Structural Equations with Latent Variables

Session 12, Lecture 8

11/22/06

Outline

- Examples using AMOS
- Identification
-         a) review
- Three useful types of SEMs
  - MIMIC model
  - multitrait-multimethod SEM
  - causal indicators
- Group comparison models
  - two-way interactions across two groups
  - two-way interactions across more than two groups
  - three-way interactions across two groups
  - three-way interactions across more than two groups

Rules for Variances

- Var(A + B)=Var(A)+Var(B)+2Cov(A,B)
- Var(A - B)= Var(A)+Var(B)-2Cov(A,B)

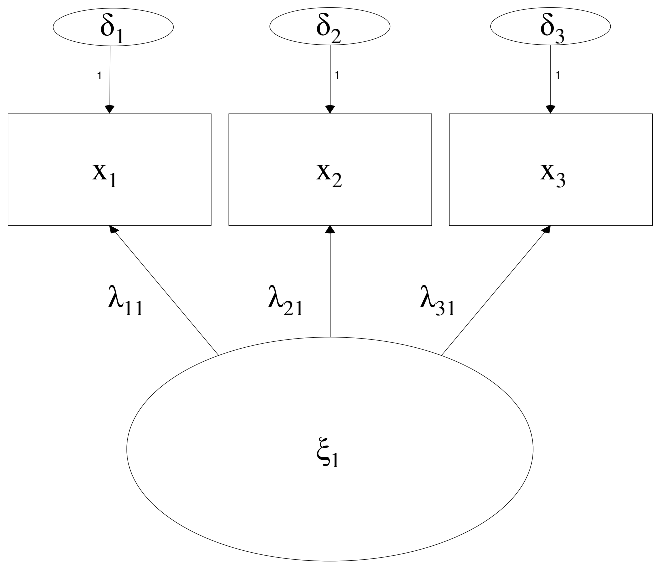
Let c be any constant (like, 5)

- – Var(c+A)=0 + Var(A) (constants don't vary)
- – Var(cA) = c^2Var(A)

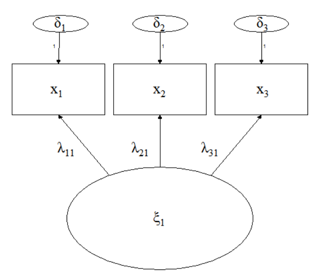
Rules for Covariances

- Cov(X,Y)=E(XY)-E(X)E(Y)

- Cov(c,X)=0
- Cov(cX,Y)=cCov(X,Y)
- Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)



Equations for path diagram:



$x_1 = \lambda_{11}\xi_1 + \delta_1$
$x_2 = \lambda_{21}\xi_1 + \delta_2$
$x_3 = \lambda_{31}\xi_1 + \delta_3$

Writing out equations:
$Var(x_1) = Cov(\lambda_{11}\xi_1 + \delta_1, \lambda_{11}\xi_1 + \delta_1)$ *distribute (FOIL)*

$= Cov(\lambda_{11}\xi_1, \lambda_{11}\xi_1) + Cov(\lambda_{11}\xi_1, \delta_1) + Cov(\delta_1, \lambda_{11}\xi_1) + Cov(\delta_1, \delta_1)$

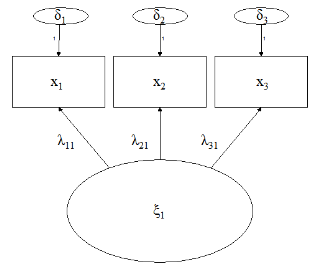
$= Var(\lambda_{11}\xi_1) + 2Cov(\lambda_{11}\xi_1, \delta_1) + Var(\delta_1)$
** $Var(cX) = c^2(X)$, also, assume that δ is uncorrelated with X*

$= \lambda_{11}^2 Var(\xi_1) + 0 + Var(\delta_1)$

Similarly,

$Var(x_2) = \lambda_{21}^2 Var(\xi_1) + Var(\delta_2)$

$Var(x_3) = \lambda_{31}^2 Var(\xi_1) + Var(\delta_3)$



$Cov(x_1,x_2) = Cov(\lambda_{11}\xi_1 + \delta_1, \lambda_{21}\xi_1 + \delta_2)$

$= Cov(\lambda_{11}\xi_1, \lambda_{21}\xi_1) + Cov(\lambda_{21}\xi_1, \delta_2) + Cov(\delta_1, \lambda_{21}\xi_1) + Cov(\delta_1, \delta_2)$

No covariance between ξs,δs *and* δs *are uncorrelated*

$= \lambda_{11}\lambda_{21} Cov(\xi_1,\xi_1) + 0 + 0 + 0 = \lambda_{21}\lambda_{11} Var(\xi_1)$

Similarly,

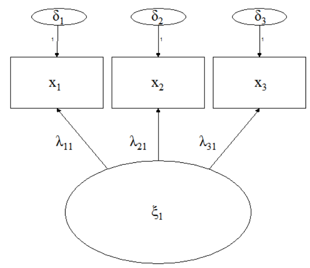
$Cov(x_3,x_1) = \lambda_{31}\lambda_{11} Var(\xi_1)$

$Cov(x_3,x_2) = \lambda_{31}\lambda_{21} Var(\xi_1)$

Why do all these calculations?

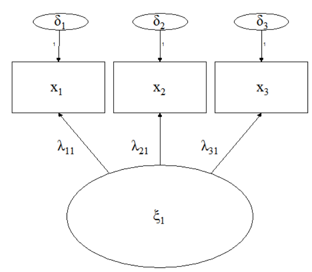
- Facilitates deeper understanding
- Highlights model assumptions
- Aids in determining if model is identified

Is model identified?



- 3 indicator rule?
  - √ at least one factor
  - √ at least 3 indicators
  - √ each x pointed to by only one ξ
  - √ δs are uncorrelated

- But, on a deeper level, why is a 3 indicator model identified?



How many things are we estimating?

vars of exog. vars: 1
vars of errors for endog: 3
direct effects: 3
<u>double-headed arrows: 0 +</u>
Total to be Estimated: **7**

λs: (direct effects) we know:

$Cov(x_2,x_1) = \lambda_{21}\lambda_{11} Var(\xi_1)$

$Cov(x_3,x_1) = \lambda_{31}\lambda_{11} Var(\xi_1)$

$Cov(x_3,x_2) = \lambda_{31}\lambda_{21} Var(\xi_1)$

4 unknowns, 3 equations

$$\lambda_{21} = \frac{\lambda_{11} Cov(x_3, x_1)}{Cov(x_3, x_2)}$$

$$\lambda_{31} = \frac{\lambda_{11} Cov(x_2, x_1)}{Cov(x_3, x_2)}$$

Variances of δs

We know:

$\delta_1 = x_1 - \lambda_{11}\xi_1$ $\quad$ $Var(\delta_1) = Var(x_1) - \lambda_{11}^2\, Var(\xi_1)$

$\delta_2 = x_2 - \lambda_{21}\xi_1$ $\quad$ $Var(\delta_2) = Var(x_2) - \lambda_{21}^2\, Var(\xi_1)$

$\delta_3 = x_3 - \lambda_{31}\xi_1$ $\quad$ $Var(\delta_3) = Var(x_3) - \lambda_{31}^2\, Var(\xi_1)$

Variance of ξ_1:

We know:

$Var(x_1) = \lambda_{11}^2 Var(\xi_1) + Var(\delta_1)$:

$$Var(\xi_1) = \frac{Var(x_1) - Var(\delta_1)}{\lambda_{11}^2}$$

$Var(x_2) = \lambda_{21}^2 Var(\xi_1) + Var(\delta_2)$

$$Var(\xi_1) = \frac{Var(x_2) - Var(\delta_2)}{\lambda_{21}^2}$$

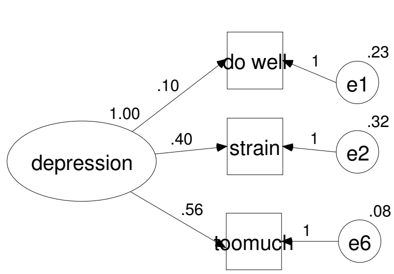
Verdict: as things stand, we don't have enough information to Estimate all the parameters. *Per t-rule: 3(4)/2 = 6*

Solution: Fix one of the parameters values to 1

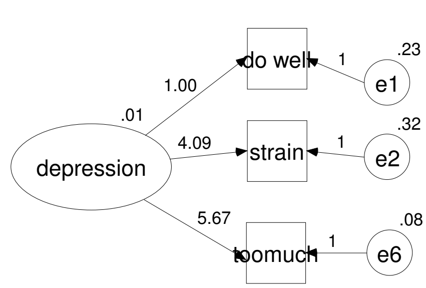
Set $Var(\xi_1)$ to 1, or a λ e.g., λ_{11} to 1

Unstandardized estimates

Set $Var(\xi_1)$ to 1

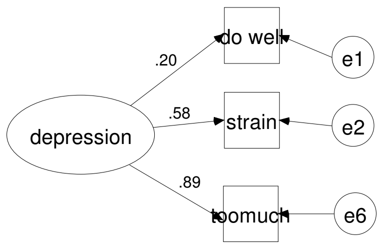


Set λ_{11} to 1

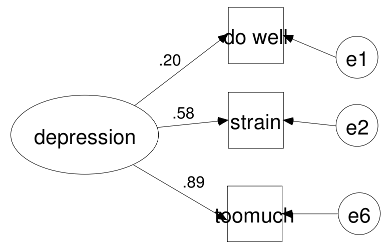


standardized estimates

Set Var(ξ_1) to 1



Set λ_{11} to 1



It doesn't matter what you fix (with standardized estimates).

General identification rules for confirmatory factor analysis:

1) t-rule

2) Three indicator rule. Model passes rule if:
- a) it has three or more indicators per latent variable
- b) each row of Λ_x with one and only one nonzero element
- c) a diagonal Θ_δ

3) Two indicator rule. Model passes rule if:
- a) factor complexity of each x is one
- b) no zero elements in Φ
- c) Θ_δ is diagonal

General identification rules for models that incorporate both structure equations and measurement models.

a) t-rule

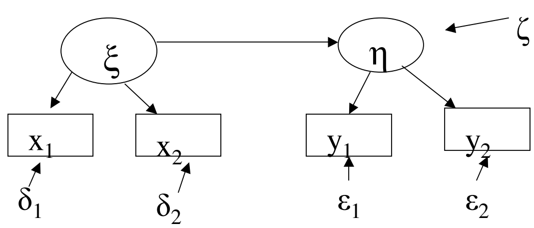
b) two-step rule

1) step 1:

2) step 2:

If both the measurement model and the structural model are identifiable then the model as a whole is identifiable.

Two-Step Rule examples

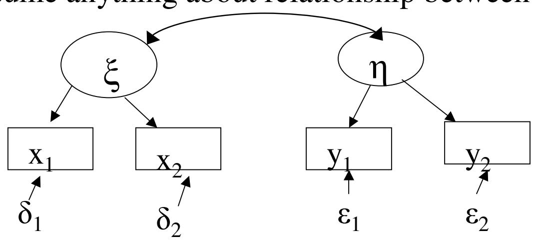


$$\eta = \gamma_1 \xi + \zeta$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \xi + \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}$$

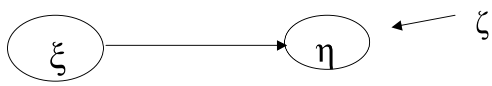
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \lambda_3 \\ \lambda_4 \end{bmatrix} \eta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

Step 1: Measurement Model (has observed variables), doesn't assume anything about relationship between latent variables:



OK per
2-indicator rule

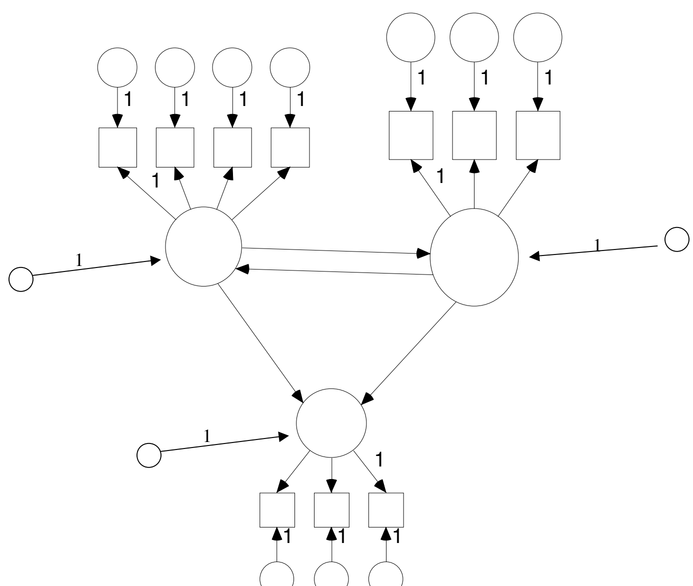
Step 2: Structural Model consists of the relationships between latent constructs – think of the latent variables as observed (in That their variances are known)



OK per
Null B rule

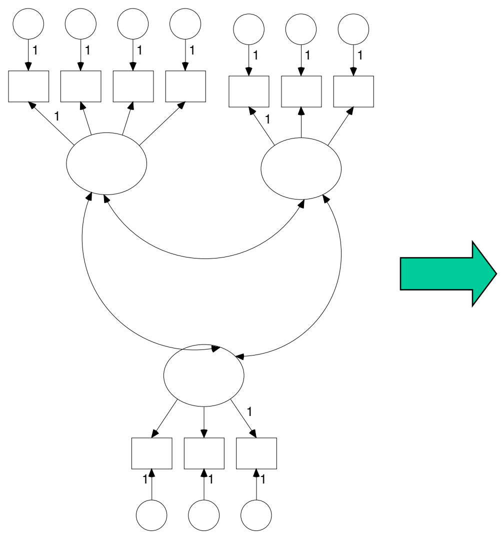
Bollen, pg 326-9

TWO STEP RULE EXAMPLES:

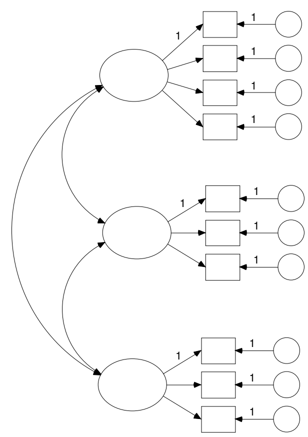


Bollen, pg 324

STEP 1: CFA



OK per
3-indicator rule



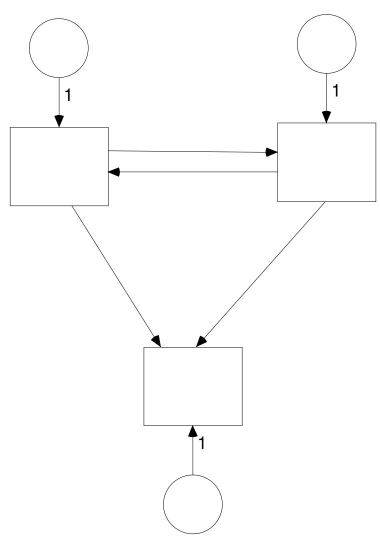
STEP 2: Structural Part

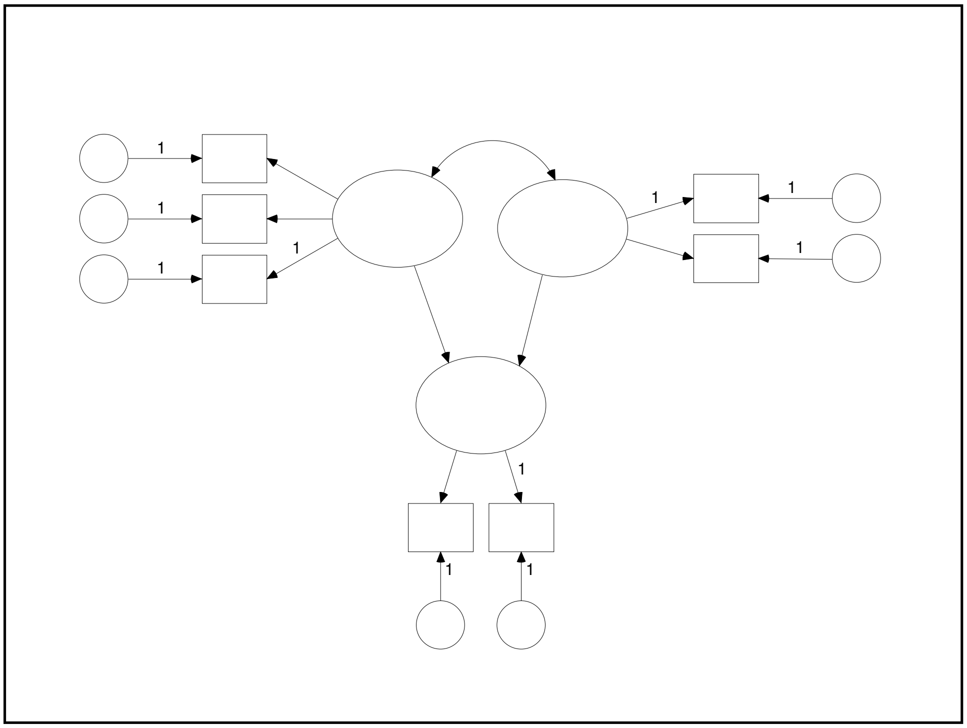
Null B-rule:
Recursive Rule:

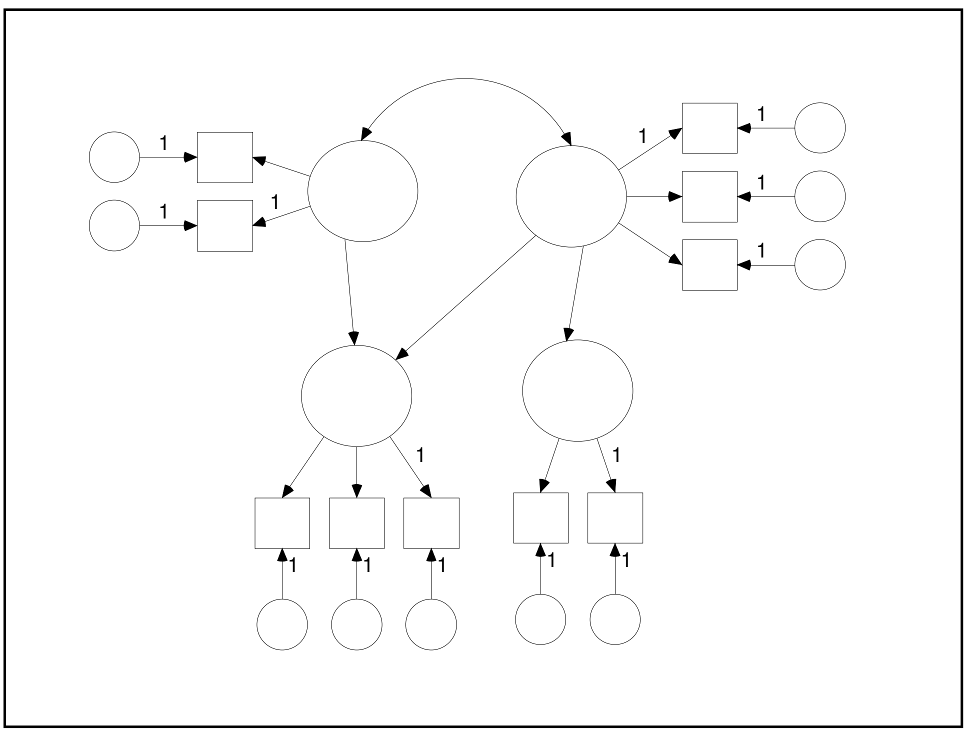
T-Rule:

vars of exog. vars:	0
vars of errors for endog:	3
direct effects:	4
double-headed arrows:	0 +
	7

Sample Moments = (3*4)/2=6

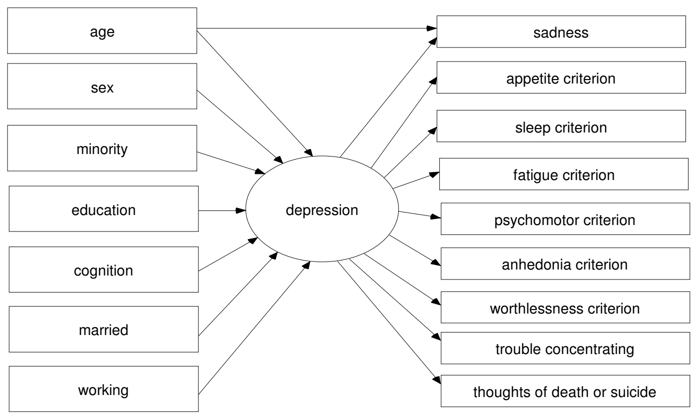



1
1
1
1
1
1
1
1
1
1
1


1
1
1
1
1
1
1
1
1
1
1
1
1
1
1

MIMIC (Multiple Indicators, Multiple Causes)

(from on Gallo, J.J., P. V. Rabins, and J. C. Anthony. 1999. *Psychological Medicine* 29:341-350)



Three major components to the model: causes (left) indicators (right) Latent construct (middle).

Quality of Life Research, **7**, pp. 387–397

The proxy problem: child report versus parent report in health-related quality of life research

N. C. M. Theunissen*, T. G. C. Vogels, H. M. Koopman, G. H. W. Verrips, K. A. H. Zwinderman, S. P. Verloove-Vanhorick and J. M. Wit

Figure 3. MTMM Pearson correlations between the child and parent reports (methods) for seven HRQoL scales (traits).

(methods)	(triats)	CBODY	CMOTOR	CAUTO	CCOGNIT	CSOCIAL	CEMOPOS	CEMONEG	PBODY	PMOTOR	PAUTO	PCOGNIT	PSOCIAL	PEMOPOS	PEMONEG
		Child report							Parent report						
Child report	CBODY	*1.00*													
	CMOTOR	0.47	*1.00*												
	CAUTO	0.32	0.61	*1.00*											
	CCOGNIT	0.39	0.46	0.38	*1.00*										
	CSOCIAL	0.30	0.36	0.33	0.45	*1.00*									
	CEMOPOS	0.23	0.31	0.29	0.28	0.37	*1.00*								
	CEMONEG	0.38	0.34	0.26	0.43	0.48	0.29	*1.00*							
Parent report	PBODY	***0.61***	0.29	0.23	0.25	0.21	0.16	0.27	*1.00*						
	PMOTOR	0.24	***0.50***	0.38	0.24	0.19	0.22	0.19	0.35	*1.00*					
	PAUTO	0.18	0.39	***0.48***	0.20	0.20	0.16	0.14	0.26	0.55	*1.00*				
	PCOGNIT	0.23	0.24	0.14	***0.61***	0.21	0.17	0.24	0.27	0.31	0.28	*1.00*			
	PSOCIAL	0.21	0.24	0.26	0.28	***0.51***	0.29	0.33	0.23	0.30	0.31	0.33	*1.00*		
	PEMOPOS	0.17	0.22	0.19	0.23	0.31	***0.44***	0.27	0.22	0.33	0.25	0.26	0.40	*1.00*	
	PEMONEG	0.29	0.23	0.19	0.27	0.36	0.31	***0.55***	0.32	0.24	0.20	0.27	0.49	0.38	*1.00*

Identification of multi-trait multi-method SEMs

Must have at least three traits and three methods.

(Alwin, 1973)

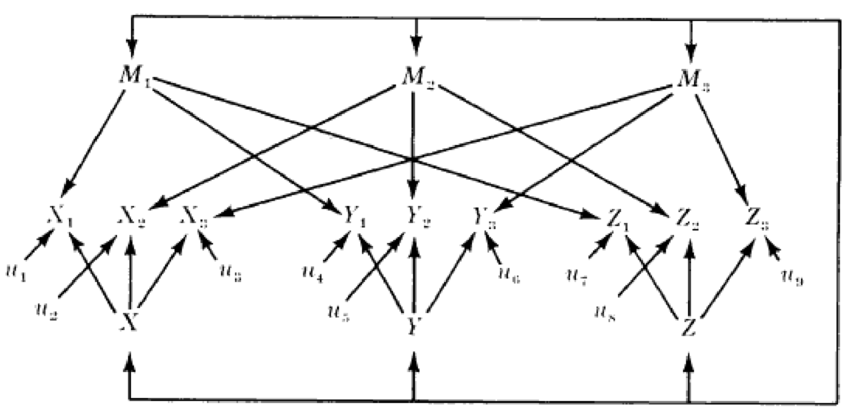
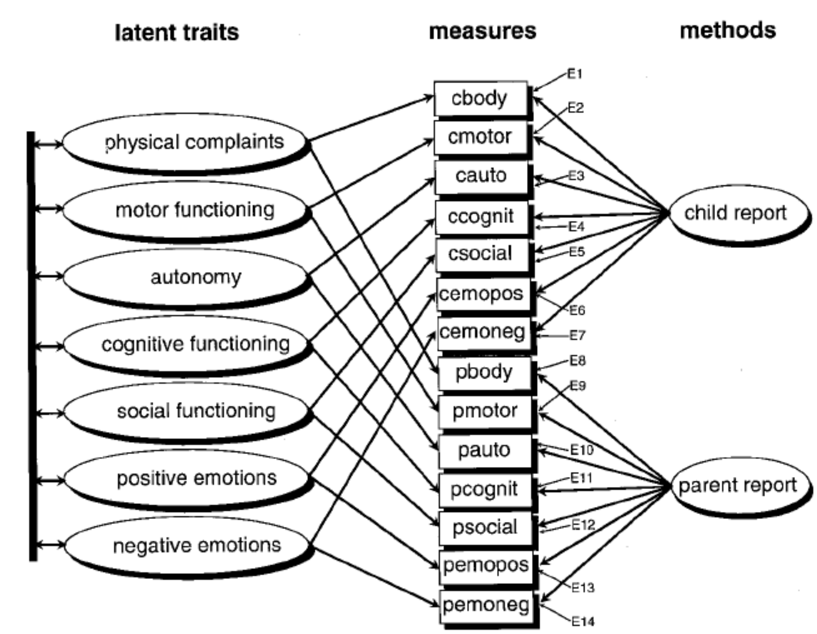


Figure 1. Path diagram for the multitrait-multimethod matrix ($p = 3$, $m = 3$).



Multiple-Group Comparisons (Interaction Effects)

1) Used to study interaction effects
2) SEMs well-suited to study interaction effects because they take measurement error into account
   a) example: suppose you had a four-indicator measure of self-esteem and you hypothesized that it had a different relationship with depression for high SES and low SES adults [a dichotomous variable]. How would you go about testing this hypothesis with regular regression analysis?
      1) first, you would have to *a priori* add the self-esteem indicators
      2) second, you would have to use product terms, which compound problems of reliability.
         a) we know that the $V_{(t1)}=\rho_1 V_{(\rho 1)}$ and $V_{(t2)}=\rho_2 V_{(\rho 2)}$
         b) so, then, the true variance of V1 times V2 (if they are not correlated) = $\rho_1\rho_2 V_{(\rho 1)} V_{(\rho 2)}$
   b) recap:
      1) first, with an SEM it is possible to measure latent variables
      2) second, with an SEM it is possible to perform group comparisons and avoid the compounding of measurement error that is inherent in product terms.

Multiple-Group Comparisons: An Interaction Across Two Groups

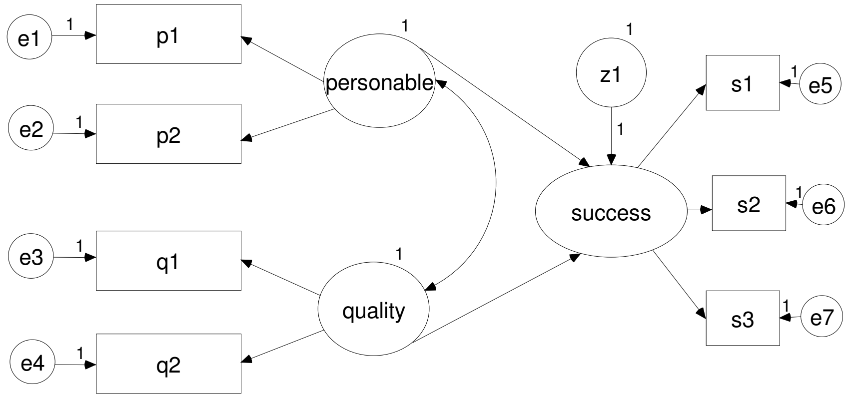
1) Does gender modify the relationship between personableness and the democratic candidate's success in winning a debate?
   a) check to make sure that the measurement models do not differ between groups.

   b) calculate model fit using a multiple group solution with no constraints across groups (except for measurement model)

   c) calculate model fit using a multiple group solution with an across-group constraint imposed to reflect the interaction effect.

   d) calculate the difference in model fit by comparing the fit index for the constrained solution with the fit index for the unconstrained solution.

Example: Ratings of a political debate

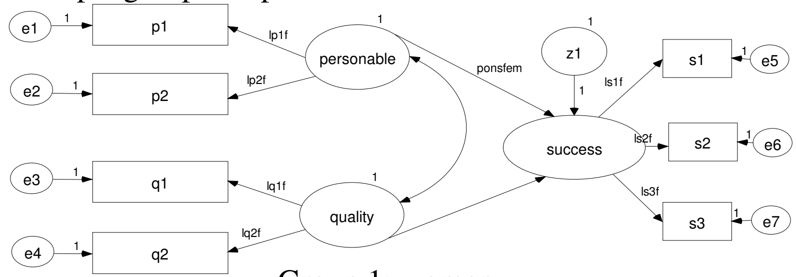


Number of parameters to estimate = 17

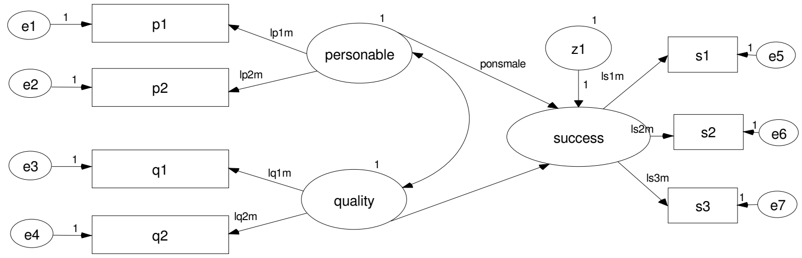
Straight arrows: 9 variances: 7

Curved arrows: 1

Multiple group comparison: Men and Women



Group 1: women



Group 2: men

Number of parameters to estimate: (17*2 = 34)

Multiple group comparison:

Saturated model: No constraints in parameters

Default model: All model parameters differ across groups (males and females)

Model #2: Measurement model, in which cfas are constrained to be equal across groups (males and females)

a) lp1f=lp1m; lp2f=lp2m;
lq1f=lq1m; lq2f=lq2m;
ls1f=ls1m; ls2f=ls2m; ls3f=ls3m

Model #3: Same as Model #2, except that parameter of interest is constrained to be equal across groups

a) additional constraint: ponsmale=ponsfem

Question 1: Measurement Model

Before looking to see if the relationship between "personableness" and "success" is the same for males and females, we first need to makes sure that the latent constructs of "personableness" and "success" are the same.

Question 1: Measurement Models

Females

			Estimate
success	<---	personable	.746
success	<---	quality	.391
p1	<---	personable	.868
p2	<---	personable	.816
q1	<---	quality	.744
q2	<---	quality	.896
s1	<---	success	.973
s2	<---	success	.970
s3	<---	success	.974

Males

			Estimate
success	<---	personable	.523
success	<---	quality	.548
p1	<---	personable	.841
p2	<---	personable	.863
q1	<---	quality	.845
q2	<---	quality	.856
s1	<---	success	.940
s2	<---	success	.929
s3	<---	success	.972

Question 1: Measurement Models

P value

Assuming that the **Default model** model is correct, the probability of getting a discrepancy as large as 15.257 is .851.

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	34	15.257	22	.851	.693
Model Number 2	27	16.792	29	.965	.579

Likelihood Ratio Test (for nested models):
Default: if AIC = 83.26, and s=34, then –2LL=83.26-68=15.26
#2: if AIC=70.79, and s=27, then –2LL=70.79-54=16.79

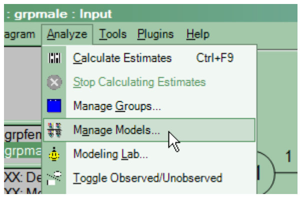
Difference = 1.53~$\chi 2$ with 7 d.f.
So, constrained model is not a significantly worse fit.
Caveat: this is an asymptotic result which needs a large N.

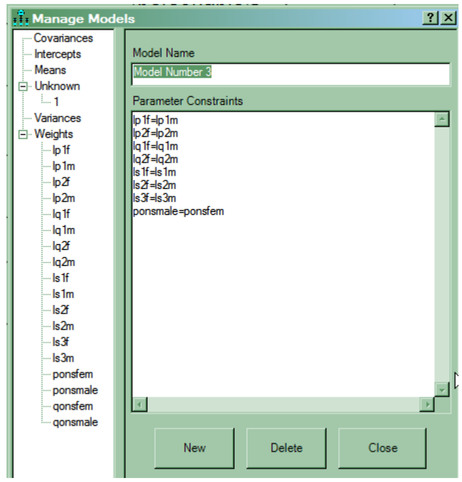
Question 2: Structural Model

OK, so assuming the measurement model is the same across Males and females, now we can ask if the relationship between Personableness and success is the same across gender.

Fit two models (both with equal measurement models)
1) Where ponsfem and ponsmale are both estimated (already done, this was model # 2
2) Where ponsfem and ponsmale are forced to be equal (model 3)

Constraining a Weight





Question 2: Structural Model

Regression Weights: (grpfem - Model Number 2)

			Estimate	S.E.	C.R.	P	Label
success	<---	personable	1.581	.182	8.676	***	ponsfem
success	<---	quality	.841	.149	5.663	***	qonsfem

Regression Weights: (grpmale - Model Number 2)

			Estimate	S.E.	C.R.	P	Label
success	<---	personable	.770	.120	6.411	***	ponsmale
success	<---	quality	.802	.122	6.549	***	qonsmale

Question 2: Structural Model

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	34	15.257	22	.851	.693
Model Number 2	27	16.792	29	.965	.579
Model Number 3	26	41.668	30	.076	1.389
Saturated model	56	.000	0		
Independence model	14	1752.220	42	.000	41.720

Likelihood Ratio Test (for nested models):
#2: if AIC=70.79, and s=27, then –2LL=70.79-54=16.79
#3: if AIC=93.67, and s=26, then –2LL=93.67-52=41.67

Difference = 24.88~$\chi 2$ with 1 d.f.
So, constrained model is significantly worse fit.
Caveat: this is an asymptotic result which needs a large N.

Multiple group comparison for more than two groups

Compare previous model across African-American, Hispanic, and white respondents

Default model: compare measurement model to saturated model

Model 2: Constrain parameter of interest to be the same across all groups

Results:

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	37	54.217	47	0.218	1.154
Model Number 2	35	96.085	49	0.000	1.961
Saturated model	84	0.000	0		
Independence model	21	1966.186	63	0.000	31.209

Three-way interaction across two groups:

For example, is difference in parameter different across male and female democrats in comparison to male and female republicans?

Procedure: Constrain male/female difference in parameters to be equal across political groups, and see if difference is significantly different when this constraint is not included

Three-way interaction across more than two groups:

For example, is difference in parameter different across male and females different for republicans, democrats, and/or independents?

Procedure: Constrain male/female difference in parameters to be equal across all political groups, and see if difference is significantly different when this constraint is not included