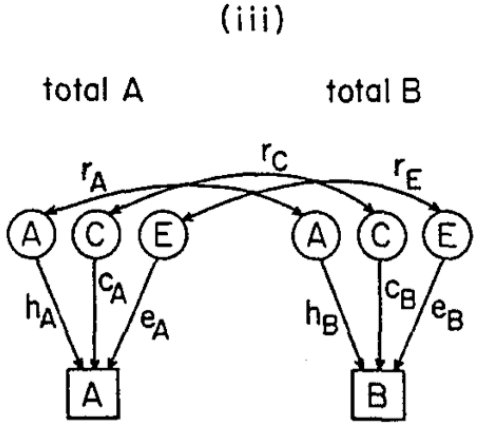
Attention is called to a common misinterpretation of a bivariate Cholesky analysis as if it were a common and specific factor analysis. It is suggested that an initial Cholesky behavior genetic analysis should often be transformed into a different form for interpretation. Formulas are provided for four transformations in the bivariate case.

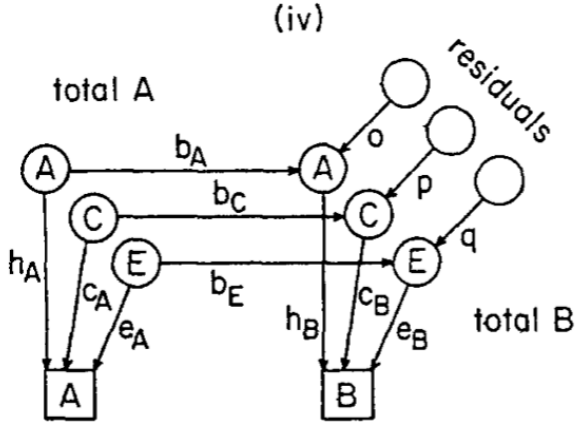
All equivalent

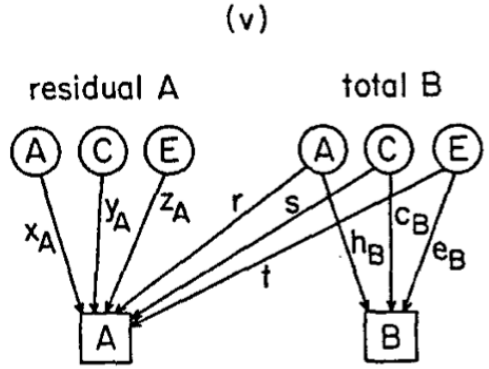




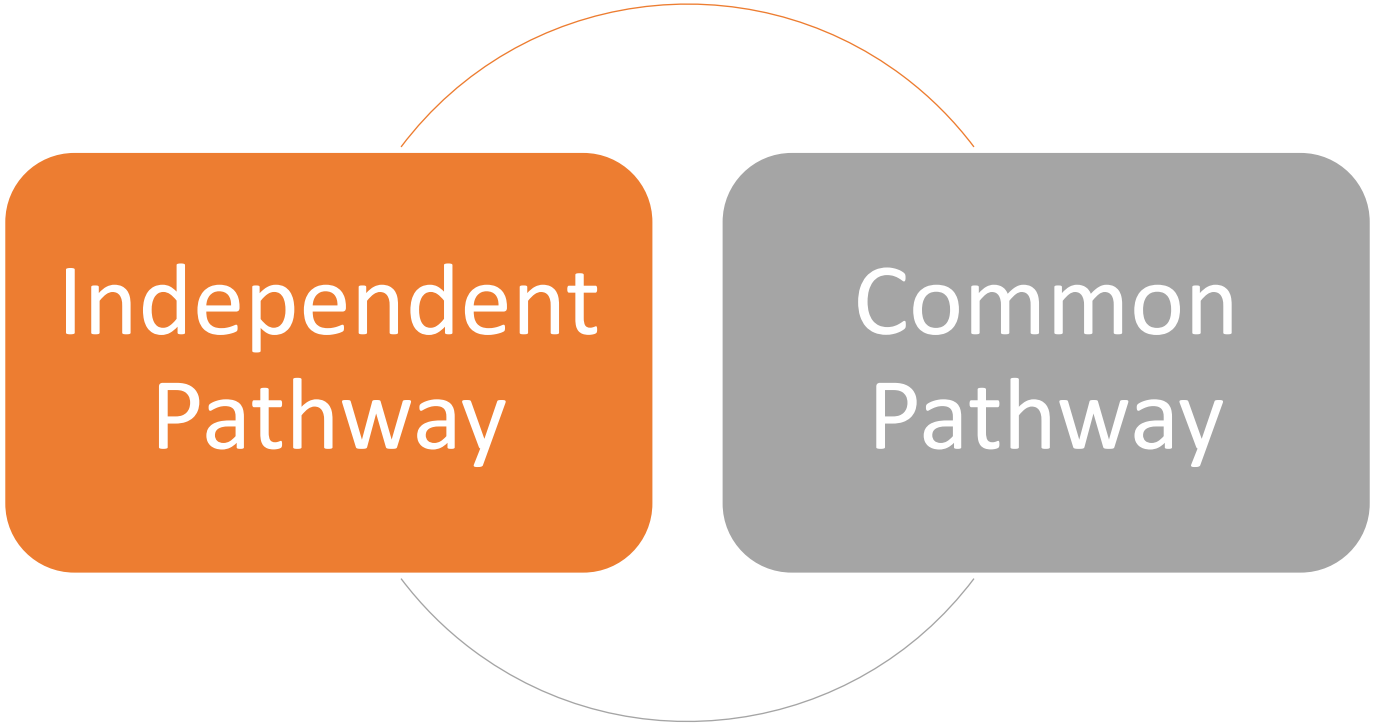




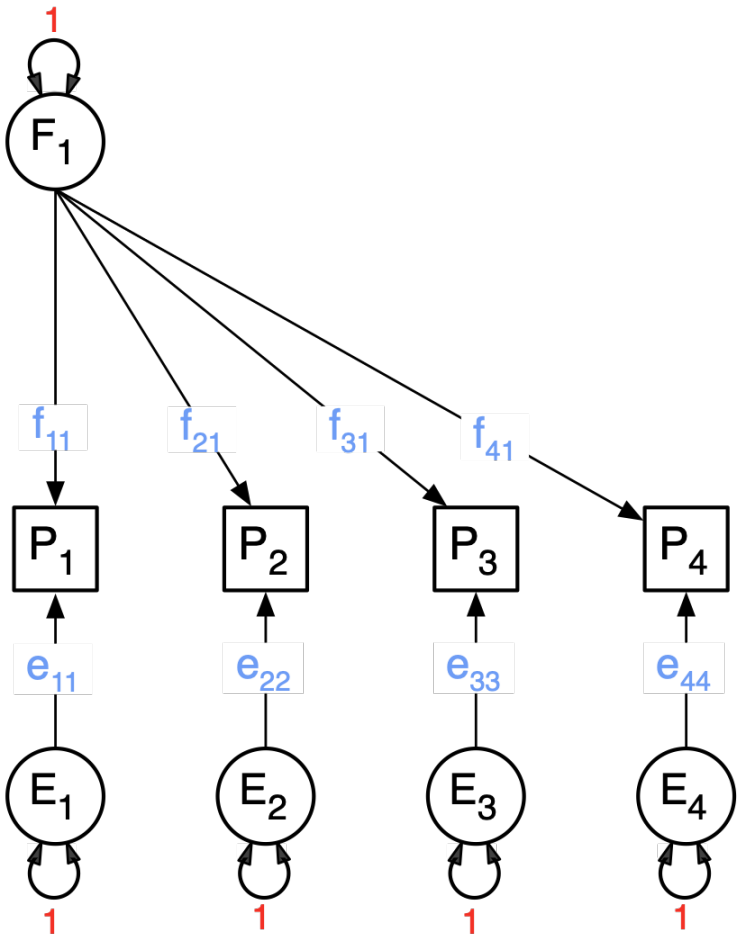




Theoretical Models



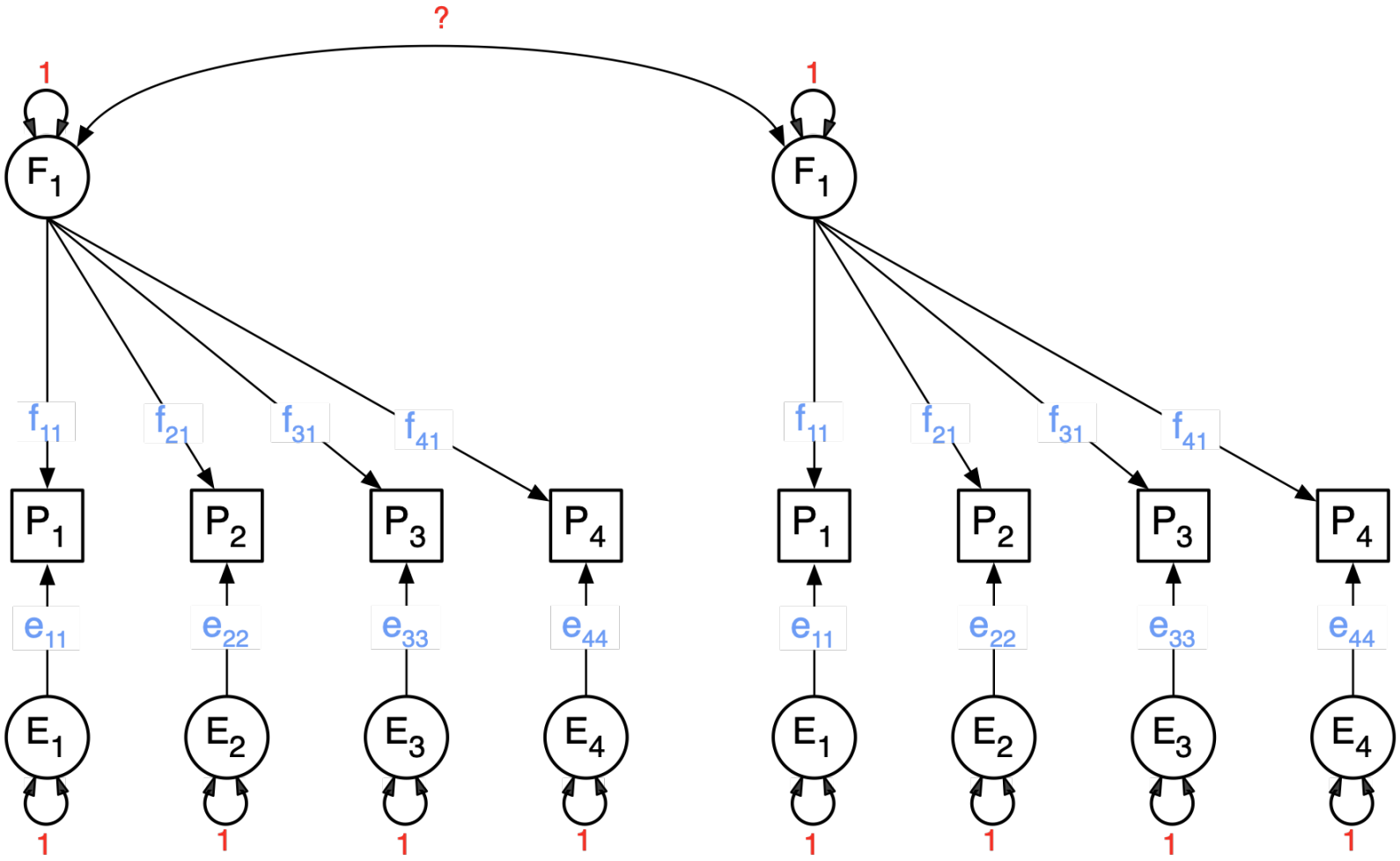
General factor + Residuals



$$\begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_4 \end{array} \begin{array}{c} F_1 \\ \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{41} \end{bmatrix} \end{array}$$

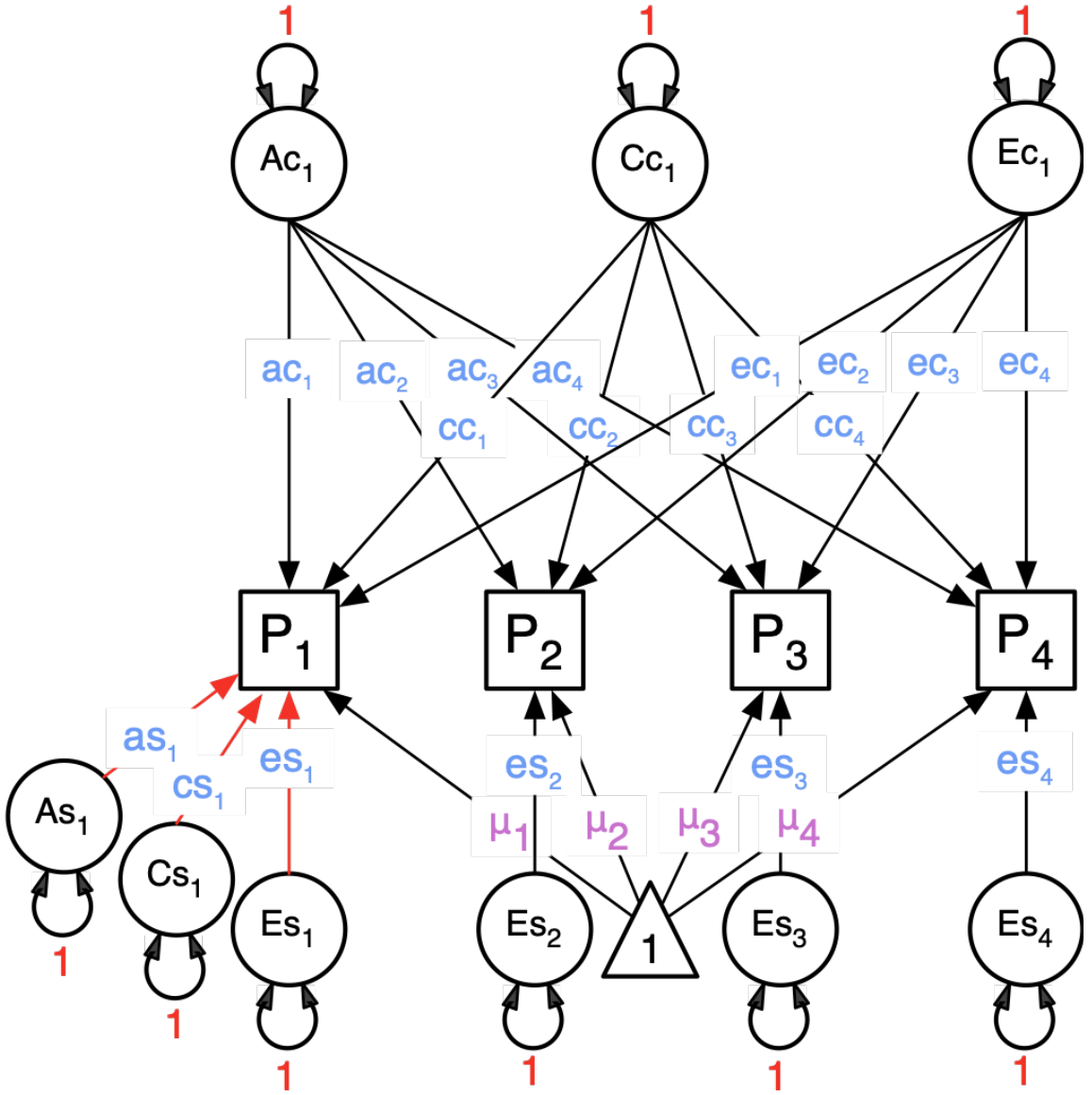
$$\begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_4 \end{array} \begin{bmatrix} e_{11} & 0 & 0 & 0 \\ 0 & e_{22} & 0 & 0 \\ 0 & 0 & e_{33} & 0 \\ 0 & 0 & 0 & e_{44} \end{bmatrix}$$

In Twins we duplicate this for a pair (or more), and then duplicate the result for MZ and DZ (or more)

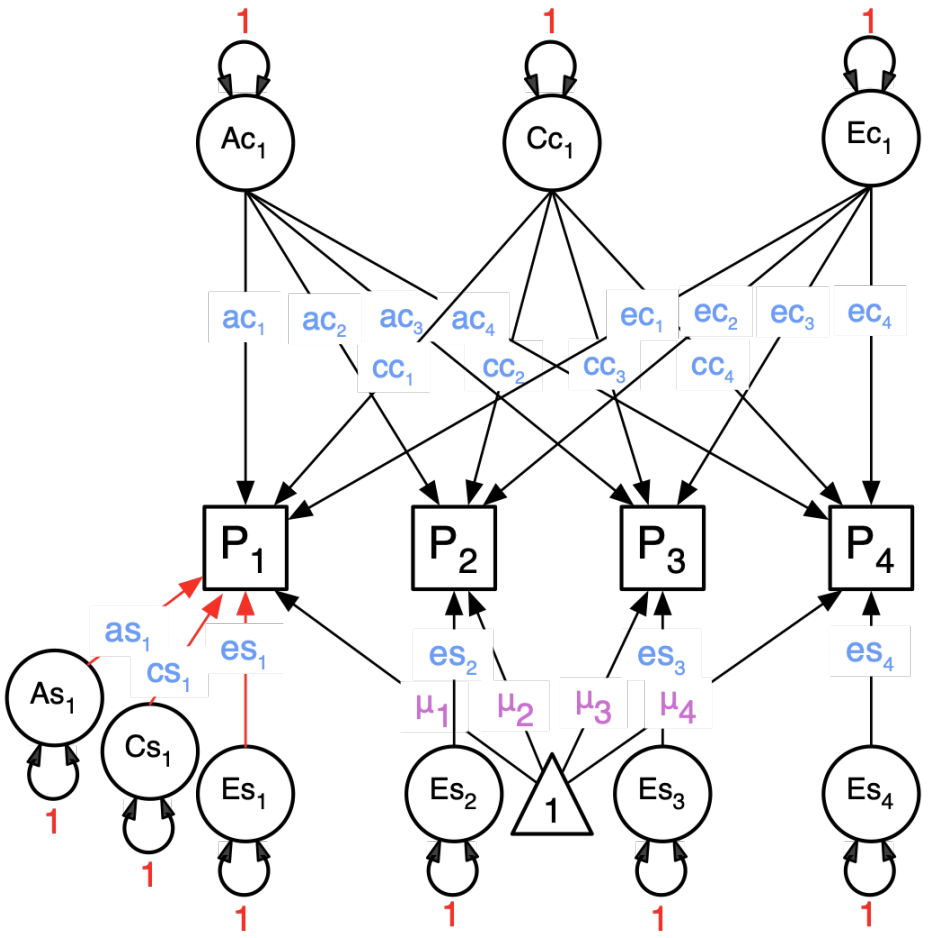


Independent Pathway Model

- Biometric model
- Allows covariance structures for A, C and E to differ among each other



General A, C, and E factors, and A, C, and E Specific factors



$$\begin{matrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{matrix} \begin{bmatrix} es_{11} & 0 & 0 & 0 \\ 0 & es_{22} & 0 & 0 \\ 0 & 0 & es_{33} & 0 \\ 0 & 0 & 0 & es_{44} \end{bmatrix}$$

$$\begin{bmatrix} as_{11} & 0 & 0 & 0 \\ 0 & as_{22} & 0 & 0 \\ 0 & 0 & as_{33} & 0 \\ 0 & 0 & 0 & as_{44} \end{bmatrix}$$

$$\begin{bmatrix} cs_{11} & 0 & 0 & 0 \\ 0 & cs_{22} & 0 & 0 \\ 0 & 0 & cs_{33} & 0 \\ 0 & 0 & 0 & cs_{44} \end{bmatrix}$$

Total A Covariance: we add up the components

ac %*% t(ac)

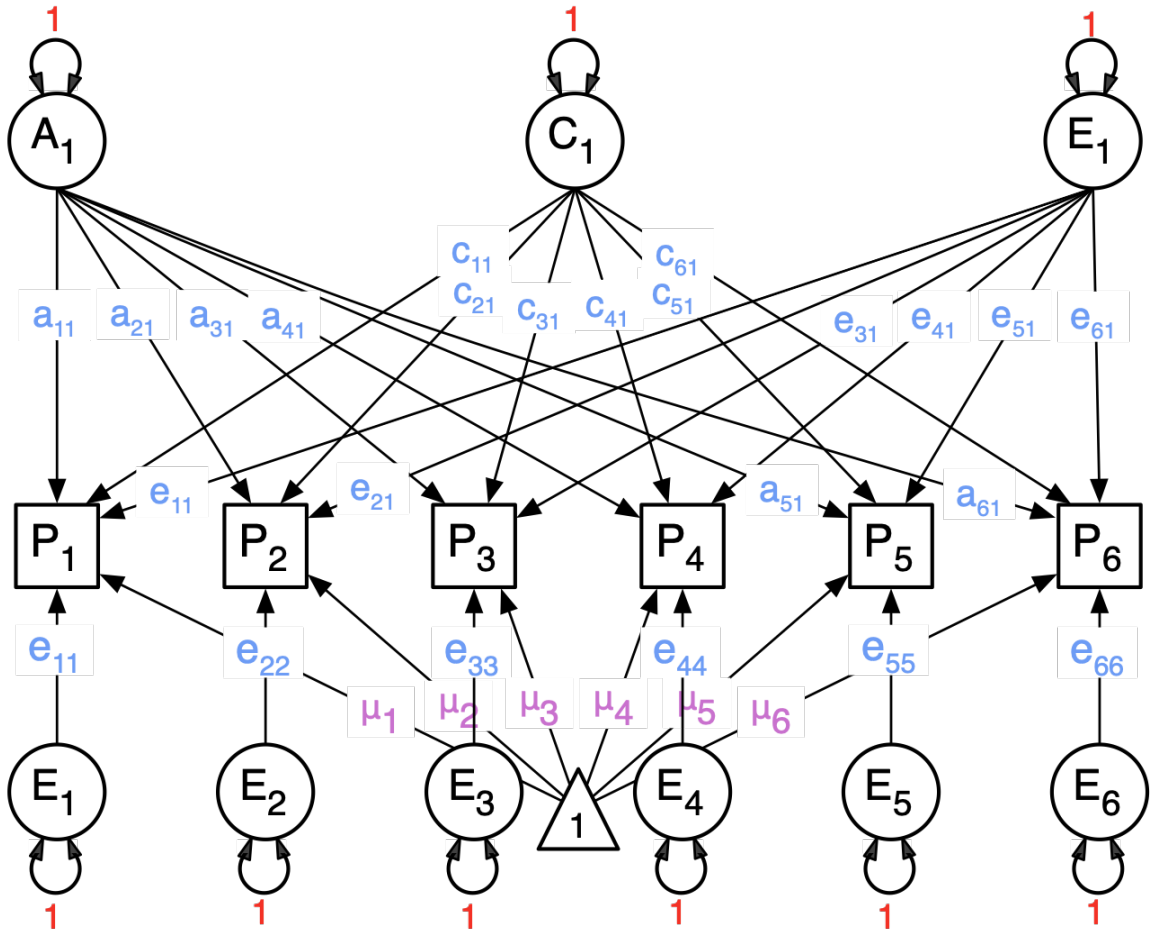
$$\begin{bmatrix} ac_{11}^2 & ac_{11}ac_{21} & ac_{11}ac_{31} & ac_{11}ac_{41} \\ ac_{21}ac_{11} & ac_{21}^2 & ac_{21}ac_{31} & ac_{21}ac_{41} \\ ac_{31}ac_{11} & ac_{31}ac_{21} & ac_{31}^2 & ac_{31}ac_{41} \\ ac_{41}ac_{11} & ac_{41}ac_{21} & ac_{41}ac_{31} & ac_{41}^2 \end{bmatrix} + \begin{bmatrix} as_{11}^2 & 0 & 0 & 0 \\ 0 & as_{22}^2 & 0 & 0 \\ 0 & 0 & as_{33}^2 & 0 \\ 0 & 0 & 0 & as_{44}^2 \end{bmatrix} = \begin{bmatrix} ac_{11}^2 + as_{11}^2 & ac_{11}ac_{21} & ac_{11}ac_{31} & ac_{11}ac_{41} \\ ac_{21}ac_{11} & ac_{21}^2 + as_{22}^2 & ac_{21}ac_{31} & ac_{21}ac_{41} \\ ac_{31}ac_{11} & ac_{31}ac_{21} & ac_{31}^2 + as_{33}^2 & ac_{31}ac_{41} \\ ac_{41}ac_{11} & ac_{41}ac_{21} & ac_{41}ac_{31} & ac_{41}^2 + as_{44}^2 \end{bmatrix}$$

as %*% t(as)

object: CovA
matrix name:A

```
covA      <- mxAlgebra( expression=ac %*% t(ac) + as %*% t(as), name="A" )
```

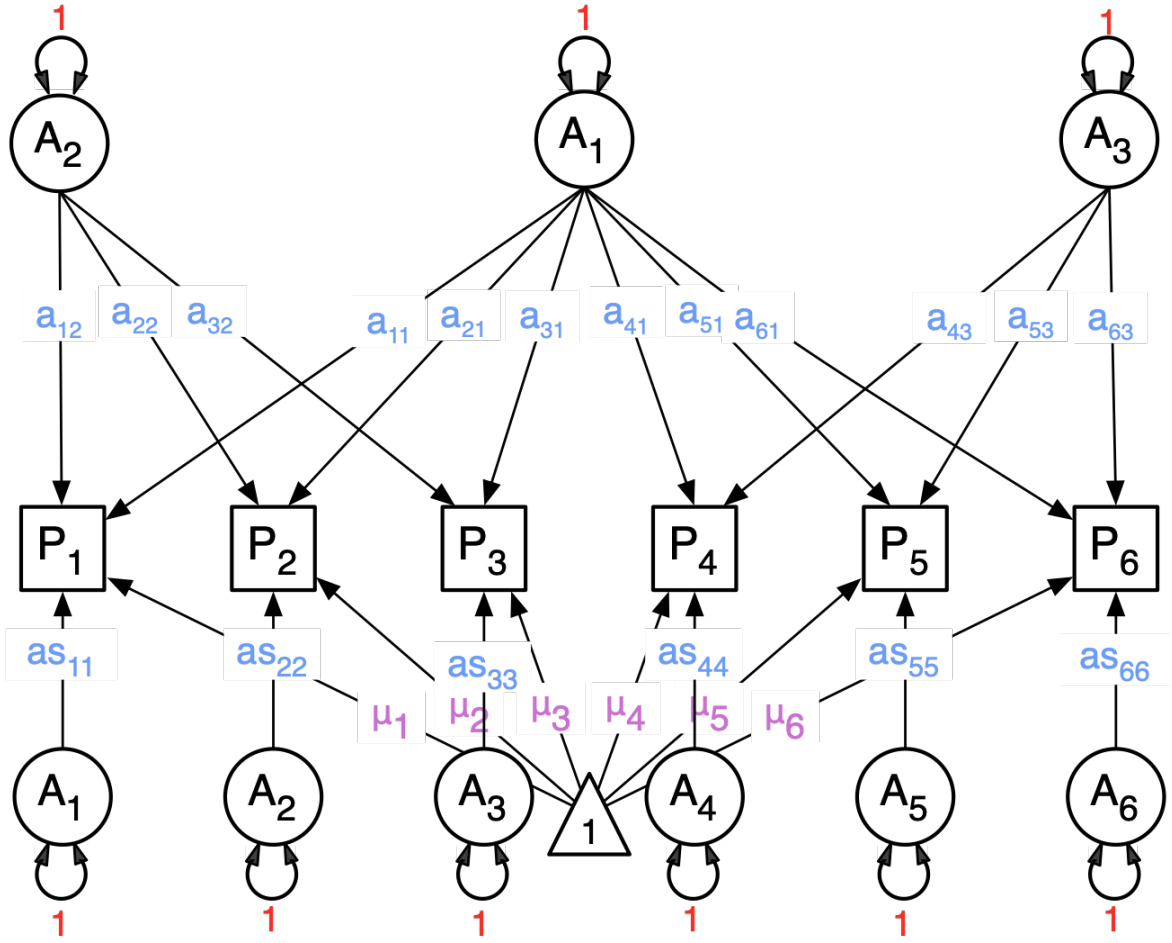
IP Model: Can we add more general factors?



Identification

- Be careful when adding factors: total parameters per source of variance can not exceed (nv*(nv+1))/2
- For a common factor with only 2 indicators the two factor loadings on the latent factor need to be equated OR instead a correlation could be estimated between the residual factors (of the same source of variance) on the two indicators.

Bi-factor model implemented as 3 factors with a bi-factor configuration (1 general)



Common Pathway Model

- Psychometric model
- Same covariance structure for A, C and E

Behavior Genetics, Vol. 20, No. 5, 1990

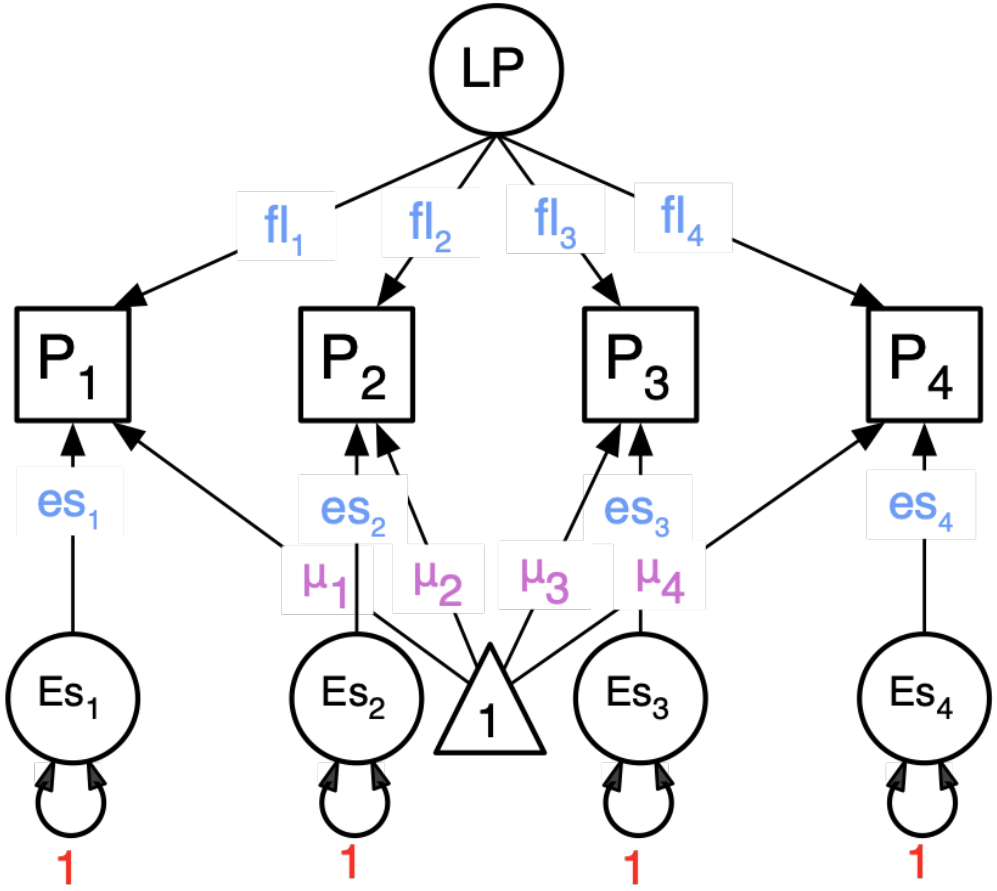
Alternative Common Factor Models for Multivariate Biometric Analyses

J. J. McArdle[1] and H. H. Goldsmith [2]

Received 4 Apr. 1988—Final 15 May 1990

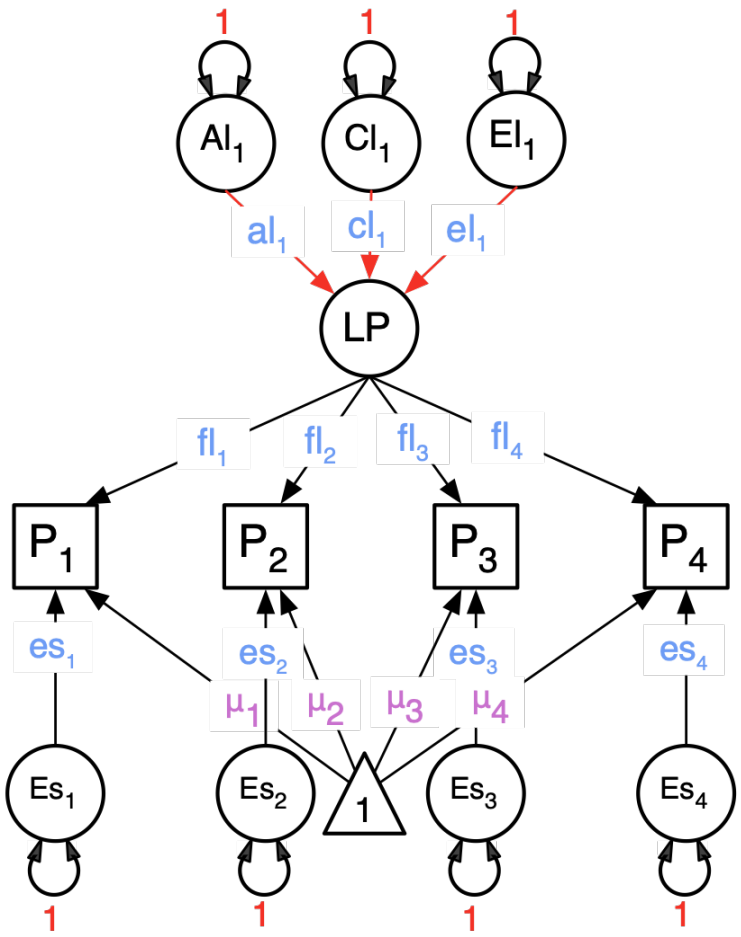
In prior research we have shown how linear structural equation models and computer programs (e.g., LISREL) may be simply and directly used to provide alternatives for the traditional biometric twin design. We use structural equations and path models to define biometric group differences, we write traditional common-factor models in the same way, and then we take a detailed look at some alternative multivariate and biometric models. We contrast the biometric-factors covariance structure approach used by Loehlin and Vandenberg (1968), Martin and Eaves (1977), and others with the psychometric-factors approach used by McArdle et al. (1980) and others. We use the multivariate primary mental abilities data on monozygotic (MZ) and dizygotic (DZ) twins from Loehlin and Vandenberg (1968) to detail fundamental differences in model specification and results. We extend both multivariate biometric approaches using exploratory and confirmatory multiple-factor models. These comparisons show that each alternative multivariate methodology has useful features for empirical applications.

Unlike the IP model, we implement latent traits, and these load (Factor Loadings) on our manifests (“items”)



$$\begin{matrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{matrix} \begin{bmatrix} fl_{11} \\ fl_{21} \\ fl_{31} \\ fl_{41} \end{bmatrix}$$

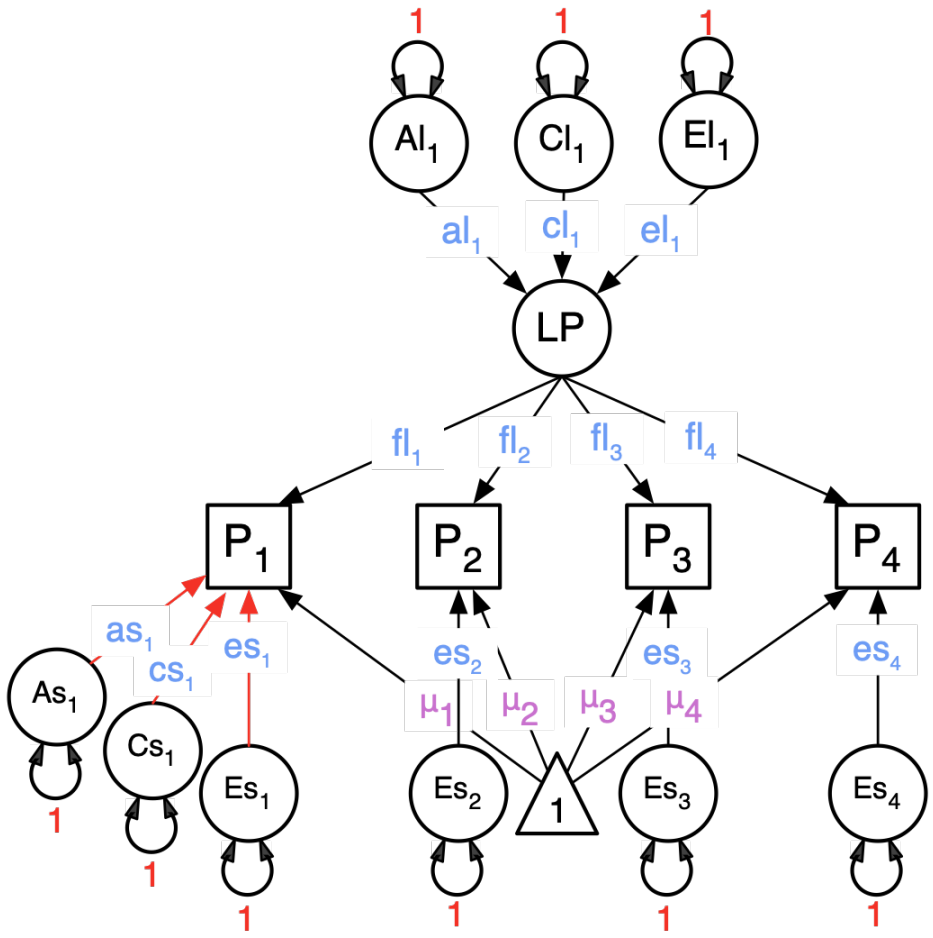
Explaining a Latent Phenotype with ACE components



$$\begin{bmatrix} al_{11} \end{bmatrix} \quad \begin{bmatrix} cl_{11} \end{bmatrix} \quad \begin{bmatrix} el_{11} \end{bmatrix}$$

$$\begin{matrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{matrix} \begin{bmatrix} fl_{11} \\ fl_{21} \\ fl_{31} \\ fl_{41} \end{bmatrix}$$

CP Model: Latent common trait + ACE Specifics



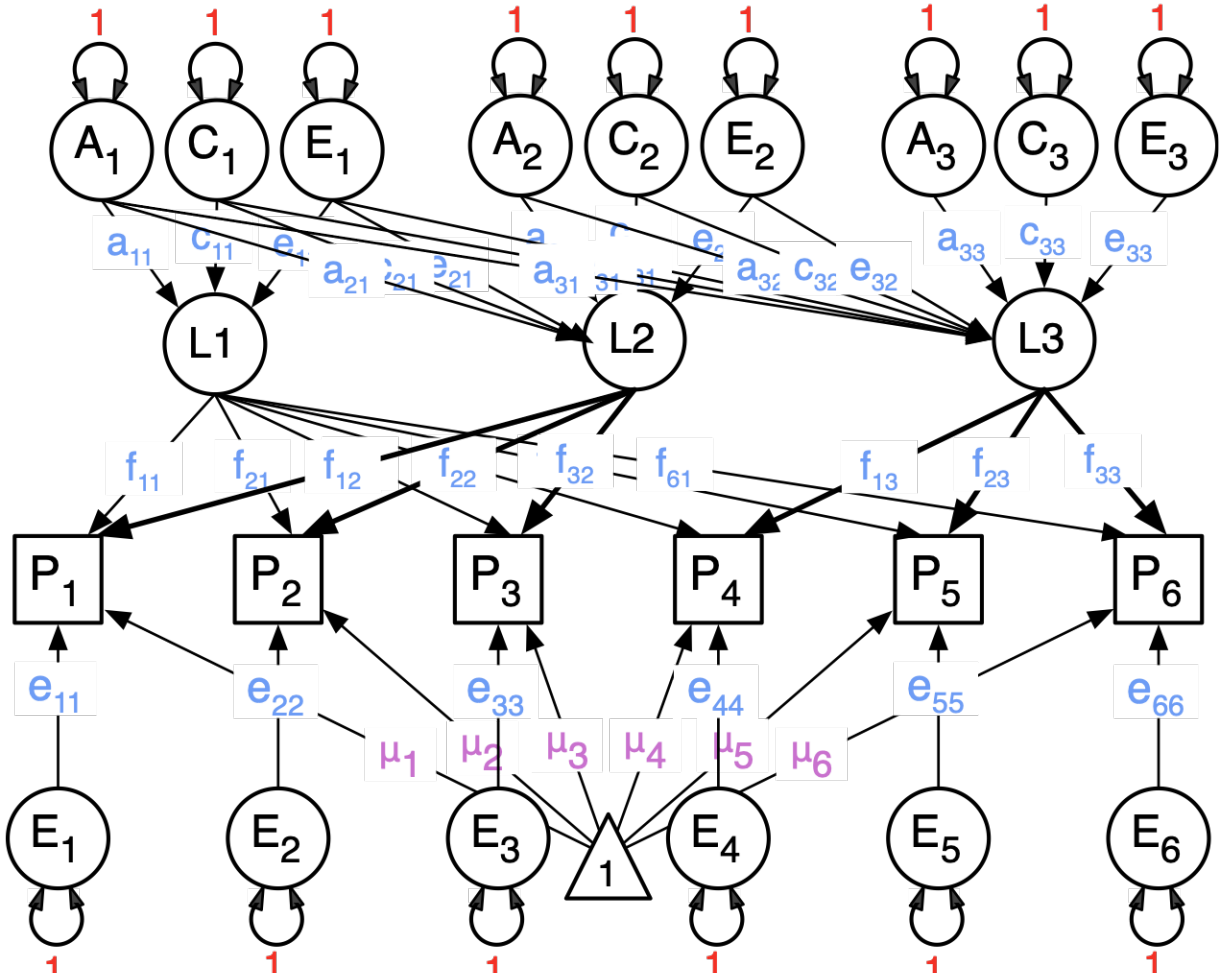
$$\begin{matrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{matrix} \begin{bmatrix} es_{11} & 0 & 0 & 0 \\ 0 & es_{22} & 0 & 0 \\ 0 & 0 & es_{33} & 0 \\ 0 & 0 & 0 & es_{44} \end{bmatrix}$$

$$\begin{bmatrix} as_{11} & 0 & 0 & 0 \\ 0 & as_{22} & 0 & 0 \\ 0 & 0 & as_{33} & 0 \\ 0 & 0 & 0 & as_{44} \end{bmatrix}$$

$$\begin{bmatrix} cs_{11} & 0 & 0 & 0 \\ 0 & cs_{22} & 0 & 0 \\ 0 & 0 & cs_{33} & 0 \\ 0 & 0 & 0 & cs_{44} \end{bmatrix}$$

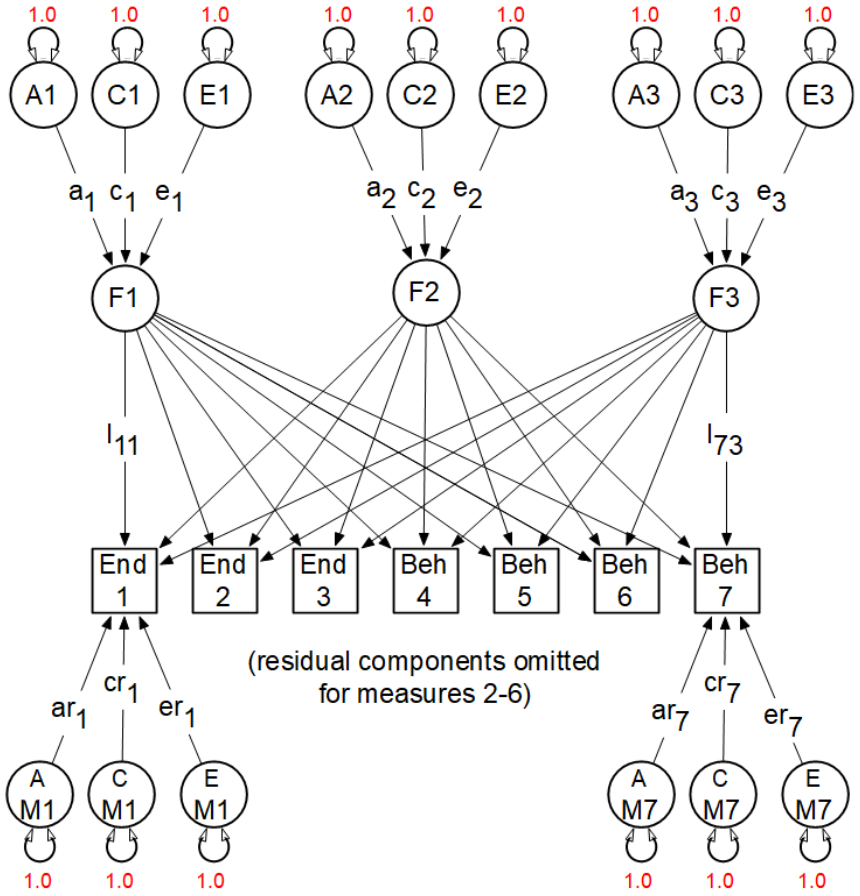
Extensible: 3-factor CP model

(not nearly used widely enough!)

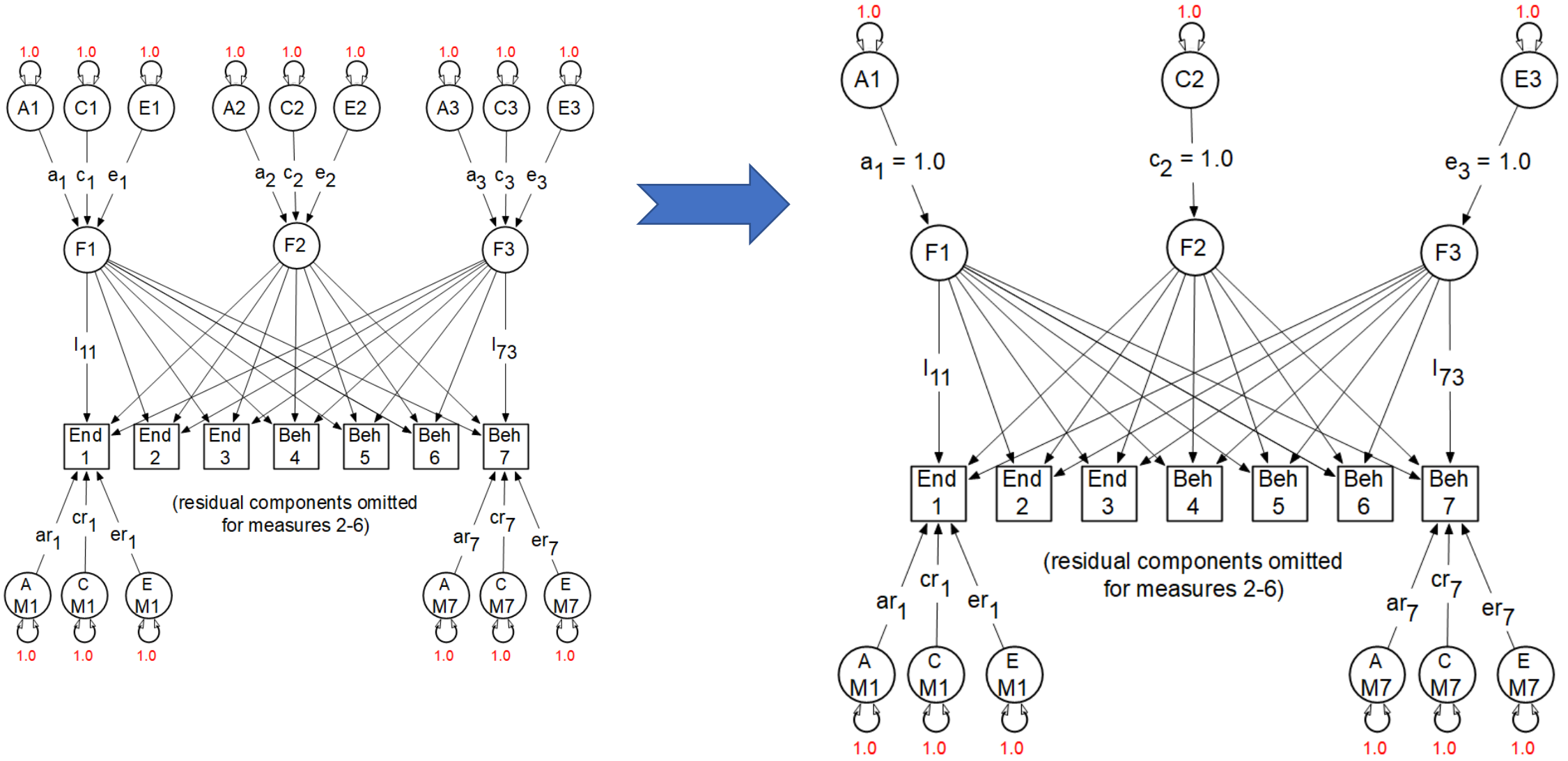


Nesting: Can we get from a CP model to an IP model by deleting paths?

- Comparisons require nesting
  - (or AIC)

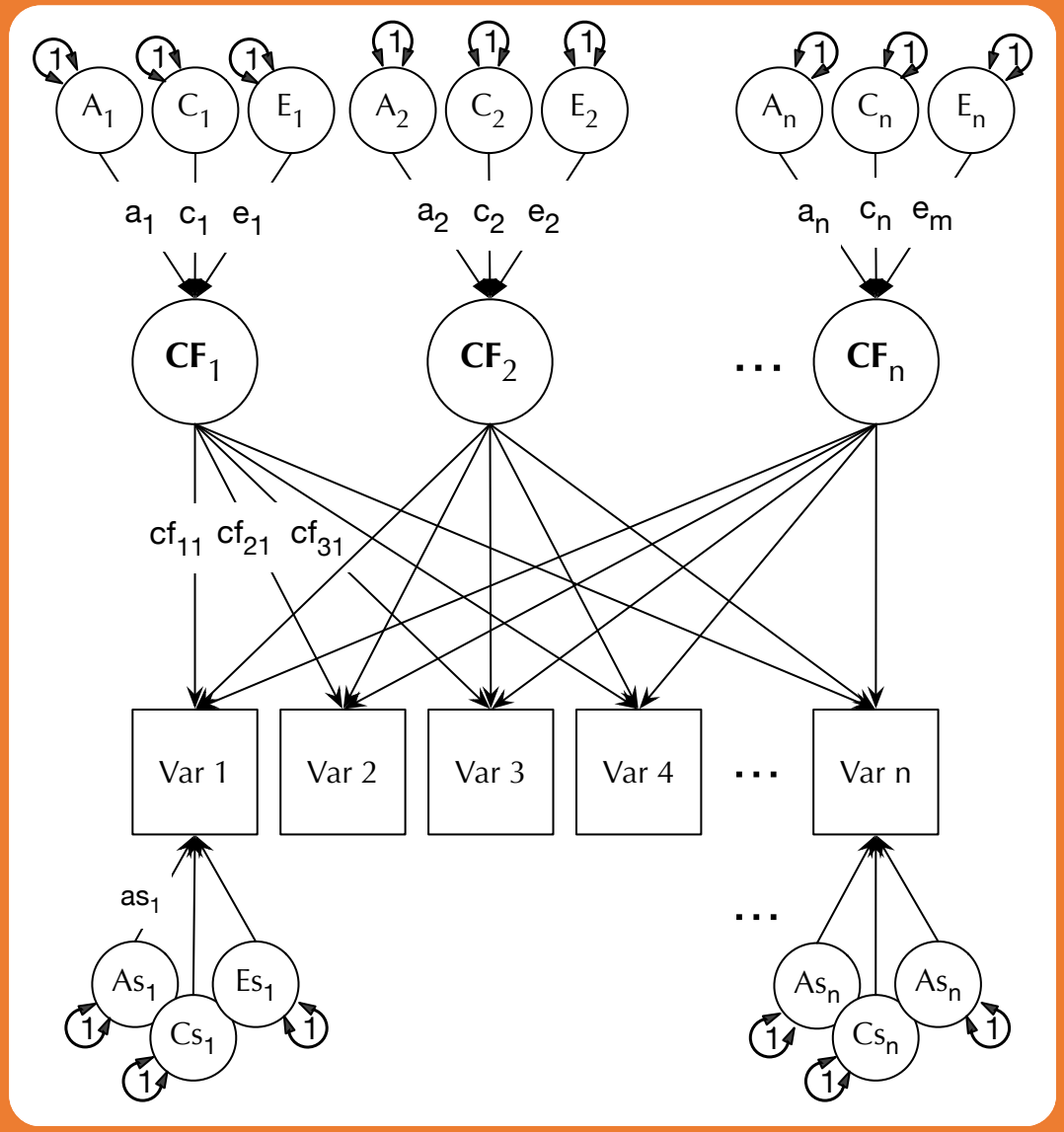


Nesting: Independent Pathway: 3 factors

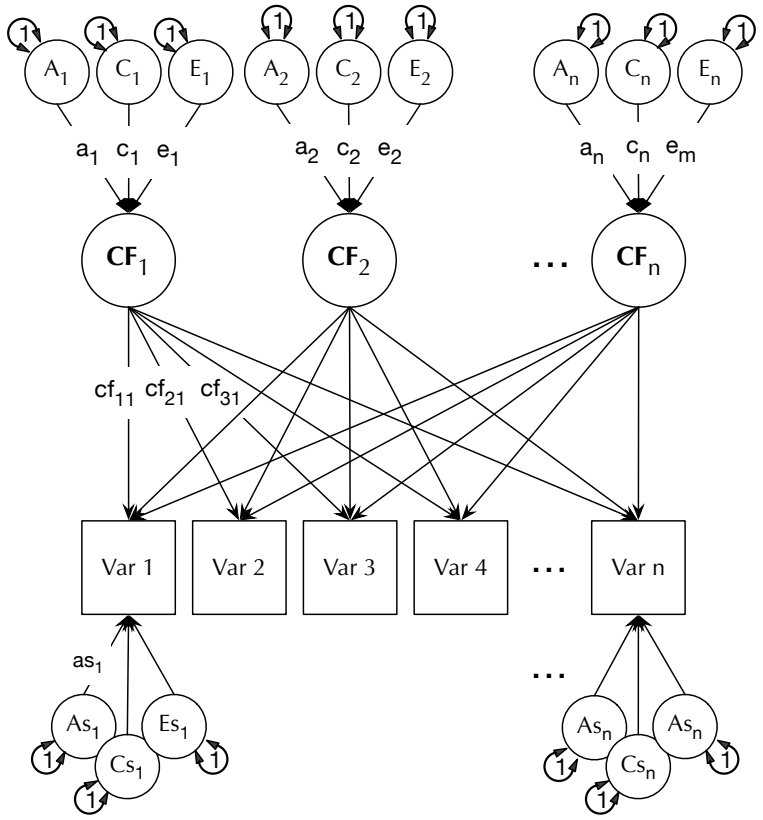


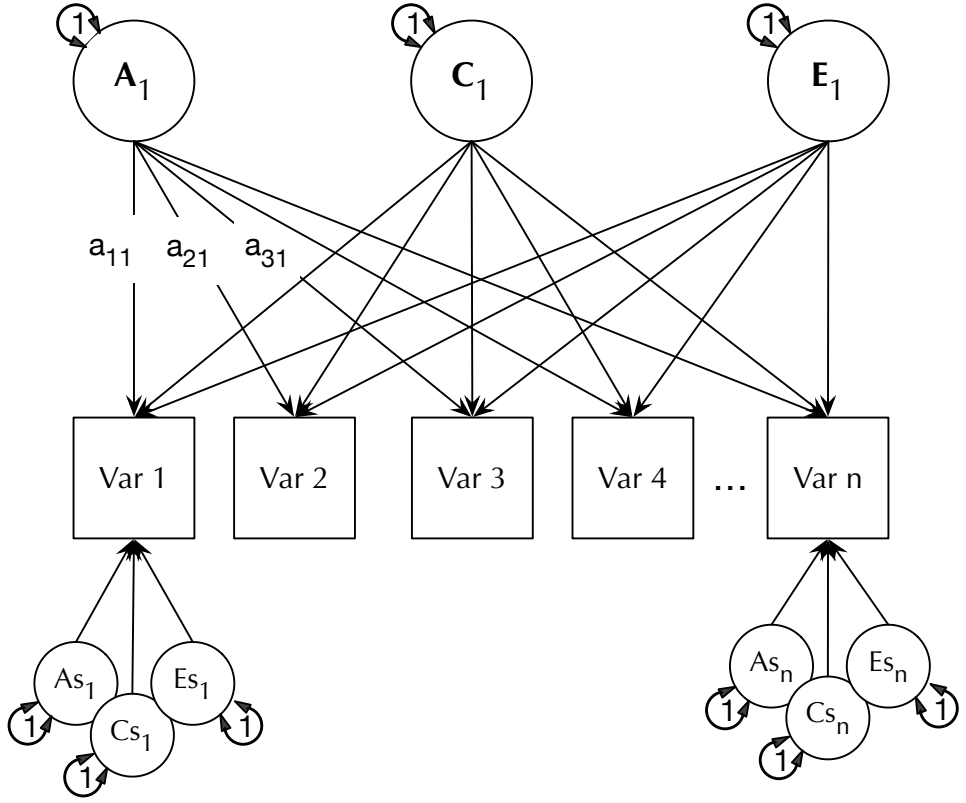
umxCP & umxIP practical

Timothy C. Bates



Theoretical models of phenotypes





OpenMx note

- Some problems are solved more readily by some optimizers than others.
- In this case, SLSQP is often better than CSOLNP
- umx_set_optimizer("SLSQP")

umxCP parameters

- umxCP(name= "CP", **selDVs**, **dzData**, **mzData**, **sep**, **nFac** = 1)

- Other parameters:
  - freeLowerA = FALSE, freeLowerC = FALSE, freeLowerE = FALSE, correlatedA = FALSE, equateMeans = TRUE, dzAr = 0.5, dzCr = 1, addStd = T, addCI = TRUE, autoRun = getOption("umx_auto_run"), optimizer = NULL

umxCP()

1. Use umxCP() to build m1: A common factor model of the GFF
2. Does it fit significantly worse than a base model (m0) built with umxACEv() ?

```
data(GFF)
base=c("gff", "hap", "sat", "AD", "SOMA", "SOC")
tmp= umx_scale_wide_twin_data(base, sep="_T", data=GFF)
mzData= subset(tmp, zyg_2grp == "MZ")
dzData= subset(tmp, zyg_2grp == "DZ")
```

Running umxCP

m1= umxCP(selDVs= base, sep="_T", mzData=mzData, dzData= dzData)

What does table about “common factor” mean?

- 'log Lik.' 30945.15 (df=33)
- Common Factor paths

	A	C	E
Common.factor.1	0.7	0.38	0.61

Loading of each trait on the Common Factors

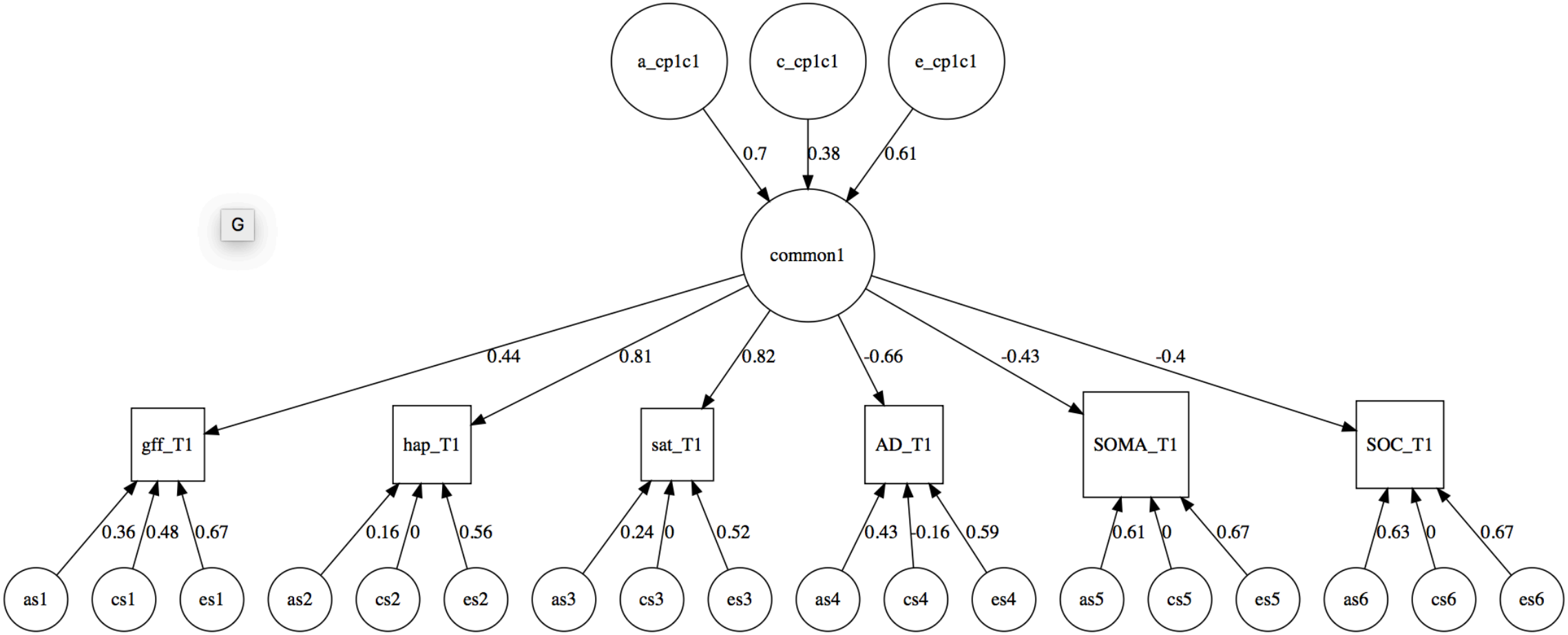
	CP1
gff	0.44
hap	0.81
sat	0.82
AD	-0.66
SOMA	-0.43
SOC	-0.40

Standardized specific loadings

	gff	hap	sat	AD	SOMA	SOC
Specific a	0.36	0.16	0.24	0.43	0.61	0.63
Specific c	0.48	.	.	0.16	.	.
Specific e	0.67	0.56	0.52	0.59	0.67	0.67

Standardized Specific Loadings

	As1	As2	As3	As4	As5	As6	Cs1	Cs2	Cs3	Cs4	Cs5	Cs6	Es1	Es2	Es3
gff	0.36						0.48						0.67		
hap	.	0.16					.	.					.	0.56	
sat	.	.	0.24				.	.	.				.	.	0.52
AD	.	.	.	0.43			.	.	.	-0.16			.	.	.
SOMA	.	.	.	.	0.61		.	.	.	.	.		.	.	.
SOC	.	.	.	.	.	0.63	.	.	.	.	.	.	.	.	.


G
a_cp1c1
c_cp1c1
e_cp1c1
0.7
0.38
0.61
common1
0.44
0.81
0.82
-0.66
-0.43
-0.4
gff_T1
hap_T1
sat_T1
AD_T1
SOMA_T1
SOC_T1
0.36
0.48
0.67
0.16
0
0.56
0.24
0
0.52
0.43
-0.16
0.59
0.61
0
0.67
0.63
0
0.67
as1
cs1
es1
as2
cs2
es2
as3
cs3
es3
as4
cs4
es4
as5
cs5
es5
as6
cs6
es6

Improve the CP model fit

- Make `m2 = umxCP()` model with 2 common factors
  - What did you need to set to do that?

- Is it better than the 1 factor model?
  - What umx function lets you compare models?

- Does it (still) fit significantly worse than the baseline model?
  - Can you compare more than one model?

Identification: How many common factors are allowed?

- Sources of variance allow us to fit parameters
  - Sources of variance = number of phenotypes = “nv” here
- Total parameters/nv must be < $(nv * (nv + 1))/2$
  - Why nv*nv+1/2?
- For a single common factor with only 2 indicators, you would have to equate the factor loadings of the 2 indicators to identify the model
  - Alternatively, remove the common factor and add a correlated residual
- If in doubt, try `mxCheckIdentification()`
  - When not identified, lists the offending parameters

umxCompare

- umxCompare(**base**, **comparison**, all = TRUE, digits = 3)

- umxCompare(**base** = m1, **comparison** =c(m2, m3))
- umxCompare(**base** = c(m1, m2), **comparison** =c(m2, m3))

Questions

- Is a better AIC higher or lower?
- Is AIC -2432 or -1980 to be preferred?

Common Pathway fit with different #s of factors

- ✓ Build a 1-factor model
- ✓ Build a 2-factor model
- Build a 3-factor model
- Test for a 1 factor vs a 2 factor vs a 3 factor CP?

Common Pathway **Modification**

- Test for each common factor being A and E (only – no C)?
  - Need to drop C from the *common factors*
- Test for every specific factor being only E?
  - *Hypothesis*: Residuals are measurement error (or at least unshared effects)
  - Need to drop what from the residuals?
- Fit an ADE model
  - Does it fit as well or better than ACE?
  - What parameter needed to be changed?

Functions that will help: **parameters()**

parameters(x, pattern = ".*",

digits = 2, thresh = c("all", "above", "below", "NS", "sig"), b = NULL)

parameters(m1, pattern = "cp")

parameters(m1, pattern="cp")

```
name Estimate
7         a_cp_r1c1     0.70
8         c_cp_r1c1     0.38
9         e_cp_r1c1     0.61
28 cp_loadings_r1c1     0.45
29 cp_loadings_r2c1     0.83
30 cp_loadings_r3c1     0.82
31 cp_loadings_r4c1    -0.70
32 cp_loadings_r5c1    -0.46
33 cp_loadings_r6c1    -0.39
```

Functions that will help: **umxModify**

umxModify(**lastFit**, **update**, master, **regex** = FALSE, free = FALSE, value = 0, newlabels, freeToStart = NA, name, **comparison** = FALSE, autoRun = TRUE)

Common Pathway Modification

- Test for each common factor being A and C (only – no E)?
  - Need to drop E from the common factors

Common Pathway Modification

- Test for every specific factor being only E?
  - Hypothesis: The residuals are measurement error (or at least unshared effects)
  - Need to drop what from the residuals?

Common Pathway Modification

- Fit an ADE model
  - What parameter needed to be changed?
  - Does it fit as well or better than ACE?

Common Pathway Modification: AP (advanced placement question) ☺

- A friend thinks common environment determines one common factor, while the others are a mixture of A and C
  - Can you test her hypothesis?

Same again for umxIP (independent pathway)

- Build a 1 factor IP model
  - `ip1 = umxIP()`
- Compare a 1 factor vs a 2 factor vs a 3 factor IP?
- Test for each A influence having the same loading?
  - What does that imply?
- Can you test for every specific factor being only E?
- Can you fit an ADE model with umxIP?