结构方程模型及其应用
Structural Equation Model and Its Applications

I 简介
I Introduction

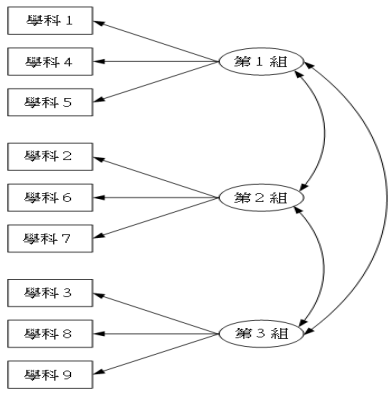
侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

100个推理测验分数
21, 31, 32, 05, 06, 09, 10, 22, 29, 18,
11, 01, 39, 92, 23, 27, 93, 97, 30, 02,
96, 40, 53, 78, 04, 98, 36, 07, 08, 24,
54, 55, 77, 99, 34, 03, 86, 87, 59, 60,
15, 62, 63, 43, 52, 28, 79, 58, 65, 95,
81, 85, 57, 14, 17, 33, 16, 19, 20, 37,
25, 69, 84, 61, 64, 68, 70, 42, 45, 72,
83, 89, 44, 38, 47, 71, 00, 73, 12, 35,
82, 56, 75, 41, 46, 49, 50, 94, 66, 67,
76, 51, 88, 90, 74, 13, 26, 80, 48, 91
均值**Mean=53**，标准差**SD(Std Dev)=15**

100名学生在9个不同学科间的相关系数
(correlation coefficient matrix)

9 个不同学科间观察所得的相关矩阵 S

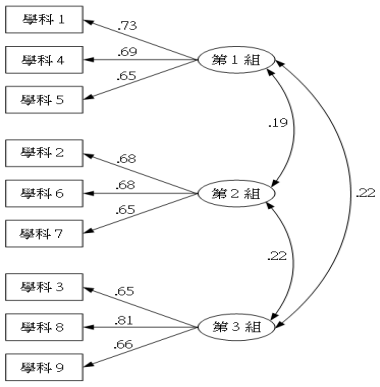
学科	1	2	3	4	5	6	7	8	9
学科 1	1.00								
学科 2	.12	1.00							
学科 3	.08	.08	1.00						
学科 4	.50	.11	.08	1.00					
学科 5	.48	.03	.12	.45	1.00				
学科 6	.07	.46	.15	.08	.11	1.00			
学科 7	.05	.44	.15	.12	.12	.44	1.00		
学科 8	.14	.17	.53	.14	.08	.10	.06	1.00	
学科 9	.16	.05	.43	.10	.06	.08	.10	.54	1.00



再生/隐含矩阵 (reproduced/implied matrix)

依据输入的相关矩阵 S 和模型 M_1 所得的再生矩阵 Σ

学科	1	2	3	4	5	6	7	8	9
学科 1	1.00								
学科 2	.10	1.00							
学科 3	.11	.10	1.00						
学科 4	.50	.09	.10	1.00					
学科 5	.48	.09	.09	.45	1.00				
学科 6	.10	.46	.10	.09	.09	1.00			
学科 7	.09	.44	.09	.09	.08	.44	1.00		
学科 8	.13	.12	.53	.12	.12	.12	.11	1.00	
学科 9	.11	.10	.43	.10	.10	.10	.09	.54	1.00

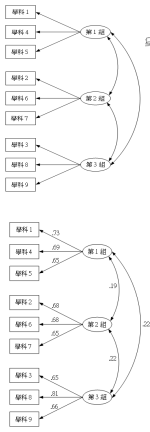


9 个不同学科间观察所得的相关矩阵 S

学科	1	2	3	4	5	6	7	8	9
学科 1	1.00								
学科 2	.12	1.00							
学科 3	.08	.08	1.00						
学科 4	.50	.11	.08	1.00					
学科 5	.48	.03	.12	.45	1.00				
学科 6	.07	.46	.15	.08	.11	1.00			
学科 7	.05	.44	.15	.12	.12	.44	1.00		
学科 8	.14	.17	.53	.14	.08	.10	.06	1.00	
学科 9	.16	.05	.43	.10	.06	.08	.10	.54	1.00

依据输入的相关矩阵 S 和模型 M_1 所得的再生矩阵 Σ

学科	1	2	3	4	5	6	7	8	9
学科 1	1.00								
学科 2	.10	1.00							
学科 3	.11	.10	1.00						
学科 4	.50	.09	.10	1.00					
学科 5	.48	.09	.09	.45	1.00				
学科 6	.10	.46	.10	.09	.09	1.00			
学科 7	.09	.44	.09	.09	.08	.44	1.00		
学科 8	.13	.12	.53	.12	.12	.12	.11	1.00	
学科 9	.11	.10	.43	.10	.10	.10	.09	.54	1.00

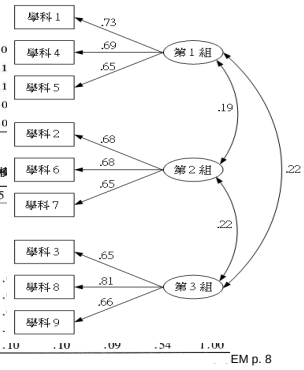


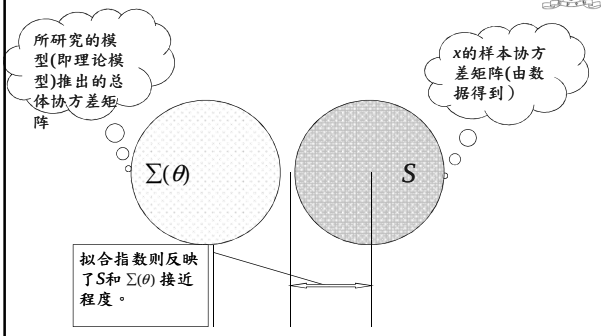
9 个不同学科间观察所得的相关矩阵 S

学科	1	2	3	4	5	6	7	8	9
学科 1	1.00								
学科 2	.12	1.00							
学科 3	.08	.08	1.00						
学科 4	.50	.11	.08	1.00					
学科 5	.48	.03	.12	.45	1.0				
学科 6	.07	.46	.15	.08	.1				
学科 7	.05	.44	.15	.12	.1				
学科 8	.14	.17	.53	.14	.0				
学科 9	.16	.05	.43	.10	.0				

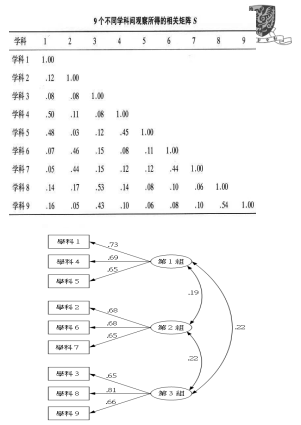
依据输入的相关矩阵 S 和模

学科	1	2	3	4	5
学科 1	1.00				
学科 2	.10	1.00			
学科 3	.11	.10	1.00		
学科 4	.50	.09	.10	1.00	
学科 5	.48	.09	.09	.45	1.
学科 6	.10	.46	.10	.09	.
学科 7	.09	.44	.09	.09	.
学科 8	.13	.12	.53	.12	.
学科 9	.11	.10	.43	.10	.10





- **检查模型的准确性(accuracy)和简洁性(parsimony)**
  - **拟合优度指数（goodness of fit index），简称为拟合指数(fit index): χ^2、NNFI、CFI**
  - **df=[不重复元素non-duplicating elements, p(p+1)/2] – [估计参数 estimated parameters]**
  - **在前面例子 df =9 x 10/2 – 21 = 24**



样本相关（或协方差）矩阵
correlation/covariance matrix S

一个或多个有理据的可能模型 (alternative models)

输入 Input

SEM 程式program (e.g., AMOS, LISREL)

依据 S 及指定模型
找出与 S 相距最小的 Σ

输出 Output

Σ、各路径参数（因子负荷loading、因子相关系数factor correlations等）
各种拟合指数

结构方程模型及其应用
Structural Equation Model and Its Applications

II 探索性与验证性因子分析
II Exploratory vs Confirmatory Factor Analysis

侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

Additional notes on EFA

Exploratory Factor Analyses (EFA)

- In the above analyses, we have a structure in mind to test, this process is called confirmatory factor analysis (CFA)
- It is also possible that we have no "theory" in mind to test, i.e., we have the following research questions:
  - How many cluster of subjects are there? How do these 9 subjects relate to each of these clusters (factors)?
  - Which of these subjects are more closely related/correlated than others?

Using SPSS, run the following program

```
MATRIX DATA variables=ROWTYPE_ v1 to v9.
begin data.
N 100 100 100 100 100 100 100 100 100
SD 1 1 1 1 1 1 1 1 1
CORR 1.00
CORR 0.12 1.00
CORR 0.08 0.08 1.00
CORR 0.50 0.11 0.08 1.00
CORR 0.48 0.03 0.12 0.45 1.00
CORR 0.07 0.46 0.15 0.08 0.11 1.00
CORR 0.05 0.44 0.15 0.12 0.12 0.44 1.00
CORR 0.14 0.17 0.53 0.14 0.08 0.10 0.06 1.00
CORR 0.16 0.05 0.43 0.10 0.06 0.08 0.10 0.54 1.00
END data.
Factor matrix in(corr=*)/plot=eigen.
```


The output:

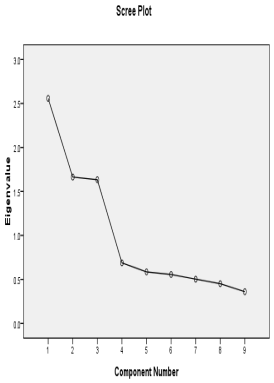
Principal Components Analysis

Eigenvalues and Eigenvectors

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
EV	2.56	1.66	1.63	0.69	0.59	0.56	0.50	0.45	0.36
% Var	28.4	18.5	18.2	7.6	6.5	6.2	5.6	5.0	4.0
Cum%	28.4	46.9	65.1	72.7	79.2	85.4	91.0	96.0	100.

3 rules to determine number of factors

- EV(eigenvalue特征值) ≧ 1
- scree test碎石: greatest change in slope
- meaningful dimensions



Assume 3 factors, we run the following program and obtain further information

```
MATRIX DATA variables=ROWTYPE_ v1 to v9.
begin data.
N 100 100 100 100 100 100 100 100 100
SD 1 1 1 1 1 1 1 1 1
CORR 1.00
CORR 0.12 1.00
CORR 0.08 0.08 1.00
CORR 0.50 0.11 0.08 1.00
CORR 0.48 0.03 0.12 0.45 1.00
CORR 0.07 0.46 0.15 0.08 0.11 1.00
CORR 0.05 0.44 0.15 0.12 0.12 0.44 1.00
CORR 0.14 0.17 0.53 0.14 0.08 0.10 0.06 1.00
CORR 0.16 0.05 0.43 0.10 0.06 0.08 0.10 0.54 1.00
END data.
Factor matrix in(corr=*)/criteria=fac(3)/rotate=varimax.
Factor matrix in(corr=*)/criteria=fac(3)/rotate=promax.
```


factors are assumed to be uncorrelated正交

The Output:

Varimax-Rotated Factor Loadings

	Factor 1	Factor 2	Factor 3	Unique Var
VAR 1	0.10	0.82	0.03	0.68
VAR 2	0.06	0.05	0.79	0.64
VAR 3	0.79	0.04	0.11	0.63
VAR 4	0.06	0.80	0.08	0.65
VAR 5	0.03	0.79	0.06	0.63
VAR 6	0.07	0.05	0.79	0.64
VAR 7	0.06	0.07	0.78	0.62
VAR 8	0.85	0.08	0.07	0.73
VAR 9	0.80	0.07	0.01	0.65

Promax-Rotated Factor Loadings

Factors allowed to be correlated斜交

	Factor 1	Factor 2	Factor 3	Unique Var
VAR 1	0.04	0.82	-0.03	0.68
VAR 2	0.00	0.00	0.80	0.64
VAR 3	0.79	-0.03	0.05	0.63
VAR 4	0.00	0.80	0.03	0.65
VAR 5	-0.03	0.80	0.01	0.63
VAR 6	0.10	-0.01	0.80	0.64
VAR 7	-0.01	0.02	0.79	0.62
VAR 8	0.85	0.01	0.00	0.73
VAR 9	0.81	0.01	-0.05	0.65

Factor Correlations

	Factor 1	Factor 2	Factor 3
Factor 1	1.00		
Factor 2	0.16	1.00	
Factor 3	0.16	0.14	1.00

EFA (exploratory FA)	CFA (confirmatory FA)
No specific idea on how variable are related	Have some guess (hypotheses) on relations among variables (e.g., Variables 2, 6, 7 should load on Factor 3)
determine number of factors using Eigenvalue (EV ≧ 1 or scree test)	know beforehand the number of factors
each item loaded on ALL factors, though some loadings are small	Items loaded on targeted factors only

结构方程模型及其应用

Structural Equation Model and Its Applications

III 原理

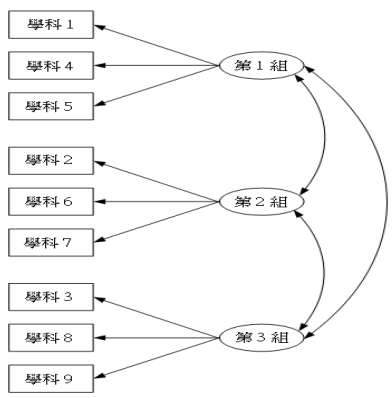
III Principle

侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

100名学生在9个不同学科间的相关系数
(correlation coefficient matrix)

9 个不同学科间观察所得的相关矩阵 S

学科	1	2	3	4	5	6	7	8	9
学科 1	1.00								
学科 2	.12	1.00							
学科 3	.08	.08	1.00						
学科 4	.50	.11	.08	1.00					
学科 5	.48	.03	.12	.45	1.00				
学科 6	.07	.46	.15	.08	.11	1.00			
学科 7	.05	.44	.15	.12	.12	.44	1.00		
学科 8	.14	.17	.53	.14	.08	.10	.06	1.00	
学科 9	.16	.05	.43	.10	.06	.08	.10	.54	1.00

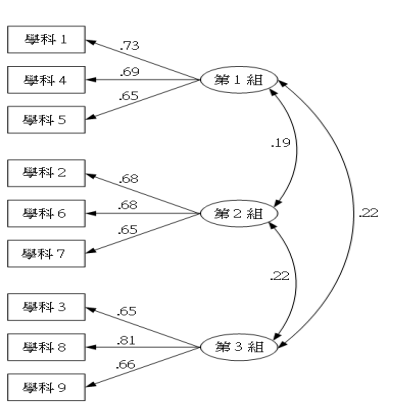


9 个不同学科间观察所得的相关矩阵 S

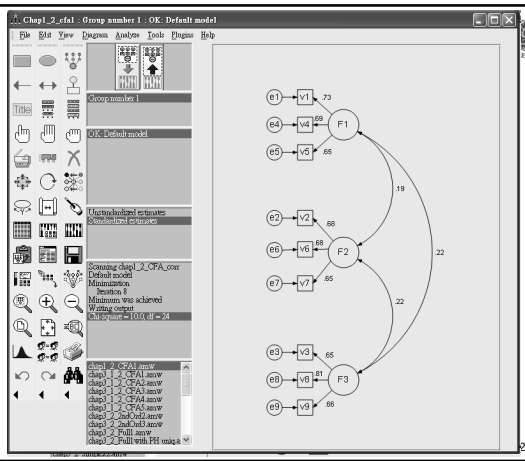
学科	1	2	3	4	5	6	7	8	9
学科 1	1.00								
学科 2	.12	1.00							
学科 3	.08	.08	1.00						
学科 4	.50	.11	.08	1.00					
学科 5	.48	.03	.12	.45	1.00				
学科 6	.07	.46	.15	.08	.11	1.00			
学科 7	.05	.44	.15	.12	.12	.44	1.00		
学科 8	.14	.17	.53	.14	.08	.10	.06	1.00	
学科 9	.16	.05	.43	.10	.06	.08	.10	.54	1.00

依据输入的相关矩阵 S 和模型 M_1 所得的再生矩阵 Σ

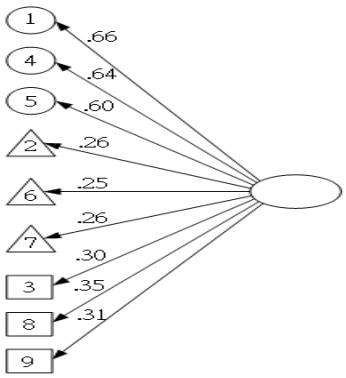
学科	1	2	3	4	5	6	7	8	9
学科 1	1.00								
学科 2	.10	1.00							
学科 3	.11	.10	1.00						
学科 4	.50	.09	.10	1.00					
学科 5	.48	.09	.09	.45	1.00				
学科 6	.10	.46	.10	.09	.09	1.00			
学科 7	.09	.44	.09	.09	.08	.44	1.00		
学科 8	.13	.12	.53	.12	.12	.12	.11	1.00	
学科 9	.11	.10	.43	.10	.10	.10	.09	.54	1.00



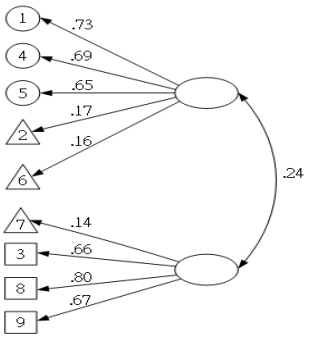
KT HAU SEM p. 25



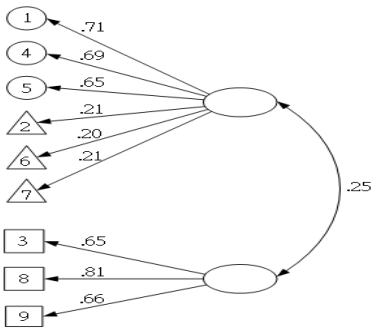
KT HAU SEM p. 26



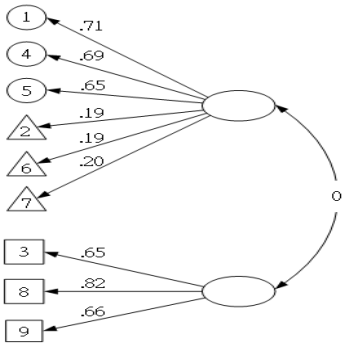
KT HAU SEM p. 27



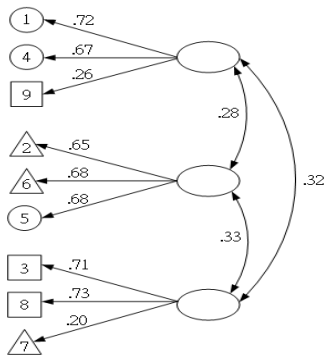
KT HAU SEM p. 28

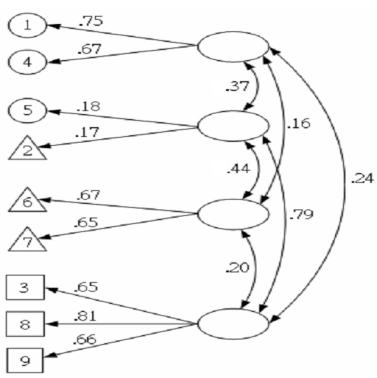


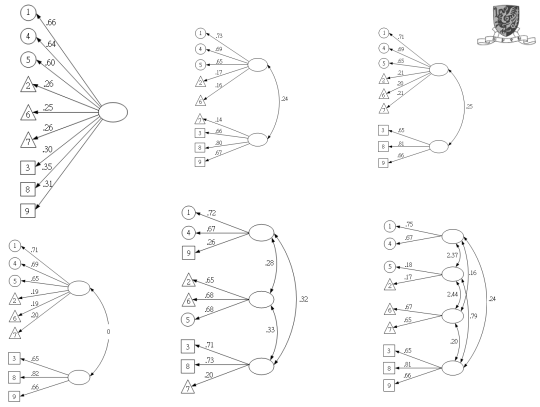
KT HAU SEM p. 29



KT HAU SEM p. 30







模型	df	χ^2	NNFI	CFI	(no. of estimated parameters) 需要估计的参数个数
M1	24	40	.973	.982	21 = 9 Load + 9 Uniq + 3 Corr
M2	27	503	.294	.471	18 = 9 Load + 9 Uniq
M3	26	255	.647	.745	19 = 9 Load + 9 Uniq + 1 Corr
M4	26	249	.656	.752	19 = 9 Load + 9 Uniq + 1 Corr
M5	27	263	.649	.727	18 = 9 Load + 9 Uniq
M6	24	422	.337	.558	21 = 9 Load + 9 Uniq + 3 Corr
M7	21	113	.826	.898	24 = 9 Load + 9 Uniq + 6 Corr

模型比较 (Model Comparison)

- 自由度(df), 拟合程度 (fit), 不能保证最好，可能存在更简洁(parsimonious) 又拟合(fit)得很好的模型
- **输入Input:**
  - 相关（或协方差）矩阵correlation/covariance matrix S
  - 一个或多个有理据的可能模型 (alternative models)
- **输出Output:**
  - 既符合某指定模型，又与 S 差异最小的矩阵 Σ
  - 估计各路径参数parameter（因子负荷loading、因子相关系数factor correlations等）。
  - 计算出各种拟合指数(goodness of fit indexes)

样本相关（或协方差）矩阵
correlation/covariance matrix S

一个或多个有理据的可能模型 (alternative models)

输入
Input

SEM 程式program
(e.g., LISREL)

依据 S 及指定模型
找出与 S 相距最小的 Σ

输出
Output

Σ、各路径参数（因子负荷loading、因子相关系数factor correlations等）
各种拟合指数

结构方程模型的重要性

- Structural Equation Model，SEM
- Covariance Structure Modeling，CSM
- AMOS = Analysis of Moment Structures
- LInear Structural RELationship ， LISREL
  - (EQS, Mplus, etc.)

结构方程模型的结构

- 测量模型 (measurement model)

$$\boldsymbol{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

$$\boldsymbol{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

$\boldsymbol{x}$ ——外源指标exogenous（如6项社经指标）

$\boldsymbol{y}$—内生指标endogenous（如语、数、英成绩）

$\boldsymbol{\Lambda}_x$，$\boldsymbol{\Lambda}_y$ —因子负荷矩阵 (loading)

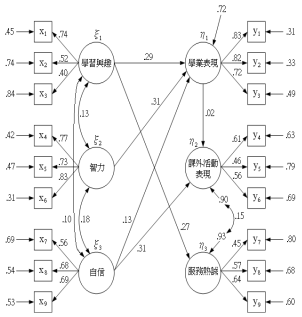
$\boldsymbol{\delta}$，$\boldsymbol{\varepsilon}$ —误差项 (uniqueness, measurement errors)

- 结构模型 (structural model)

$$\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$$

结构方程模型的优点

- 同时处理多个因变量(many dependent variables)
- 同时估计因子结构 factor structure和因子关系



- 容许自变量 independent variable和因变量 dependent variable含测量误差 measurement error [传统方法（如回归 regression）假设自变量 independent variable没有误差]

英文			中文		
观察得分 observed score X	**真分数** true score T_X	**误差** error e	**观察得分** observed score Y	**真分数** true score T_y	**误差** error e
8	7	+1	5	3	+2
5	6	-1	6	7	-1
7	5	+2	9	7	+2
9	8	+1	5	8	-3
.	.	.	.	.	.

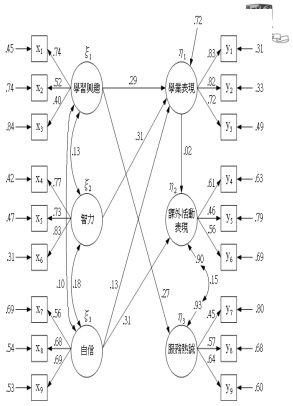
$X = T_X + e$ $\quad Y = T_y + e$

if $r(X, Y) = 0.5$

$r(T_X, T_y) = 0.5 / [(r_{Xt\text{-}t})(r_{Yt\text{-}t})]^{1/2}$

$= 0.71$ (assume $r_{t\text{-}t} = 0.7$)

- 容许更大弹性的测量模型
- 估计整个模型的拟合程度 model fit[用以比较不同模型]
- SEM包括:回归分析 regression、因子分析(验证性因子分析CFA、 探索性因子分析EFA)、 t 检验t-test、方差分析ANOVA、比较各组因子均值group mean comparison、交互作用模型 interaction、实验设计expt design



结构方程模型及其应用

Structural Equation Model and Its Applications

IV 验证性因子分析

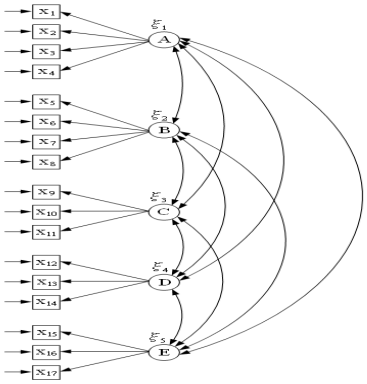
IV Confirmatory Factor Analysis (Run 1)

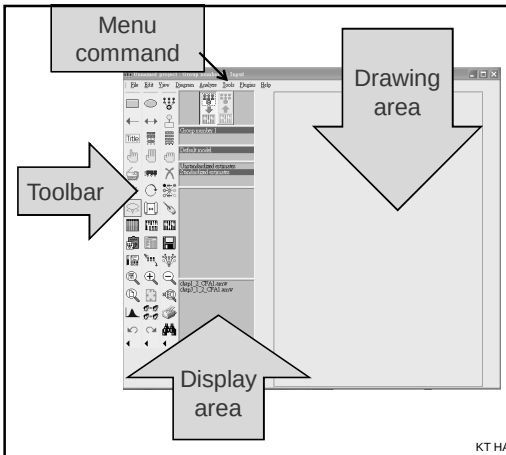
侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

验证性因子分析
Confirmatory Factor Analysis, CFA

(i)17个题目:
学习态度及取向
(ii) A、B、C、D、E
(iii) 4、4、3、3、3题
(iv) 350个学生

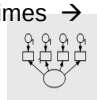
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	.1																
2	.24	.1															
3	.38	.35	.1														
4	.02	.03	.04	.1													
5	.15	.19	.14	.02	.1												
6	.17	.15	.20	.01	.42	.1											
7	.20	.13	.12	.00	.40	.21	.1										
8	.32	.32	.21	.03	.10	.10	.07	.1									
9	.10	.17	.12	.02	.15	.18	.23	.13	.1								
10	.14	.16	.15	.03	.14	.19	.18	.18	.37	.1							
11	.14	.15	.19	.01	.18	.30	.13	.08	.38	.38	.1						
12	.18	.16	.24	.02	.14	.21	.21	.22	.06	.23	.18	.1					
13	.19	.20	.15	.01	.14	.24	.09	.24	.15	.21	.21	.45	.1				
14	.18	.21	.18	.03	.25	.18	.18	.18	.22	.12	.24	.28	.35	.1			
15	.08	.18	.16	.01	.22	.20	.22	.12	.12	.16	.21	.25	.20	.26	.1		
16	.12	.16	.25	.02	.15	.12	.20	.14	.17	.20	.14	.20	.15	.20	.50	.1	
17	.20	.16	.18	.04	.25	.14	.21	.17	.21	.21	.23	.15	.21	.22	.29	.41	.1





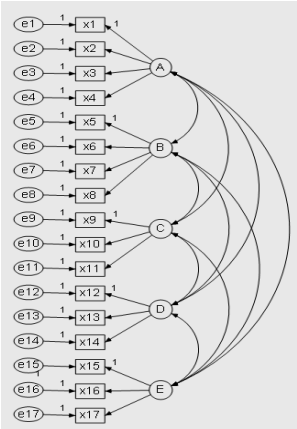
To draw the CFA diagram

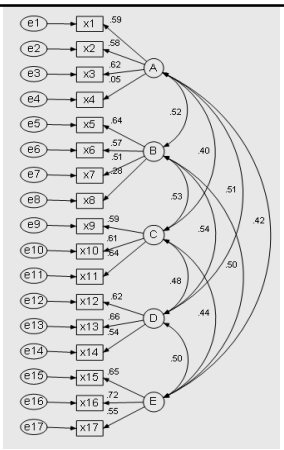
1. Click on toolbar → click/drop on drawing area, click 4 times →



2. Click , click on the factor circle on the drawing area 3 times, to rotate

3. Repeat, add paths, etc. to become:





To adjust parameters displayed:

Results

Analysis Summary

Date and Time

Date: 2009年5月6日

Time: 下午 07:36:56

Title Chap 3 CFA

Notes for Group (Group number 1)

The model is recursive

Sample size = 350

Variable counts (Group number 1)

Number of variables in your model: 39

Number of observed variables: 17

Number of unobserved variables: 22

Number of exogenous variables: 22

Number of endogenous variables: 17

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	22	0	0	0	0	22
Labeled	0	0	0	0	0	0
Unlabeled	12	10	22	0	0	44
Total	34	10	22	0	0	66

Computation of degrees of freedom (Default model)

Number of distinct sample moments: 153

Number of distinct parameters to be estimated: 44

Degrees of freedom (153 - 44): 109

Result (Default model)

Minimum was achieved

Chi-square = 194.570

Degrees of freedom = 109

Probability level = .000

参数 (parameter)

Standard Error SE

Maximum Likelihood Estimates -- Regression Weights:

t-value=参数/SE=.980/.145=6.764

			Estimate	S.E.	C.R.	P
x1	<---	A	1.000			
x2	<---	A	.980	.145	6.764	***
x3	<---	A	1.050	.150	7.007	***
x4	<---	A	.089	.110	.807	.419
x5	<---	B	1.000			
x6	<---	B	.888	.124	7.157	***
x7	<---	B	.794	.115	6.881	***
x8	<---	B	.436	.126	3.469	***
x9	<---	C	1.000			
x10	<---	C	1.045	.145	7.205	***
x11	<---	C	1.098	.152	7.229	***
x12	<---	D	1.000			
x13	<---	D	1.057	.132	8.032	***
x14	<---	D	.875	.133	6.557	***
x15	<---	E	1.000			
x16	<---	E	1.100	.127	8.677	***
x17	<---	E	.844	.119	7.087	***

Standardized Regression Weights:

			Estimate
x1	<---	A	.592
x2	<---	A	.580
x3	<---	A	.621
x4	<---	A	.053
x5	<---	B	.642
x6	<---	B	.571
x7	<---	B	.510
x8	<---	B	.280
x9	<---	C	.585
x10	<---	C	.612
x11	<---	C	.643
x12	<---	D	.621
x13	<---	D	.656
x14	<---	D	.543
x15	<---	E	.654
x16	<---	E	.719
x17	<---	E	.552

Covariances: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P
A	<-->	B	.199	.040	4.934	***
A	<-->	C	.139	.034	4.102	***
A	<-->	D	.187	.039	4.798	***
A	<-->	E	.163	.036	4.518	***
B	<-->	C	.198	.039	5.096	***
B	<-->	D	.215	.041	5.276	***
B	<-->	E	.208	.042	4.895	***
C	<-->	D	.174	.036	4.833	***
C	<-->	E	.169	.037	4.573	***
D	<-->	E	.201	.041	4.845	***

Correlations:

			Estimate
A	<-->	B	.525
A	<-->	C	.402
A	<-->	D	.510
A	<-->	E	.423
B	<-->	C	.527
B	<-->	D	.541
B	<-->	E	.496
C	<-->	D	.481
C	<-->	E	.442
D	<-->	E	.496

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
x17	.305
x16	.518
x15	.428
x14	.295
x13	.430
x12	.385
x11	.413
x10	.374
x9	.343
x8	.078
x7	.260
x6	.325
x5	.412
x4	.003
x3	.386
x2	.336
x1	.350

Modification Indices -- Covariances:

			M.I.	Par Change
e15	<-->	e17	6.18	-.10
e10	<-->	e14	5.65	-.10
e9	<-->	e12	8.63	-.12
e8	<-->	D	6.57	.09
e8	<-->	B	16.58	-.15
e8	<-->	A	24.86	.17
e7	<-->	e13	5.77	-.10
e7	<-->	e9	4.33	.09
e6	<-->	e11	8.19	.12
e6	<-->	e7	5.92	-.10
e5	<-->	e8	5.34	-.10
e5	<-->	e7	5.86	.10
e3	<-->	e16	5.61	.09
e2	<-->	e8	10.38	.15
e1	<-->	e8	10.40	.15

Regression Weights:

			M.I.	Par Change
x17	<---	C	4.38	.20
x17	<---	x5	4.28	.10
x16	<---	x11	4.25	-.09
x16	<---	x6	4.17	-.09
x16	<---	x5	4.05	-.09
x14	<---	x9	4.06	.10
x14	<---	x5	4.37	.10
x13	<---	x7	5.82	.11
x12	<---	x9	6.25	-.12
x11	<---	x6	5.22	.11
x9	<---	x12	6.68	-.12
x8	<---	D	10.10	.32
x8	<---	A	20.61	.49
x8	<---	x13	8.72	.15
x8	<---	x12	6.25	.13
x8	<---	x3	5.88	.13
x8	<---	x2	20.83	.24
x8	<---	x1	20.90	.24
x6	<---	x11	6.40	.12
x5	<---	x12	4.14	-.09
x5	<---	x8	4.83	-.10
x2	<---	x8	9.85	.15
x1	<---	x8	9.54	.15

Maximum Modification Index is 20.61 for Element x8 ← A

修正指数:该参数由固定改为自由估计， 会减少的数值

Model Fit Summary 拟合优度统计量

Sat. nPar = p x (p+1)/2
Ind. nPar = p

0.56 x (N-1) = chi
0.56x349=194.57

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	44	194.57	109	.00	1.79
Saturated model	153	.00	0		
Independence model	17	1104.79	136	.00	8.12

Model	RMR	GFI	AGFI	PGFI
Default model	.05	.94	.92	.67
Saturated model	.00	1.00		
Independence model	.19	.60	.55	.53

Oldest (LISREL) indexes:
RMR, GFI, AGFI, (SRMR)

TLI = NNFI

CFI = RNI

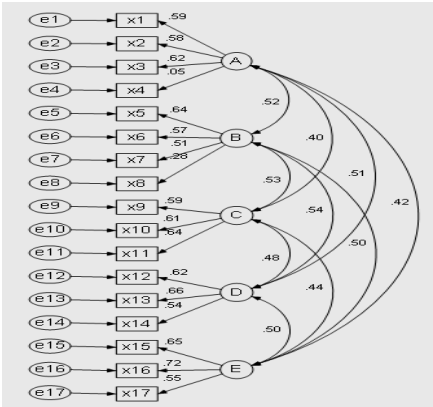
Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.82	.78	.91	.89	.91
Saturated model	1.00		1.00		1.00
Independence model	.00	.00	.00	.00	.00

Model	PRATIO	PNFI	PCFI
Default model	.80	.66	.73
Saturated model	.00	.00	.00
Independence model	1.00	.00	.00

pRatio = df (hypo model)/df (null model)
= (153 - 44) / (153-17)
= 109 / 136 = 0.801

PCFI = pRatio x CFI
= .80 x .91
= .73

Completely Standardized Solution

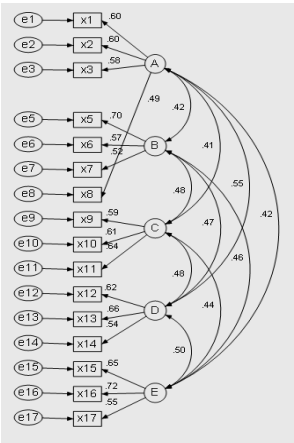


U Amos p. 61

结果解释

- x4在A的负荷loading很小 (0.05)，但在其他因子的修正指数（MI, modification index）也不高 (0.00, 0.02, 0.00, 0.02 in B,C,D,E)
  - 不从属A，也不归属其他因子
- Q8在B的负荷不高（0.28），但在A的MI是20.6，可能归属A
- 因子间相关很高 (0.40 至 0.54)
- 模型拟合fit相当好：χ^2(109) =194.57，RMSEA＝0.047, TLI = .89, CFI= .91。
- 仔细检查题目内容后，删去Q4, Q8归入A

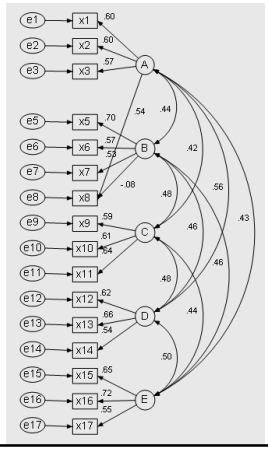
KT HAU SEM p. 62



模型修正 M_B

- Q8归属A，因子负荷很高（.49），
- χ^2(94) =149.51，RMSEA＝.041，TLI＝.93，CFI = .94。

KT HAU SEM p. 63



模型修正 M_C

- χ^2 (93)= 148.61, RMSEA＝.041, TLI = .93, CFI = .94。
- Q8在A负荷为 .54，在B负荷为 -.08
- 因为概念上Q8应与B成正相关，故不合理。而且这负荷相对低，所以我们选择 M_B
- 通常，每题只归属一个因子

KT HAU SEM p. 64

修正modification前后模型的拟合指数比较

模型	df	χ^2	RMSEA	NNFI	CFI	註
M-A	109	195	.046	.89	.91	原模型
M-B	94	150	.041	.93	.94	删Q4,Q8-A
M-C	93	149	.041	.93	.94	删Q4,Q8-A,B
MB-2	99	152	.039	.94	.95	2阶因子

KT HAU SEM p. 65

结构方程模型及其应用

Structural Equation Model and Its Applications

IV 验证性因子分析

IV Confirmatory Factor Analysis (Run 2)

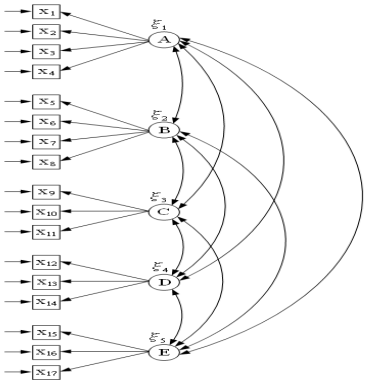
侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

KT HAU SEM p. 66

验证性因子分析
Confirmatory Factor Analysis, CFA

(i)17个题目:
学习态度及取向
(ii) A、B、C、D、E
(iii) 4、4、3、3、3题
(iv) 350个学生

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1																
2	.24	1															
3	.38	.35	1														
4	.02	.03	.04	1													
5	.15	.19	.14	.02	1												
6	.17	.15	.20	.01	.42	1											
7	.20	.13	.12	.00	.40	.21	1										
8	.32	.32	.21	.03	.10	.10	.07	1									
9	.10	.17	.12	.02	.15	.18	.23	.13	1								
10	.14	.16	.15	.03	.14	.19	.18	.18	.37	1							
11	.14	.15	.19	.01	.18	.30	.13	.08	.38	.38	1						
12	.18	.16	.24	.02	.14	.21	.21	.22	.06	.23	.18	1					
13	.19	.20	.15	.01	.14	.24	.09	.24	.15	.21	.21	.45	1				
14	.18	.21	.18	.03	.25	.18	.18	.18	.22	.12	.24	.28	.35	1			
15	.08	.18	.16	.01	.22	.20	.22	.12	.12	.16	.21	.25	.20	.26	1		
16	.12	.16	.25	.02	.15	.12	.20	.14	.17	.20	.14	.20	.15	.20	.50	1	
17	.20	.16	.18	.04	.25	.14	.21	.17	.21	.21	.23	.15	.21	.22	.29	.41	1



In SPSS, prepare the raw data or corr matrix

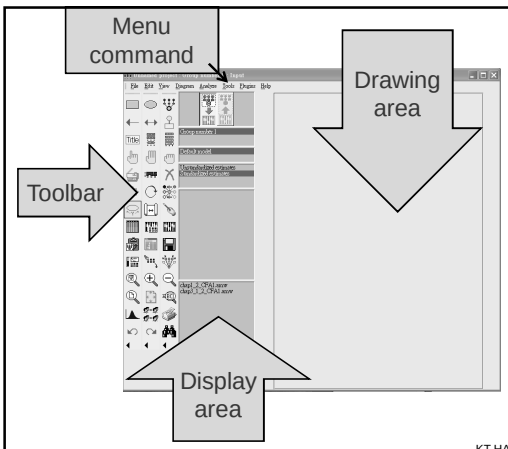
```
MATRIX DATA variables=ROWTYPE_ v1 to v17.
begin data.
N 350 350 350 350 350 350 350 350 350 350 350
   350 350 350 350 350 350
SD 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
CORR 1
CORR .34 1
CORR .38 .35 1
CORR .02 .03 .04  1
CORR .15 .19 .14 .02  1
CORR .17 .15 .20 .01 .42  1
CORR .20 .13 .12 .00 .40 .21  1
CORR .32 .32 .21 .03 .10 .10 .07 1
CORR .10 .17 .12 .02 .15 .18 .23 .13  1
CORR .14 .16 .15 .03 .14 .19 .18 .18 .37  1
CORR .14 .15 .19 .01 .18 .30 .13 .08 .38 .38  1
CORR .18 .16 .24 .02 .14 .21 .21 .22 .06 .23 .18  1
CORR .19 .20 .15 .01 .14 .24 .09 .24 .15
   .21 .21 .45  1
CORR .18 .21 .18 .03 .25 .18 .18 .18 .22
   .12 .24 .28 .35  1
CORR .08 .18 .16 .01 .22 .20 .22 .12 .12
   .16 .21 .25 .20 .26  1
CORR .12 .16 .25 .02 .15 .12 .20 .14 .17
   .20 .14 .20 .15 .20 .50  1
CORR .20 .16 .18 .04 .25 .14 .21 .17 .21
   .21 .23 .15 .21 .22 .29 .41  1
END data.save outfile =
   'D:\AMOS\chap3CFA.sav'.
```


In Amos: to start

Ref: Amos xx.x User's Guide and Programming Reference Guide in X:\Program Files\Amos xx.x\Documentation\

Either:

- start → program → Amos xx.x → Amos graphics
- double click *.amw
- drag/drop *.amw from explorer to Amos graphics
- Start → programs → Amos xx.x → view path diagrams → double-click *.amw
- in SPSS → analyze → Amos xx.x

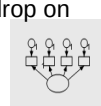


draw observed	draw unobserved	draw latent/add ind
draw path	draw covariance	add unique variable
figure caption	list var in model	list var in data set
select one object	select all object	deselect all objects
duplicate objects	move objects	erase objects
change shape obj	rotate indicators	reflect indicators
move parm. Values	reposition path diag	touch up variable
select data files	analysis properties	calculate estimates
cc path diag/clip bd	view text	save current diag
Object properties	drag prop obj →obj	preserve symmetries
zoom in select area	View a smaller area	View a larger area
Show entire page	resize diag fit page	examine with a loupe
Bayesian	multiple-gp analysis	Print selected graph
undo prev change	undo previous undo	specification search
toolbar option	toolbar option	toolbar option

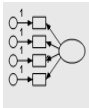
View input path diagram (model specification)	View output path diagram
Group number	
Model selection	
Unstandardized /standardized estimates	
Interaction process	
*.amw files	

To draw the CFA diagram

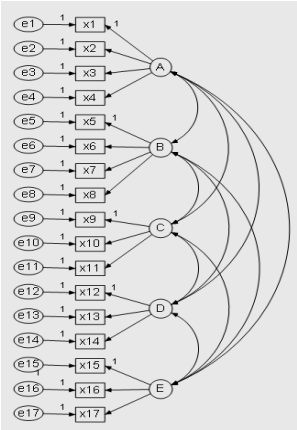
1. Click on toolbar → click/drop on drawing area, click 4 times →



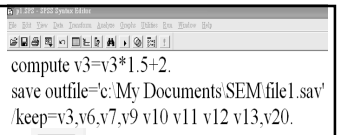
2. Click , click on the factor circle on the drawing area 3 times, to rotate



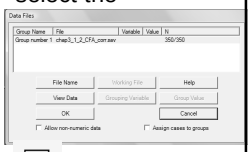
3. Repeat, add paths, etc. to become:



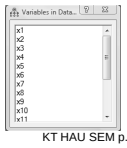
1. Generate *.sav file

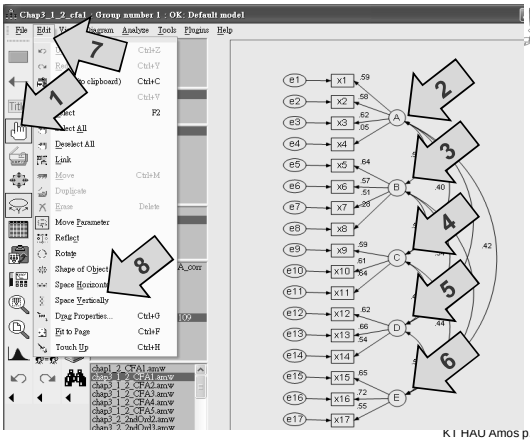


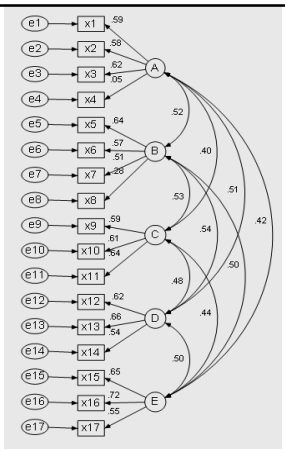
2. Click "Select data files" select the *.sav file



3. Click "list variables in data set" drag and drop the variable to the box/circle in the diagram, right click to add names, then



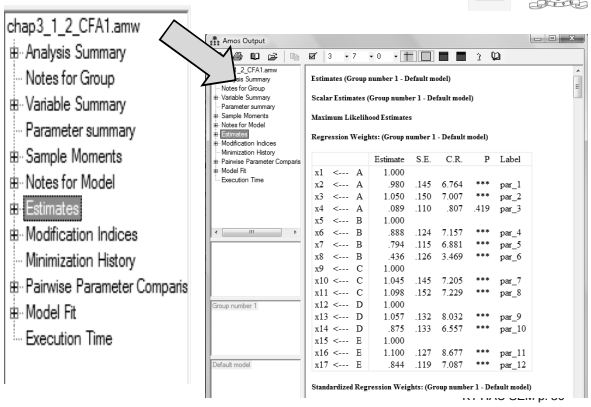




To adjust parameters displayed:

- **什么情况下固定 (fixed, FI)？**
  - 两个变量（指标或因子）间没有关系，将元素固定为0
    - 例如，不从属，将因子负荷固定为0。又如，因子和因子没有相关，↔ 固定为0。
  - 需要设定因子的度量单位(set metric/scale)
    - 因子没有单位 (metric)，无法计算。
    - 一种将所有因子的方差固定为1(或其他常数)，简称为固定方差法 (fixed variance method)
    - 一种是在每个因子中选择一个负荷固定为1(或其他常数)，简称为固定负荷法(fixed loading), as in Amos for CFA
- **什么情况下设定为自由(free,FR)**:所有需要估计的参数

To see results: Click “view text”



Results

Analysis Summary

Date and Time

Date: 2009年5月6日

Time: 下午 07:36:56

Title Chap 3 CFA

Notes for Group (Group number 1)

The model is recursive

Sample size = 350

Variable Summary (Group number 1)

Your model contains the following variables (Group number 1)

Observed, endogenous variables

x1
x2
x3
x4
x5
x6
x7
x8
x9
x10
x11
x12
x13
x14
x15
x16
x17

Unobserved, exogenous variables

A
e1
e2
e3
e4
B
e5
e6
e7
e8
C
e9
e10
e11
D
e12
e13
e14
E
e15
e16
e17

Variable counts (Group number 1)

Number of variables in your model: 39

Number of observed variables: 17

Number of unobserved variables: 22

Number of exogenous variables: 22

Number of endogenous variables: 17

Parameter summary (Group number 1)

	Weights	Covariances	Variances	Means	Intercepts	Total
Fixed	22	0	0	0	0	22
Labeled	0	0	0	0	0	0
Unlabeled	12	10	22	0	0	44
Total	34	10	22	0	0	66

Computation of degrees of freedom (Default model)

Number of distinct sample moments: 153

Number of distinct parameters to be estimated: 44

Degrees of freedom (153 - 44): 109

Result (Default model)

Minimum was achieved

Chi-square = 194.570

Degrees of freedom = 109

Probability level = .000

参数 (parameter) Standard Error SE

Maximum Likelihood Estimates -- Regression Weights:

t-value=参数/SE=.980/.145=6.764

			Estimate	S.E.	C.R.	P
x1	<---	A	1.000			
x2	<---	A	.980	.145	6.764	***
x3	<---	A	1.050	.150	7.007	***
x4	<---	A	.089	.110	.807	.419
x5	<---	B	1.000			
x6	<---	B	.888	.124	7.157	***
x7	<---	B	.794	.115	6.881	***
x8	<---	B	.436	.126	3.469	***
x9	<---	C	1.000			
x10	<---	C	1.045	.145	7.205	***
x11	<---	C	1.098	.152	7.229	***
x12	<---	D	1.000			
x13	<---	D	1.057	.132	8.032	***
x14	<---	D	.875	.133	6.557	***
x15	<---	E	1.000			
x16	<---	E	1.100	.127	8.677	***
x17	<---	E	.844	.119	7.087	***

Standardized Regression Weights:

			Estimate
x1	<---	A	.592
x2	<---	A	.580
x3	<---	A	.621
x4	<---	A	.053
x5	<---	B	.642
x6	<---	B	.571
x7	<---	B	.510
x8	<---	B	.280
x9	<---	C	.585
x10	<---	C	.612
x11	<---	C	.643
x12	<---	D	.621
x13	<---	D	.656
x14	<---	D	.543
x15	<---	E	.654
x16	<---	E	.719
x17	<---	E	.552

Covariances: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P
A	<-->	B	.199	.040	4.934	***
A	<-->	C	.139	.034	4.102	***
A	<-->	D	.187	.039	4.798	***
A	<-->	E	.163	.036	4.518	***
B	<-->	C	.198	.039	5.096	***
B	<-->	D	.215	.041	5.276	***
B	<-->	E	.208	.042	4.895	***
C	<-->	D	.174	.036	4.833	***
C	<-->	E	.169	.037	4.573	***
D	<-->	E	.201	.041	4.845	***

Correlations: (Group number 1 - Default model)

			Estimate
A	<-->	B	.525
A	<-->	C	.402
A	<-->	D	.510
A	<-->	E	.423
B	<-->	C	.527
B	<-->	D	.541
B	<-->	E	.496
C	<-->	D	.481
C	<-->	E	.442
D	<-->	E	.496

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
A	.349	.074	4.725	***
B	.411	.086	4.796	***
C	.342	.072	4.774	***
D	.384	.075	5.098	***
E	.427	.077	5.562	***
e1	.648	.068	9.585	***
e2	.662	.068	9.808	***
e3	.612	.068	8.949	***
e4	.994	.075	13.192	***
e5	.586	.075	7.838	***
e6	.673	.067	10.034	***
e7	.738	.068	10.831	***
e8	.919	.074	12.398	***
e9	.656	.066	9.942	***
e10	.624	.066	9.397	***
e11	.585	.067	8.736	***
e12	.613	.065	9.363	***
e13	.568	.067	8.508	***
e14	.703	.067	10.512	***
e15	.570	.063	9.112	***
e16	.481	.065	7.382	***
e17	.693	.065	10.694	***

Squared Multiple Correlations:

	Estimate
x17	.305
x16	.518
x15	.428
x14	.295
x13	.430
x12	.385
x11	.413
x10	.374
x9	.343
x8	.078
x7	.260
x6	.325
x5	.412
x4	.003
x3	.386
x2	.336
x1	.350

Modification Indices -- Covariances:

		M.I.	Par Change
e15	<--> e17	6.18	-.10
e10	<--> e14	5.65	-.10
e9	<--> e12	8.63	-.12
e8	<--> D	6.57	.09
e8	<--> B	16.58	-.15
e8	<--> A	24.86	.17
e7	<--> e13	5.77	-.10
e7	<--> e9	4.33	.09
e6	<--> e11	8.19	.12
e6	<--> e7	5.92	-.10
e5	<--> e8	5.34	-.10
e5	<--> e7	5.86	.10
e3	<--> e16	5.61	.09
e2	<--> e8	10.38	.15
e1	<--> e8	10.40	.15

Regression Weights:

			M.I.	Par Change
x17	<---	C	4.38	.20
x17	<---	x5	4.28	.10
x16	<---	x11	4.25	-.09
x16	<---	x6	4.17	-.09
x16	<---	x5	4.05	-.09
x14	<---	x9	4.06	.10
x14	<---	x5	4.37	.10
x13	<---	x7	5.82	-.11
x12	<---	x9	6.25	-.12
x11	<---	x6	5.22	.11
x9	<---	x12	6.68	-.12
x8	<---	D	10.40	.32
x8	<---	A	20.61	.49
x8	<---	x13	8.72	.15
x8	<---	x12	6.25	.13
x8	<---	x3	5.88	.13
x8	<---	x2	20.83	.24
x8	<---	x1	20.90	.24
x6	<---	x11	6.40	.12
x5	<---	x12	4.14	-.09
x5	<---	x8	4.83	-.10
x2	<---	x8	9.85	.15
x1	<---	x8	9.54	.15

Maximum Modification Index is 20.61 for Element x8 ← A
修正指数:该参数由固定改为自由估计， 会减少的数值

Minimization History (Default model)

Iteration		- ve EV	Condition #	Smallest EV	Diameter	F	N Tries	Ratio
0	e	13		-.16	9999.00	1310.09	0	9999.00
1	e	0	83.44		2.15	390.84	20	.75
2	e	0	28.83		.67	308.08	5	.00
3	e	0	35.91		.80	219.15	2	.00
4	e	0	70.96		.55	196.21	1	.98
5	e	0	89.24		.16	194.60	1	.99
6	e	0	92.32		.02	194.57	1	1.01
7	e	0	91.81		.00	194.57	1	1.00

Model Fit Summary 拟合优度统计量

0.56 x (N-1) = chi
0.56x349=194.57

Sat. nPar = p x (p+1)/2
Ind. nPar = p

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	44	194.57	109	.00	1.79
Saturated model	153	.00	0		
Independence model	17	1104.79	136	.00	8.12

Model	RMR	GFI	AGFI	PGFI
Default model	.05	.94	.92	.67
Saturated model	.00	1.00		
Independence model	.19	.60	.55	.53

Oldest (LISREL) indexes:
RMR, GFI, AGFI, (SRMR)

TLI = NNFI

CFI = RNI

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.82	.78	.91	.89	.91
Saturated model	1.00		1.00		1.00
Independence model	.00	.00	.00	.00	.00

Model	PRATIO	PNFI	PCFI
Default model	.80	.66	.73
Saturated model	.00	.00	.00
Independence model	1.00	.00	.00

pRatio = df (hypo model)/df (null model)
= (153 - 44) / (153-17)
= 109 / 136 = 0.801

PCFI = pRatio x CFI
= .80 x .91
= .73

CMin – df = 194.57 - 109

Lo – Hi = confidence interval

Model	NCP	LO 90	HI 90
Default model	85.57	50.51	128.48
Saturated model	.00	.00	.00
Independence model	968.79	866.55	1078.48

0.56 x (N-1) = CMin
0.56x349=194.57

Model	FMIN	F0	LO 90	HI 90
Default model	.56	.25	.14	.37
Saturated model	.00	.00	.00	.00
Independence model	3.17	2.78	2.48	3.09

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.05	.04	.06	.64
Independence model	.14	.14	.15	.00

Model	AIC	BCC	BIC	CAIC
Default model	282.57	287.36	452.32	496.32
Saturated model	306.00	322.64	896.26	1049.26
Independence model	1138.79	1140.64	1204.37	1221.37

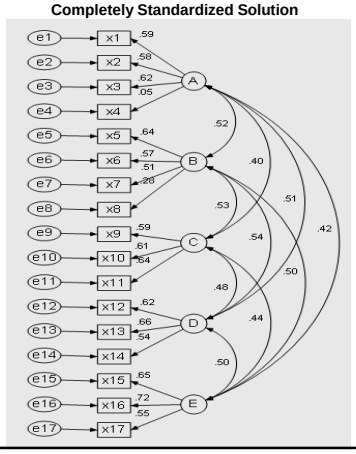
Model	ECVI	LO 90	HI 90	MECVI
Default model	.81	.71	.93	.82
Saturated model	.88	.88	.88	.92
Independence model	3.26	2.97	3.58	3.27

Model	HOELTER .05	HOELTER .01
Default model	242	263
Independence model	52	57

Execution time summary

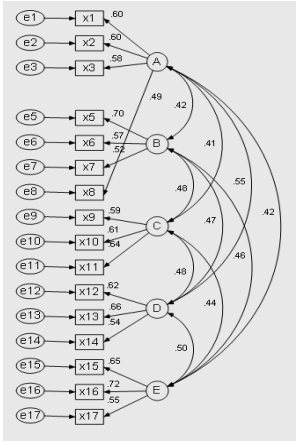
Minimization:	.00
Miscellaneous:	.13
Bootstrap:	.00
Total:	.13

Completely Standardized Solution



结果解释

- x4在A的负荷loading很小 (0.05)，但在其他因子的修正指数（MI, modification index）也不高 (0.00, 0.02, 0.00, 0.02 in B,C,D,E)
  - 不从属A，也不归属其他因子
- Q8在B的负荷不高（0.28），但在A的MI是20.6，可能归属A
- 因子间相关很高 (0.40 至 0.54)
- 模型拟合fit相当好：χ^2 (109) =194.57，RMSEA＝0.047, TLI = .89, CFI= .91。
- 仔细检查题目内容后，删去Q4, Q8归入A



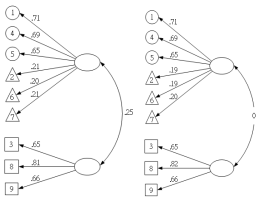
模型修正 M_B

- Q8归属A，因子负荷很高（.49），
- χ^2(94) =149.51，RMSEA＝.041，TLI＝.93，CFI = .94。

嵌套概念

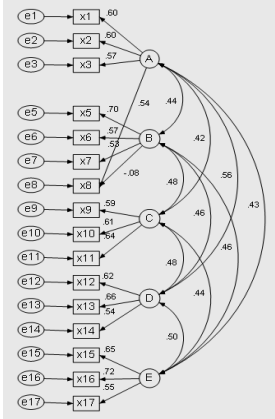
模型	df	χ^2	NNFI	CFI	(no. of estimated parameters) 需要估计的参数个数
M_X	26	249	.656	.752	19 = 9 Load + 9 Uniq + 1 Corr
M_Y	27	263	.649	.727	18 = 9 Load + 9 Uniq

$\Delta\chi^2$ (Δ df) = $\Delta\chi^2$ (1) = 14, p < .05



M_X M_Y

- 虽然没有嵌套关系，模型 M_B 比 M_A 好
- Q8同时从属A和B?



模型修正 M_C

- χ^2 (93)= 148.61, RMSEA ＝.041, TLI = .93, CFI = .94。
- Q8在A负荷为 .54，在B负荷为 -.08
- 因为概念上Q8应与B成正相关，故不合理。而且这负荷相对低，所以我们选择 M_B
- 通常，每题只归属一个因子

修正modification前后模型的拟合指数比较

模型	df	χ^2	RMSEA	NNFI	CFI	註
M-A	109	195	.046	.89	.91	原模型
M-B	94	150	.041	.93	.94	删Q4,Q8-A
M-C	93	149	.041	.93	.94	删Q4,Q8-A,B
MB-2	99	152	.039	.94	.95	2阶因子

结构方程模型及其应用
Structural Equation Model and Its Applications

V 多质多法模型
V multitrait-multimethod (MTMM)

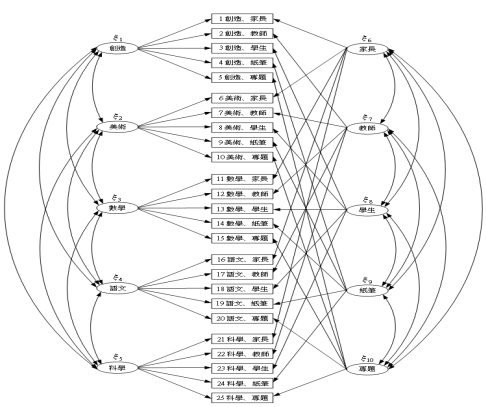
侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

多质多法模型 multitrait-multimethod(MTMM)

- 五种方法(method)：家长，教师，学生，纸笔测验，专题报告
- 五种能力(trait)：创造力，美术技巧，数学能力，语文能力，科学知识
- 25个得分（观测变量）5种方法x5种能力
- 分析方法一：相关特质相关方法 (CTCM, correlated-trait correlated-method)

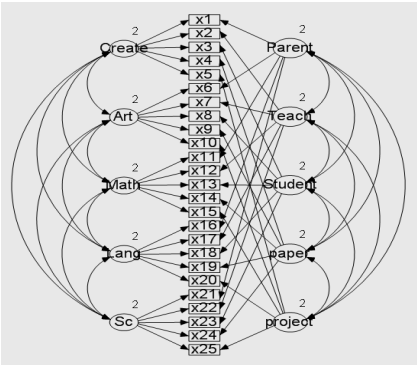
表 3-6 多质多法模型变量与因子从属关系

	特质因子					方法因子				
变量（题目）	创造	美术	数学	语文	科学	家长	教师	学生	纸笔	专题
1. 创造．家长	√					√				
2. 创造．教师	√						√			
3. 创造．学生	√							√		
4. 创造．纸笔	√								√	
5. 创造．专题	√									√
6. 美术．家长		√				√				
7. 美术．教师		√					√			
8. 美术．学生		√						√		
9. 美术．纸笔		√							√	
10.美术．专题		√								√
11.数学．家长			√			√				
12.数学．教师			√				√			
13.数学．学生			√					√		
14.数学．纸笔			√						√	
15.数学．专题			√							√
16.语文．家长				√		√				
17.语文．教师				√			√			
18.语文．学生				√				√		
19.语文．纸笔				√					√	
20.语文．专题				√						√
21.科学．家长					√	√				
22.科学．教师					√		√			
23.科学．学生					√			√		
24.科学．纸笔					√				√	
25.科学．专题					√					√




```
MATRIX DATA variables=ROWTYPE_ x1 to x25.
begin data.
N 500 500 500 500 500 500 500 500 500 500
500 500 500 500 500 500 500 500 500 500
500 500 500 500 500
SD 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
COR  1.0
COR .40 1.0
COR .44 .43 1.0
COR .39 .41 .43 1.0
COR .44 .38 .44 .45 1.0
COR .50 .21 .18 .19 .19 1.00
COR .19 .48 .22 .23 .18 .45 1.0
COR .20 .21 .53 .18 .23 .42 .43 1.0
COR .22 .19 .19 .53 .22 .41 .45 .45 1.0
COR .19 .17 .22 .19 .52 .46 .41 .39 .44 1.0
COR .49 .23 .23 .17 .23 .51 .23 .17 .23 .23 1.0
COR .24 .52 .19 .23 .19 .22 .48 .18 .19 .18 .45 1.0
COR .22 .22 .52 .22 .18 .19 .23 .51 .17 .22 .39 .43 1.0
COR .23 .18 .22 .49 .17 .24 .17 .24 .47 .18 .41 .44 .43 1.0
COR .18 .22 .23 .19 .48 .25 .21 .19 .22 .53 .44 .39 .42 .41 1.0
```

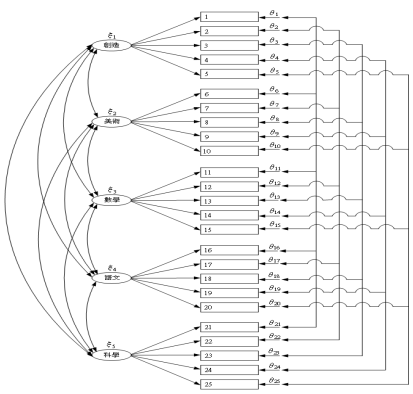
```
COR .48 .23 .18 .23 .23 .48 .18 .23 .25 .24 .55 .23 .19 .24 .23 1.0
COR .22 .51 .17 .19 .21 .19 .51 .19 .23 .23 .23 .52 .23 .19 .17 .43 1.0
COR .23 .22 .48 .22 .19 .23 .23 .53 .23 .22 .24 .22 .47 .22 .19 .45 .39 1.0
COR .19 .23 .23 .53 .22 .24 .19 .24 .53 .19 .21 .24 .23 .48 .23 .42 .46 .45 1.0
COR .20 .24 .17 .23 .49 .21 .16 .19 .19 .51 .24 .18 .24 .22 .52 .41 .43 .39 .45 1.0
COR .51 .22 .18 .19 .18 .52 .25 .24 .24 .23 .49 .24 .21 .24 .24 .53 .24 .23 .23 .24 1.0
COR .22 .53 .23 .18 .17 .25 .53 .17 .18 .24 .19 .52 .18 .19 .17 .24 .49 .19 .21 .21 .45 1.0
COR .19 .22 .47 .22 .23 .22 .23 .47 .17 .17 .23 .22 .47 .21 .24 .23 .23 .53 .24 .23 .41 .45 1.0
COR .22 .23 .22 .52 .21 .19 .17 .24 .53 .19 .19 .19 .18 .53 .22 .19 .21 .18 .51 .18 .43 .41 .45 1.0
COR .20 .24 .19 .19 .49 .18 .19 .19 .23 .52 .23 .17 .23 .19 .52 .24 .18 .24 .17 .25 .44 .42 .43 .41 1.0
END data.
save outfile='D:\My Documents\AMOS\pg\chap3_2_MTMM_corr.sav
```

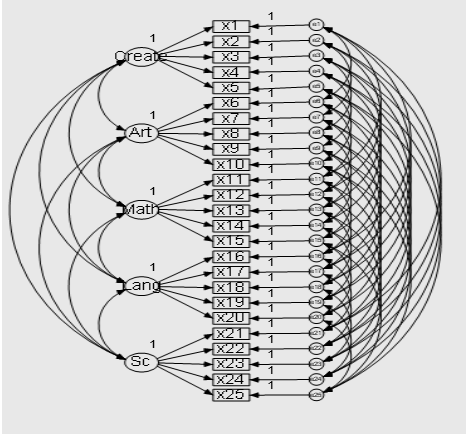
- 许多时候CTCM模型并不收敛(non-converged)，在本例中，用固定方差法，固定为1也不收敛(similarly fixed loading at 1 converges to an improper solution)，可固定为2来解决(helps only in this specific case)
- 模型复杂，过早检查解答是否正定(positive definite, proper)并不合适，所以加大迭代次数(iteration number)

多质多法模型 方法二:相关特质相关特性 correlated-trait correlated uniqueness (CTCU）

- 较大MTMM模型（如7方法×7特质）收敛机会较大



- 只留下首五个特质因子（NK＝5）
- 容许他们的特殊因子uniqueness(也称为误差, error）相关
- e.g., 第1、6、11、16、21个变量为同一个方法的分数



Checking

- proper solutions
- fit reasonably well
- assuming all items are correctly oriented (all +ve related), then
  - in CTCM: all trait and method effect loadings → all +ve;
  - in CTCU: all trait method loadings +ve; all CU +ve/zero
  - correlations among traits or among methods: depend on theory

结构方程模型及其应用
Structural Equation Model and Its Applications

VI 全模型
VI Full model

侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

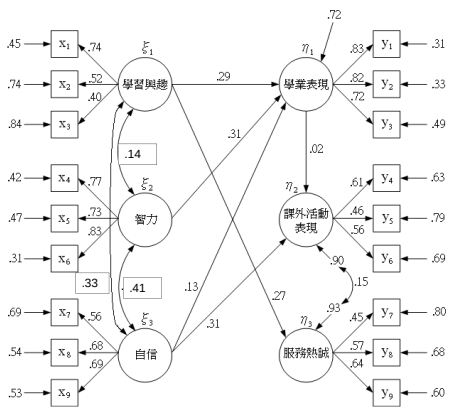
全模型(Full model)

- 兴趣(x1,2,3)、学生智力(x4,5,6)、自信(x7,8,9)如何影响学业(y1,y2,y3)、课外活动(y4,5,6)和服务热诚 (y7,8,9)? N =500
- path是factor对 factor因子的效应, 与传统的回归系数相似。
- residual (结构方程残差)是因子的残差（未被解释的部份）方差。

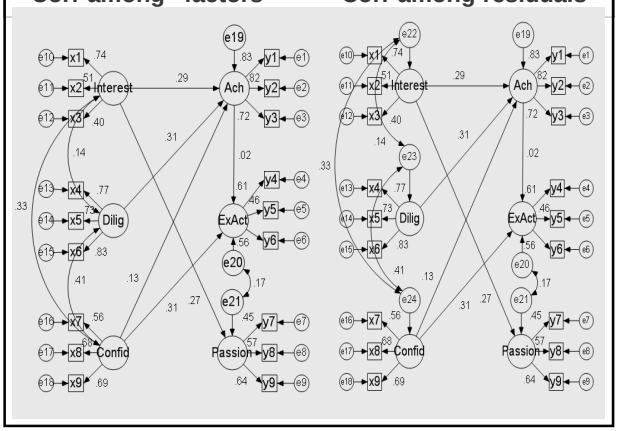

```
MATRIX DATA variables=ROWTYPE_  y1 to y9 x1 to x9.
begin data.
N 500 500 500 500 500 500 500 500 500 500
500 500 500 500 500 500 500 500
SD 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
COR 1
COR .68 1
COR .60 .58 1
COR .01 .10 .07 1
COR .12 .04 .06 .29 1
COR .06 .06 .01 .35 .24 1
COR .09 .13 .10 .05 .03 .07 1
COR .04 .08 .16 .10 .12 .06 .25 1
COR .06 .09 .02 .02 .09 .16 .29 .36 1
COR .23 .26 .19 .05 .04 .04 .08 .09 .09 1
COR .11 .13 .12 .03 .05 .03 .02 .06 .06 .40 1
```



```
COR .16 .09 .09 .10 .10 .02 .04 .12 .15 .29 .20 1
COR .24 .26 .22 .14 .06 .10 .06 .07 .08 .03 .04 .02 1
COR .21 .22 .29 .07 .05 .17 .12 .06 .06 .03 .12 .04 .55 1
COR .29 .28 .26 .06 .07 .05 .06 .15 .20 .10 .03 .12 .64 .61 1
COR .15 .16 .19 .18 .08 .07 .08 .10 .06 .15 .16 .07 .25 .25 .16 1
COR .24 .20 .16 .13 .15 .18 .19 .18 .14 .11 .07 .16 .19 .21 .22 .35 1
COR .14 .25 .12 .09 .11 .09 .09 .11 .21 .17 .09 .05 .21 .23 .18 .39 .48 1
END data.
save outfile='D:\My Documents\AMOS\pg\chap3_2_Full_corr.sav'.
```

Corr among "factors" **Corr among residuals**

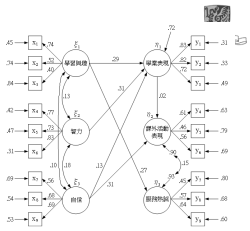


结果解释

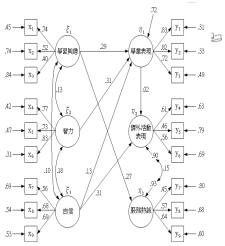
- $\chi^2(125)$=292.51，RMSEA=0.052，TLI=0.90, CFI=0.92 ，拟合不错
- interesting, MI (Passion ← ExAct) = 2.116 by Amos, but should be 292.5 - 270.1 = 22.4 (by re-running the programs and compare)，因为Passion ← ExAct 理论上不太合理，且Passion, ExAct间已有相关; thus do not change here
- 故第一个修正模型M2是让Passion ← Confidence (MI = 17.10, Parm Change= .224)自由估计，说明增加這路径是合适)(and must be meaningful substantially)。
- 然后考虑要不要减少原有路径。在各因子关系中，path(ExAct ← Ach) = 0.011（SE = 0.052, *t* = 0.214)的效应最小，可以删除该路径。将模型M2的ExAct ← Ach固定为0，变成模型M3。

- 增加自由参数free parameter（模型变复杂），模型的卡方chi-square会减少；减少自由参数（模型变简单），模型的卡方会增加。
- 如果增加自由参数后，卡方非常显著地减少，说明增加自由参数是值得的。
- 如果减少自由参数后，卡方没有显著地增加，说明减少自由参数是可取的。

element	一般设定方法
loading	(a)指标与因子有从属关系的:自由估计 (b)若用"固定负荷法"，则每个因子,选取一个负荷固定为"1"
Factor corr: exogenous factors	(a)非对角线元素:因子互有相关的位置,自由估计 (b)对角线元素: 若用"固定方差法",则:固定为"1" 若用"固定负荷法",则:自由估计
Factor corr: endogenous factors	(a)非对角线元素(cov):因子残差互有相关的位置,自由估计 (b)对角线元素(var):自由估计

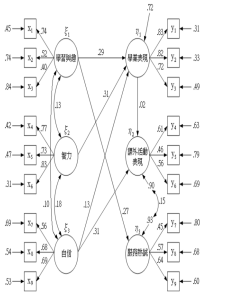


element	一般设定setting方法
Item uniqueness	(a)对角线diagonal的元素:自由估计 (b)非对角线non-diagonal的元素:固定为"零"；有特殊情况,容许额外的对应相关,该元素自由估计
directional path	因子对因子有效应effect的参数:自由估计



结构方程模型的结构

- 测量模型 (measurement model)
  $\boldsymbol{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}$
  $\boldsymbol{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$
  $\boldsymbol{x}$ ——外源指标exogenous（如6项社经指标）
  $\boldsymbol{y}$ —内生指标endogenous（如：语、数、英成绩）
  $\boldsymbol{\Lambda}_x$ $\boldsymbol{\Lambda}_y$ —因子负荷矩阵(loading)
  $\boldsymbol{\delta}$ $\boldsymbol{\varepsilon}$—误差项 (uniqueness, measurement errors)
- 结构模型 (structural model)
  $\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$



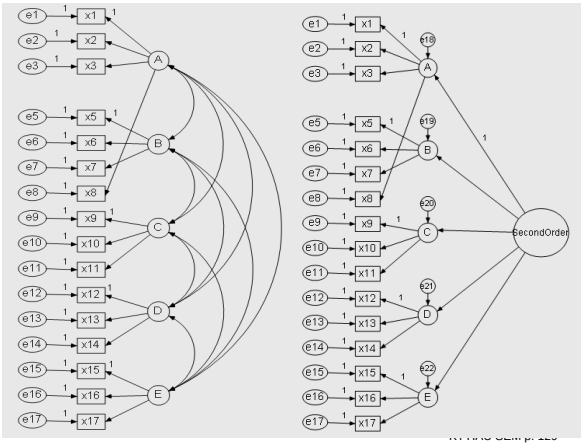
结构方程模型及其应用
Structural Equation Model and Its Applications

VII 高阶因子分析
VII High-order Factor Analysis

侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

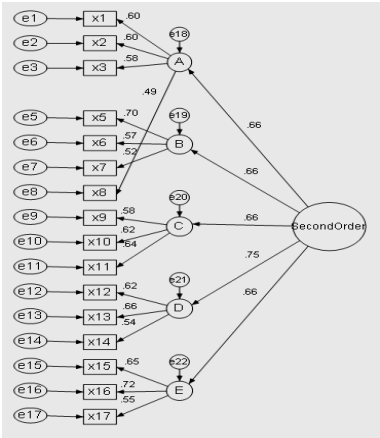
高阶因子分析 (high-order factor analysis)

- 设一阶first-order能力因子有相关，需估计的参数很多。5个一阶因子时，共有10个因子间相关。
- 设有一个普遍能力（二阶second order）因子，影响各一阶能力因子的表现。10个相关改由5个参数parameter（二阶因子与一阶因子的关系）所替代。
- 二阶因子卡方必然较大，自由度df也增加，只要增加的卡方不到显著水平nonsignificant，从模型简洁性parsimony，我们选择二阶模型second-order factor model



In SPSS, prepare the raw data or corr matrix

```
MATRIX DATA variables=ROWTYPE_ v1 to v17.
begin data.
N 350 350 350 350 350 350 350 350 350 350 350
  350 350 350 350 350 350
SD 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
CORR 1
CORR .34 1
CORR .38 .35 1
CORR .02 .03 .04 1
CORR .15 .19 .14 .02 1
CORR .17 .15 .20 .01 .42 1
CORR .20 .13 .12 .00 .40 .21 1
CORR .32 .32 .21 .03 .10 .10 .07 1
CORR .10 .17 .12 .02 .15 .18 .23 .13 1
CORR .14 .16 .15 .03 .14 .19 .18 .18 .37 1
CORR .14 .15 .19 .01 .18 .30 .13 .08 .38 .38 1
CORR .18 .16 .24 .02 .14 .21 .21 .22 .06 .23 .18 1
CORR .19 .20 .15 .01 .14 .24 .09 .24 .15
  .21 .21 .45 1
CORR .18 .21 .18 .03 .25 .18 .18 .18 .22
  .12 .24 .28 .35 1
CORR .08 .18 .16 .01 .22 .20 .22 .12 .12
  .16 .21 .25 .20 .26 1
CORR .12 .16 .25 .02 .15 .12 .20 .14 .17
  .20 .14 .20 .15 .20 .50 1
CORR .20 .16 .18 .04 .25 .14 .21 .17 .21
  .21 .23 .15 .21 .22 .29 .41 1
END data.save outfile =
  'D:\AMOS\chap3CFA.sav'.
```

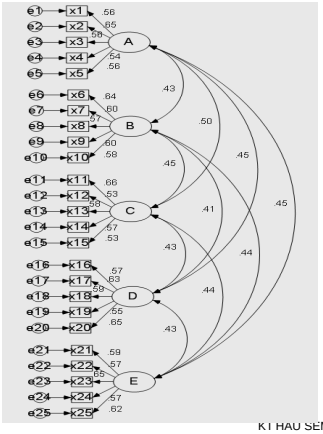



解释结果

- MB-2ord节省5个df，χ^2大致相同，其他指数index拟合fit较好
- 二阶因子与一阶因子关系（GA系数）很强(.66, .66, .66, .75, .66)
- 若一阶因子间相关很弱，没有建立二阶因子的需要
- 当模型只有3个一阶因子时（共有3个相关），二阶因子在数学上等同于equivalent一阶因子模型
- 因拟合指数fit index反映整个模型的拟合程度，一阶因子模型要有较好的拟合指数。对因子少的一阶模型（如：只含4或5个一阶因子），一般一阶与二阶拟合指数相差不大，难区分non-differentiating

另一个二阶因子模型例子

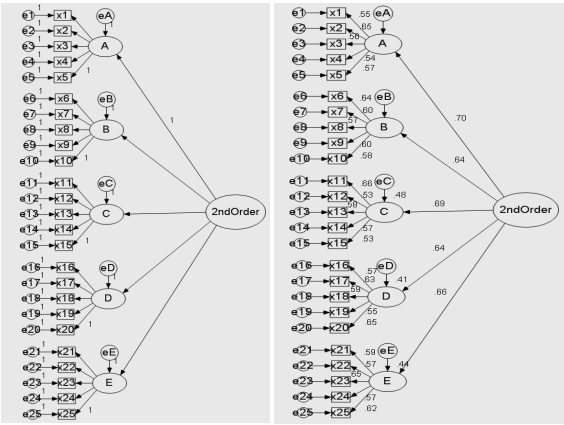
25个题:语文、数学、英语、历史和地理能力

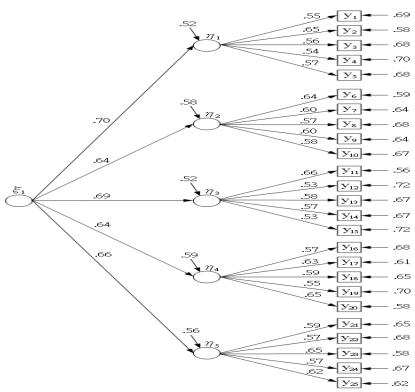



```
MATRIX DATA variables=ROWTYPE_ x1 to x25.
begin data.
N 500 500 500 500 500 500 500 500 500 500
500 500 500 500 500 500 500 500 500 500
500 500 500 500 500
SD 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
COR 1.0
COR .40 1.0
COR .44 .43 1.0
COR .39 .41 .43 1.0
COR .44 .38 .44 .45 1.0
COR .50 .21 .18 .19 .19 1.00
COR .19 .48 .22 .23 .18 .45 1.0
COR .20 .21 .53 .18 .23 .42 .43 1.0
COR .22 .19 .19 .53 .22 .41 .45 .45 1.0
COR .19 .17 .22 .19 .52 .46 .41 .39 .44 1.0
COR .49 .23 .23 .17 .23 .51 .23 .17 .23 .23 1.0
COR .24 .52 .19 .23 .19 .22 .48 .18 .19 .18 .45 1.0
COR .22 .22 .52 .22 .18 .19 .23 .51 .17 .22 .39 .43 1.0
COR .23 .18 .22 .49 .17 .24 .17 .24 .47 .18 .41 .44 .43 1.0
```



```
COR .18 .22 .23 .19 .48 .25 .21 .19 .22 .53 .44 .39 .42 .41 1.0
COR .48 .23 .18 .23 .23 .48 .18 .23 .25 .24 .55 .23 .19 .24 .23 1.0
COR .22 .51 .17 .19 .21 .19 .51 .19 .23 .23 .23 .52 .23 .19 .17 .43 1.0
COR .23 .22 .48 .22 .19 .23 .23 .53 .23 .22 .24 .22 .47 .22 .19 .45 .39 1.0
COR .19 .23 .23 .53 .22 .24 .19 .24 .53 .19 .21 .24 .23 .48 .23 .42 .46 .45 1.0
COR .20 .24 .17 .23 .49 .21 .16 .19 .19 .51 .24 .18 .24 .22 .52 .41 .43 .39 .45 1.0
COR .51 .22 .18 .19 .18 .52 .25 .24 .24 .23 .49 .24 .21 .24 .24 .53 .24 .23 .23 .24 1.0
COR .22 .53 .23 .18 .17 .25 .53 .17 .18 .24 .19 .52 .18 .19 .17 .24 .49 .19 .21 .21 .45
    1.0
COR .19 .22 .47 .22 .23 .22 .23 .47 .17 .17 .23 .22 .47 .21 .24 .23 .23 .53 .24 .23 .41
    .45 1.0
COR .22 .23 .22 .52 .21 .19 .17 .24 .53 .19 .19 .19 .18 .53 .22 .19 .21 .18 .51 .18 .43
    .41 .45 1.0
COR .20 .24 .19 .19 .49 .18 .19 .19 .23 .52 .23 .17 .23 .19 .52 .24 .18 .24 .17 .25 .44
    .42 .43 .41 1.0
END data.
save outfile='D:\My Documents\AMOS\pg\chap3_2_MTMM_corr.sav'.
```

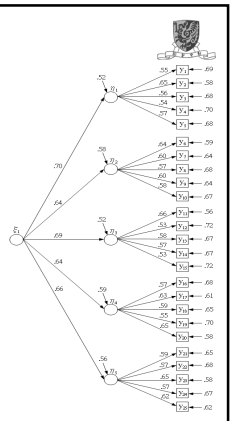





- M-1-ord: χ^2 = 464, df = 265，RMSEA = .039, TLI = .914, CFI = .924 ; 5个因子之间的相关系数在 .41至 .50之间。
- M-2-ord: 拟合优度大致相同，χ^2 = 465, df = 270, RMSEA = .038, TLI = .918, CFI = .926。按简约parsimony原则，我们应取二阶模型2-order factor model。二阶与一阶因子关系也很强（2nd order factor → 1st order factor = .70, .64, .69, .64, .66)

一般在二阶因子模型中：

- 一阶因子间不再容许相关
- 二阶与一阶因子间路径:方向是由二阶至一阶
- 二阶与一阶因子各路径中,我们取其中一个固定为1(固定负荷法)
- 对只有3个或以下一阶因子,不再构划二阶因子
- 何时宁取二阶模型，要考虑：
  - 二阶自由度较对应一阶模型为大
  - 二阶模型较对应一阶模型简单(parsimonious)
  - 二阶的χ2较一阶为大
  - 若二阶模型简化甚多，但χ2增加不多(模型拟合恶化不严重），则宁取二阶模型
- 在LISREL中设定高阶因子,可
  - 二阶因子用ξ,一阶因子用η代表
  - 二阶与一阶因子均用η代表

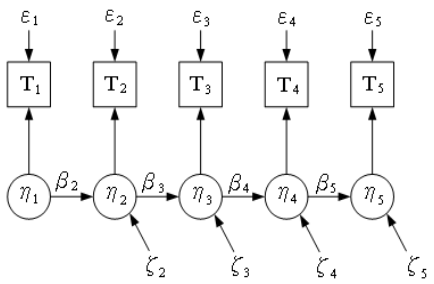


结构方程模型及其应用
Structural Equation Model and Its Applications

VIII 单纯形模型
VIII simplex model

侯傑泰(侯杰泰)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

单纯形模型(simplex model)



Simplex correlation matrix

单纯形模型相关矩阵

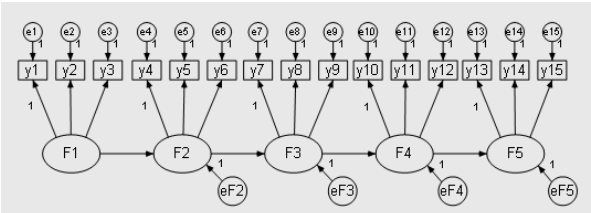
单纯形模型相关矩阵的形态

	T_1	T_2	T_3	T_4	T_5
T_1	1		↑		
T_2		1	下降→		
T_3		←下降	1	下降→	
T_4			←下降	1	
T_5			↓		1

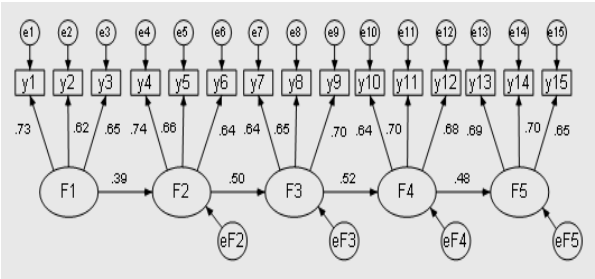
单纯形模型相关矩阵

	T_1	T_2	T_3	T_4	T_5
T_1	1	0.5	0.4	0.3	0.2
T_2	0.5	1	0.5	0.4	0.3
T_3	0.4	0.5	1	0.5	0.4
T_4	0.3	0.4	0.5	1	0.5
T_5	0.2	0.3	0.4	0.5	1

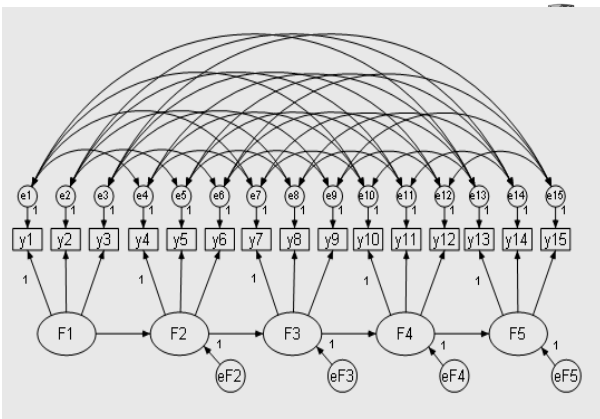
拟单纯形模型 (quasi-simplex)



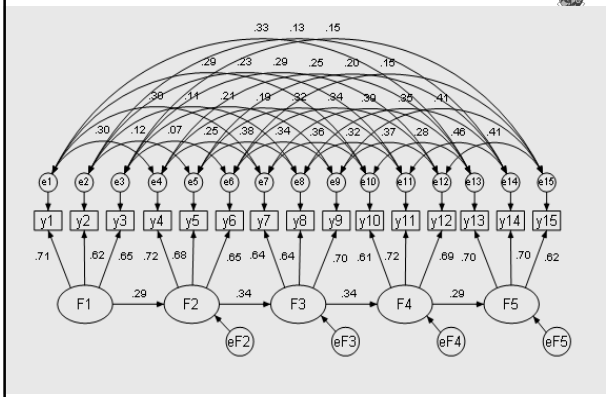

```
MATRIX DATA variables=ROWTYPE_ y1 to y15.
begin data.
N 500 500 500 500 500 500 500 500 500 500 500 500 500 500 500
SD 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
COR 1
COR .45 1
COR .47 .41 1
COR .29 .11 .17 1
COR .14 .19 .13 .49 1
COR .12 .14 .18 .46 .43 1
COR .25 .07 .08 .32 .14 .16 1
COR .09 .12 .04 .15 .33 .14 .42 1
COR .10 .07 .18 .24 .12 .35 .46 .45 1
COR .21 .04 .05 .23 .15 .08 .35 .19 .16 1
COR .09 .19 .14 .12 .25 .03 .13 .34 .15 .45 1
COR .04 .02 .22 .09 .04 .21 .12 .18 .35 .42 .49 1
COR .18 .02 .04 .16 .04 .02 .27 .12 .12 .32 .15 .16 1
COR .05 .12 .11 .05 .17 .14 .04 .29 .09 .15 .35 .13 .49 1
COR .03 .01 .13 .12 .04 .17 .02 .15 .32 .16 .12 .37 .44 .45 1
END data.
save outfile='D:\My Documents\AMOS\pg\chap3_2_simplex1_corr.sav'.
```



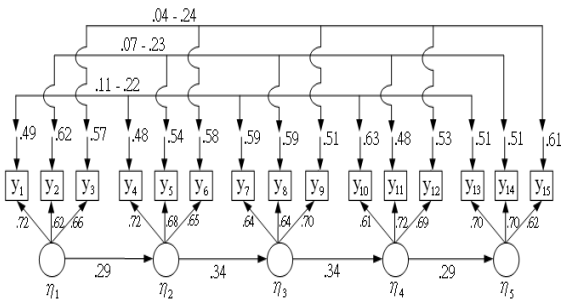
KT HAU SEM p. 145



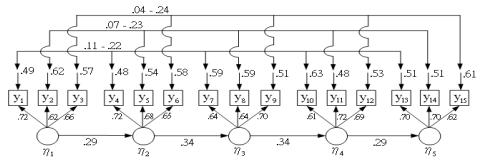
KT HAU SEM p. 146



KT HAU SEM p. 147



KT HAU SEM p. 148



对单纯或拟单纯形模型,一般而言

- 对单指标的单纯型模型, 首尾两个指标的误差方差，我们需强制为零或相等
- 多指标的拟单纯形模型,所有指针误差方差可自由估计
- 用误差相关去描述跨年相同对应学科误差的相关时，这些相关应为正值(positive)才合理
- 加了误差相关,模式的自由度减少

KT HAU SEM p. 149

结构方程模型及其应用
Structural Equation Model and Its Applications

IX 多组SEM分析
IX Multiple-group SEM

侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

KT HAU SEM p. 150

多组SEM分析 (multiple-group SEM)

- 第一类:多组验证性因子分析multiple-group CFA（或路径分析path analyses）
  - 各组（例如男、女组）的因子结构factor structure是否相同？某些路径path参数parameter/coefficient在不同的组是否有显著差异significant difference？(与比较多组回归系数regression coefficients in multiple group是否相同类似)
- 第二类:各组的因子均值是否相同(Multiple-group Mean Structure Analysis)。这与传统方差分析ANOVA相似 (通常需要先做第一类分析)

多组验证性因子分析 multiple-group CFA

1. 形态相同（configural/pattern invariance）
2. 因子负荷factor loading LX等同invariance
3. 误差方差uniqueness TD等同invariance
4. 因子方差factor variance, diagonal of PH等同invariance
5. 因子协方差factor covariance PH 等同invariance

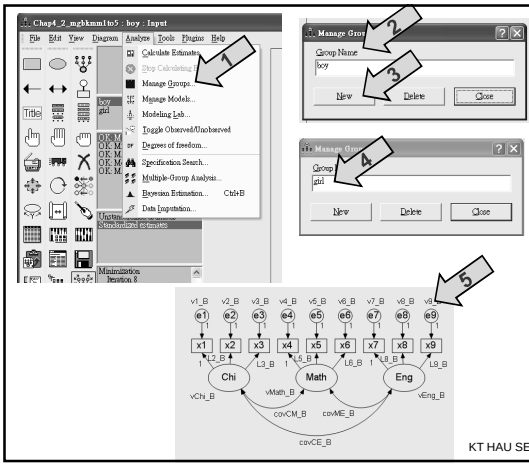
表4-2 多组验证性因子分析各模型的拟合指数

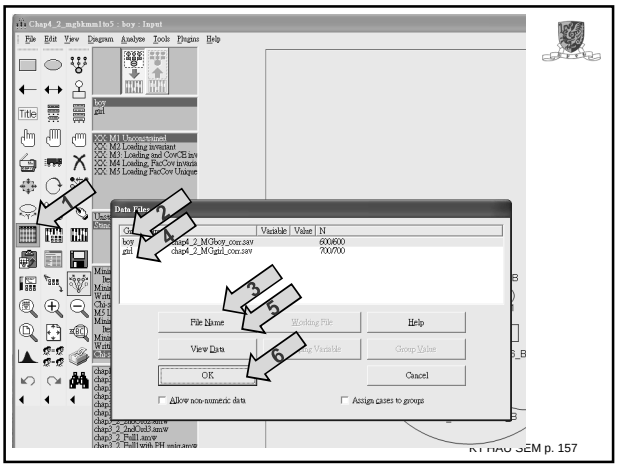
Model	df	**chi-2**	RMSEA	TLI	CFI
M0,M男生单独估计	24	**49.57**	.042	**.966**	.977
M0,F女生单独估计	24	**44.93**	.035	**.974**	.982
M1 两组同时估计, no Inv	48	**94.50**	.027	**.970**	.980
M2 Loading Inv	54	**107.19**	.028	**.969**	.977
M3 Ld, PH(3,1) Inv	55	**107.52**	.027	**.970**	.977
M4 Ld, FacCov Inv	60	**109.32**	.025	**.974**	.979
M5 Ld、FacCov、U Inv	69	**131.21**	.026	**.972**	.973
M7 Ld,FacCov,U,Intrcpt Inv; Fac meanFree	75	**132.23**	.024	**.976**	.975
M8 Ld,FacCov,U,Intrcpt, Fac mean Inv	78	**146.80**	.026	**.973**	.970

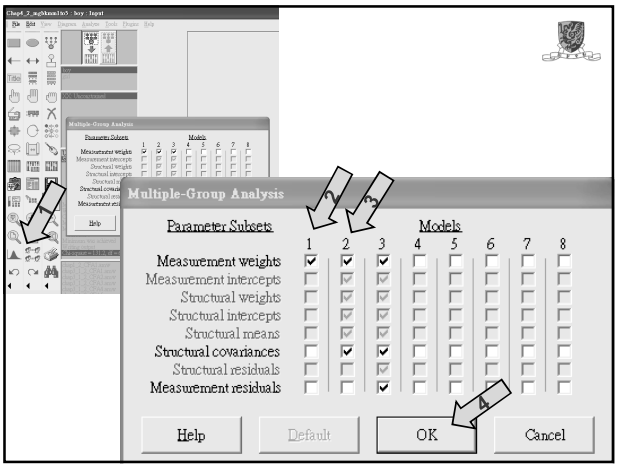

```
MATRIX DATA variables=ROWTYPE_ x1 to x9.
begin data.
N 600 600 600 600 600 600 600 600 600
SD 1.07 1.23 0.98 1.02 1.01 1.03 0.99 1.06 0.98
MEAN 2.01 2.45 2.67 3.21 3.33 3.45 2.67 2.19 2.34
COR 1
COR 0.68  1
COR 0.60  0.58 1
COR 0.05  0.10 0.07  1
COR 0.12  0.14 0.06  0.29 1
COR 0.10  0.06 0.04  0.35 0.38  1
COR 0.13  0.13 0.10  0.05 0.08  0.07  1
COR 0.04  0.10 0.16  0.10 0.12  0.06  0.35  1
COR 0.09  0.09 0.02  0.09 0.15  0.16  0.29  0.36  1
END data.
save outfile='D:\Docs\AMOS\pg\chap4_2_MGboy_corr.sav'.
```

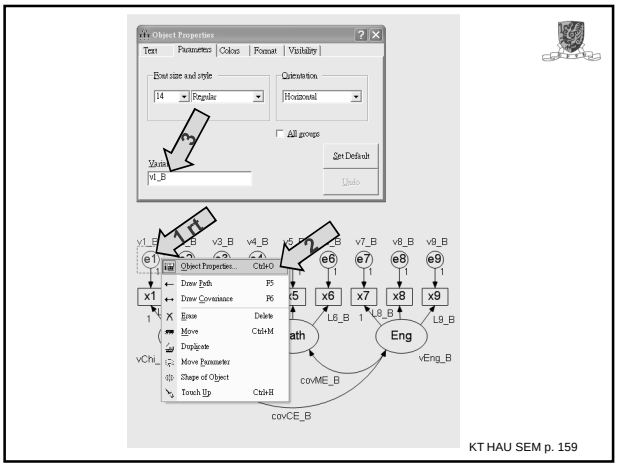


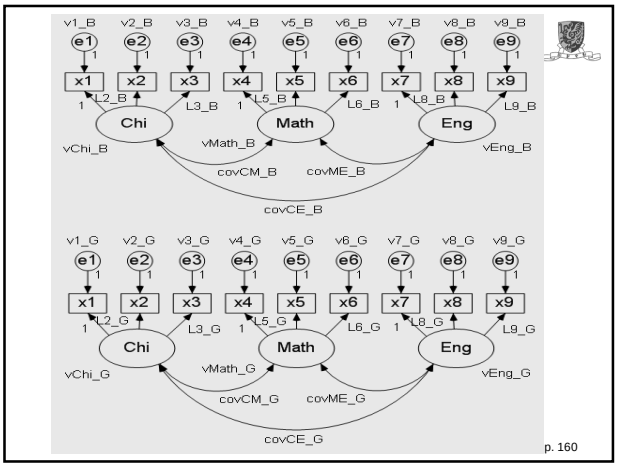
```
MATRIX DATA variables=ROWTYPE_ x1 to x9.
begin data.
N 700 700 700 700 700 700 700 700 700
SD 1.05 1.20 1.02 0.99 1.02 1.02 1.02 1.04 0.96
MEAN 2.02 2.48 2.69 3.10 3.20 3.38 2.75 2.29 2.45
COR 1
COR 0.72  1
COR 0.50  0.49 1
COR 0.10  0.08 0.10  1
COR 0.05  0.08 0.06  0.35 1
COR 0.14  0.10 0.04  0.28 0.26  1
COR 0.11  0.09 0.08  0.15 0.10  0.05  1
COR 0.10  0.15 0.16  0.04 0.08  0.10  0.27  1
COR 0.14  0.10 0.12  0.14 0.12  0.18  0.30  0.40  1
END data.
save outfile='D:\Docs\AMOS\pg\chap4_2_MGgirl_corr.sav'.
```

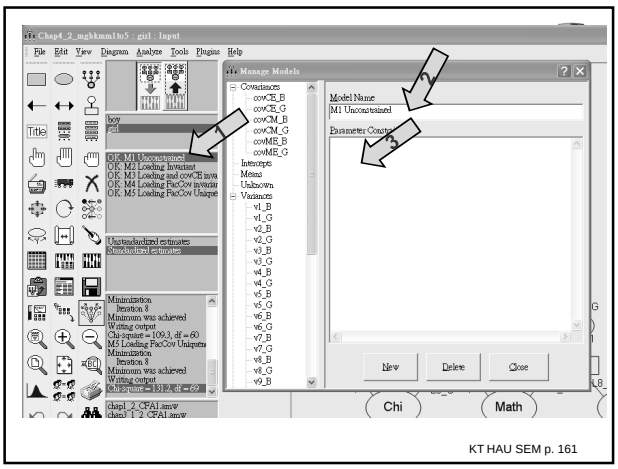



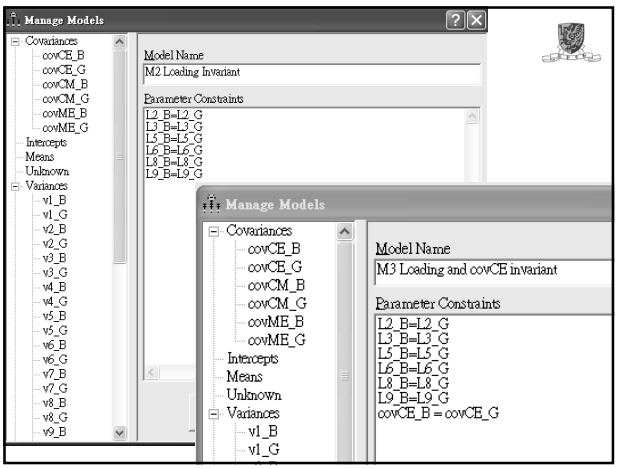

Data Files
File Name
Working File
Help
View Data
Group Variable
Group Value
OK
Cancel
Allow non-numeric data
Assign cases to groups
KT HAU SEM p. 157

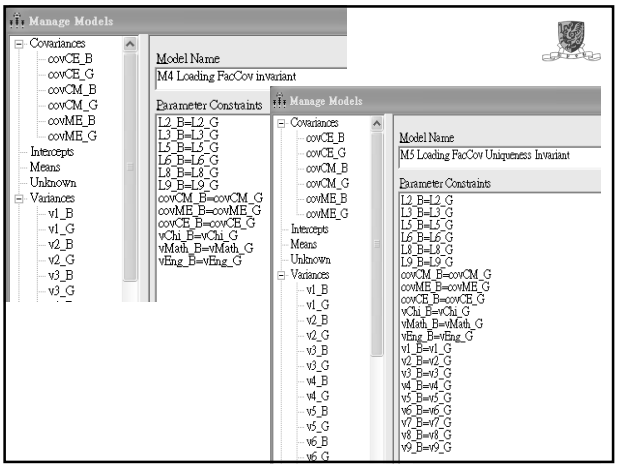

Multiple-Group Analysis
Parameter Subsets
Models
1 2 3 4 5 6 7 8
Measurement weights
Measurement intercepts
Structural weights
Structural intercepts
Structural means
Structural covariances
Structural residuals
Measurement residuals
Help
Default
OK
Cancel


Object Properties
Text
Parameters
Colors
Format
Visibility
Font size and style
Orientation
Horizontal
All groups
Set Default
Undo
Object Properties...
Draw Path
Draw Covariance
Erase
Move
Duplicate
Move Parameter
Shape of Object
Touch Up
KT HAU SEM p. 159


v1_B v2_B v3_B v4_B v5_B v6_B v7_B v8_B v9_B
e1 e2 e3 e4 e5 e6 e7 e8 e9
x1 x2 x3 x4 x5 x6 x7 x8 x9
L2_B L3_B L5_B L6_B L8_B L9_B
Chi
Math
Eng
vChi_B
vMath_B
vEng_B
covCM_B
covME_B
covCE_B
v1_G v2_G v3_G v4_G v5_G v6_G v7_G v8_G v9_G
L2_G L3_G L5_G L6_G L8_G L9_G
vChi_G
vMath_G
vEng_G
covCM_G
covME_G
covCE_G
p. 160


Manage Models
Covariances
Intercepts
Means
Unknown
Variances
Model Name
M1 Unconstrained
Parameter Constraints
New
Delete
Close
Chi
Math
KT HAU SEM p. 161

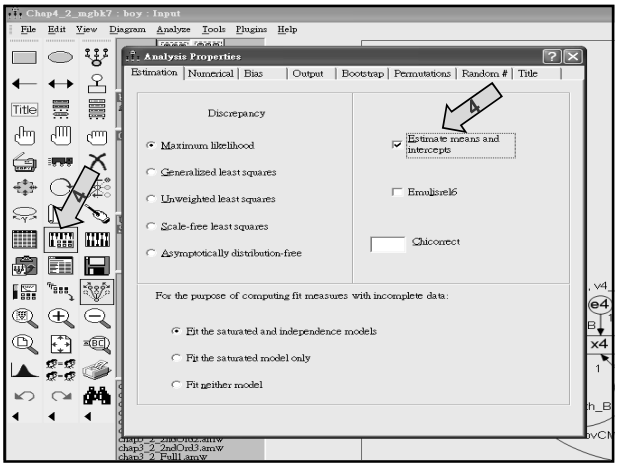

Manage Models
Covariances
covCE_B
covCE_G
covCM_B
covCM_G
covME_B
covME_G
Intercepts
Means
Unknown
Variances
v1_B
v1_G
Model Name
M2 Loading Invariant
Parameter Constraints
L2_B=L2_G
L3_B=L3_G
L5_B=L5_G
L6_B=L6_G
L8_B=L8_G
L9_B=L9_G
Manage Models
Model Name
M3 Loading and covCE invariant
Parameter Constraints
L2_B=L2_G
L3_B=L3_G
L5_B=L5_G
L6_B=L6_G
L8_B=L8_G
L9_B=L9_G
covCE_B = covCE_G

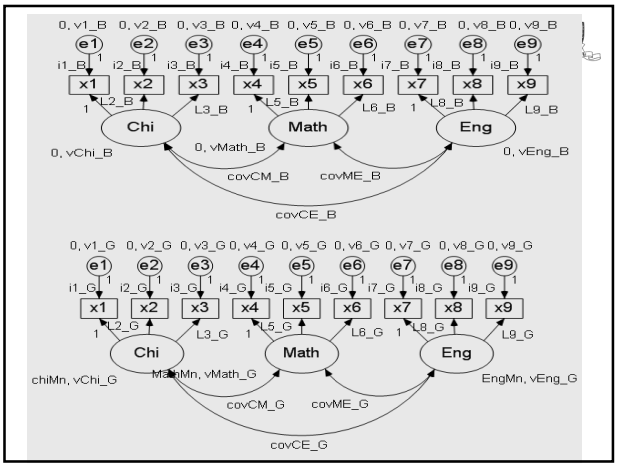


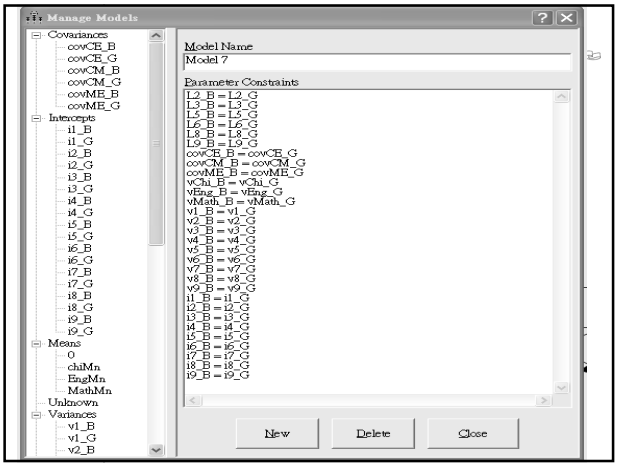
多组分析：均值结构模型
Multiple-group Mean Structure Analysis

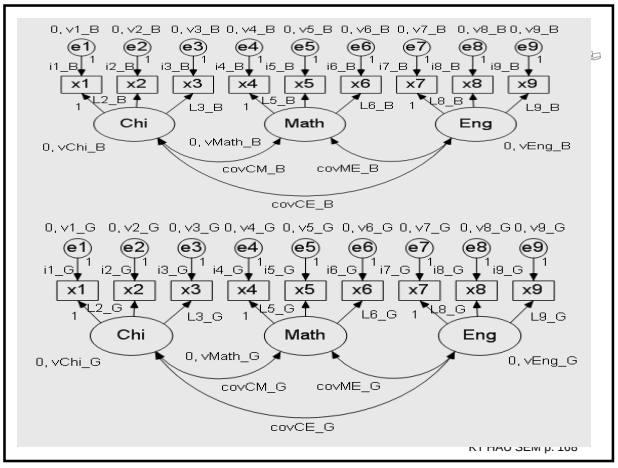
- 不同组别因子均值是否有显著差异(均值结构模型, mean structure models)
- 首先需确定各组的负荷loading相同invariance
  - 更希望因子协方差等同factor covariance,误差方差等同难实现
- 指标截距indicator interceptTX等同
  - 先让第1组的Tau自由
  - 要求其他组别Tau与第1组的相等
  - 因子均值等同 factor mean equivalence
  - 先设定第1组各因子均值为0 (Kappa)
  - 容许其他组的Kappa元素自由估计
  - 因子值>2倍SE（t>2.0）,则因子不同于第1组

KT HAU SEM p. 164









- 结果显示:
  - 第2组(女)的KA元素（即语文、数学、英语均值mean）为0.019, -0.102和0.083
  - 对应的SE为0.054, 0.041, 0.036
  - *t*-值为0.352、-2.470、2.331
  - 这表示:
    - 语文自信 -- 男(均值为0)女(均值为0.019)无差异 (t=0.352, n.s.)
    - 男生（均值为0）的数学自信高于女生（均值 ＝ -0.102, *t* = 2.470）
    - 女生的英语自信（均值 ＝ 0.083）则高于男生（均值为0, *t* = 2.331）

多组比较的次序

在SEM内,比较多组的因子均值, 一般依下述次序，遂项加上条件:

- 各组因子与指标的从属关系(形态)相同
- 各组因子负荷(loading)相同
- 各组因子间相关(协方差) (variance/covariance)相同
- 各组指标误差(特性)方差(uniqueness)相同
- 各组指标截距(Tau)相同
- 各组因子均值相同(Kappa)

多组比较检查原则

- 在检查等同条件时,未加等同条件,模型较复杂(df较小)
- 加了等同条件,模型较简单parsimonious (df较大)
- 未加”等同”条件, χ^2较小(拟合较好)
- 加”等同”条件, χ^2较大(拟合较差)
- 加”等同”条件,若模型简化甚多,但拟合优只是轻微恶化,则”等同”成立及合理
- 加”等同”条件,若模型简化不多,但拟合优严重恶化,则各组并不等同(等同不成立)

多组比较设定方法

等同检查	第一组	其它组别	检查(每组解答洽当proper外…)
型态等同	依先验(a priori)模型,估计各参数(parameter)	依先验模型,估计各自由参数	每组拟合优良good fit
Load等同	估计各自由Loading	Loading invariant	拟合优fit无严重恶化
var/covar 等同	估计各自由var/covar	var/covar invariant	拟合优fit无严重恶化
uniqueness 等同	估计各自由uniqueness	uniqueness invariant	拟合优fit无严重恶化
intercept 等同	自由估计intercept	intercept invariant	拟合优fit无严重恶化
Factor mean 等同	fixed at zero	freely estimated	每组mean与第一组(或他组)无显著差异 significant difference

结构方程模型及其应用
Structural Equation Model and Its Applications

X 专题: 结构方程建模和分析步骤
X Issues: Model Specification and analyses

侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

专题: 结构方程建模和分析步骤
Model Specification and analyses

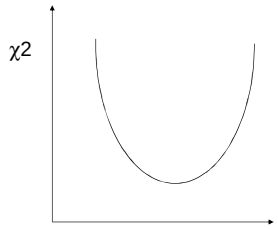
A. 验证模型与产生模型Confirmatory, Model generation

- 纯粹验证（strictly confirmatory，SC）
  - 心目中只有一个模型
  - 这类分析不多，无论接受还是拒绝，仍希望有更佳的选择
- 选择模型（alternative models，AM）
  - 从拟合的优劣，决定那个模型最为可取
  - 但我们仍常做一些轻微修改，成为MG类的分析

- 产生模型（model generating，MG)
  - 先提出一个或多个基本模型
  - 基于理论或数据，找出模型中拟合欠佳的部份
  - 修改模型，通过同一或其他样本，检查修正模型model respecification的拟合程度，目的在于产生一个最佳模型

B. 结构方程分析步骤
- 模型建构（model specification）,指定
  - 观测变量与潜变量（因子）的关系
  - 各潜变量间的相互关系（指定哪些因子间有相关或直接效应direct effect）
  - 在复杂的模型中，可以限制constrain因子负荷loading或因子相关系数等参数的数值或关系（例如，2个因子间相关系数correlation等于0.3；2个因子负荷必须相等）
- 模型拟合(model fitting)
  - 通常用 ML作模型参数的估计(versus 回归分析，通常用所最小二乘方法拟合模型，相应的参数估计称为最小二乘估计)

C. 模型评价（model assessment）
- 结构方程的解solution是否适当proper，估计是否收敛，各参数估计值是否在合理范围内（例如，相关系数在 +1与－1之内）



- 参数与预设模型的关系是否合理。当然数据分析可能出现一些预期以外的结果，但各参数绝不应出现一些互相矛盾，与先验假设有严重冲突的现象
- 检视多个不同类型的整体拟合指数，如NNFI/TLI、CFI、RMSEA 和χ^2等
- 含较多因子的复杂模型中，无论是否删去某一两个路径（固定它们为0），对整个模型拟合影响不大
- 应当先检查每一个测量模型 measurement model

- D.**模型修正（model modification）**
  - 依据理论或有关假设，提出一个或数个合理的先验模型a priori model
  - 检查潜变量（因子）与指标（题目）间的关系，建立测量模型measurement model
  - 可能增删或重组题目
  - 若用同一样本数据去修正重组测量模型，再检查新模型的拟合指数，这十分接近探索性因素分析（exploratory factor analysis，EFA），所得拟合指数，不足以说明数据支持或验证模型
  - 可以循序渐进地，每次只检查含2个因子的模型，确立测量模型部分的合理后，最后才将所有因子合并成预设的先验模型，作一个总体检查
  - 对每一模型，检查标准误、*t*值、标准化残差std residuals、修正指数Modification index MI、参数期望改变值expected change、及各种拟合指数fit index，据此修改模型并重复步骤。
  - 这最后的模型是依据某一个样本数据修改而成，最好用另一个独立样本，交互确定cross-validate

参数估计和拟合函数

Parameter Estimation and Fit Function

- 目标是找一些参数parameter使得再生/隐含implied/reproduced 协方差矩阵 $\Sigma(\theta)$与样本协方差矩阵 S “差距”最小
- 拟合透過拟合函数fit function
- 多种拟合函数，参数估计值可能不同
  - 工具变量 (IV, instrumental variable)；
  - 两阶段最小二乘 (TSLS, two-stage least squares)；
  - 无加权最小二乘 (ULS, unweighted least squares)；
  - 最大似然 (ML, maximum likelihood)；
  - 广义最小二乘 (GLS, generalized least squares)；
  - 一般加权最小二乘 (WLS, generally weighted least sq)
  - 对角加权最小二乘 (DWLS, diagonally weighted least sq)

结构方程模型及其应用
Structural Equation Model and Its Applications

XI 专题: 涉及数据的问题
XI Issues on data

侯傑泰(**侯杰泰**)
Kit-Tai Hau
香港中文大学教育心理系
Educational Psychology Dept,
The Chinese University of Hong Kong
使用时请着明出处

专题: 涉及数据的问题
issues on data

- 样本容量Sample Size
  - 样本:愈大愈好
  - 每个因子上多设计几题，预试协助删去一些不好的题目
  - 最后每个因子应有3个或更多的题目
- 数据类型Data Type
  - 绝大部份分析基于皮尔逊（Pearson）相关(假設等比/等距数据interval/ratio data)
  - 来自等级（顺序）量表（ordinal scale），改用多项（polyserial）相关系数，并与渐近方差矩阵(asymptotical covariance matrix，ACM)合用，以WLS法拟合模型，除非N很大，额外需要的ACM矩阵多不稳定

- 可否应用相关矩阵作分析？
  - SEM建立在方差和协方差分析上
  - 用相关矩阵，大多数情况下正确
  - 在某些况下并不正确(见Cudeck, 1989):
    - 限制因子方差为 1，同时限制某指标的因子负荷不等于零
    - 同一个因子，限制其两个或以上指标的因子负荷，不等于零
    - 同一个因子的两个或以上指标，限制其因子负荷相同
    - 不同因子的两个或以上指标，限制其因子负荷相同
    - 限制两个或以上内生潜变量的误差相等

专题:涉及模型拟合fitting的问题

- 忽略测量误差measurement error所引致的错误
  - 方差（变异量）variance
    - x变异量var＝ξ 变异量＋δ 误差变异量
    - 除非 δvar 等于零，传统统计高估了变量的真正变异量
  - 相关correlation和回归参数regression coefficient

$$\gamma^* = \gamma \ (\mathrm{var}(\xi)/\mathrm{var}(x))$$

$$r_{\xi\eta} = r_{xy} \Big/ \left(r_{xx} r_{yy}\right)^{1/2}$$

$$\boxed{r_{\xi\eta} = 0.5/(0.7^2)^{1/2} = 0.71}$$

- 单指标潜变量(single indicator)
  - 不能同时估计loading 与uniqueness
  - 对相关矩阵 correlation matrix

    Fix both: loading, uniqueness

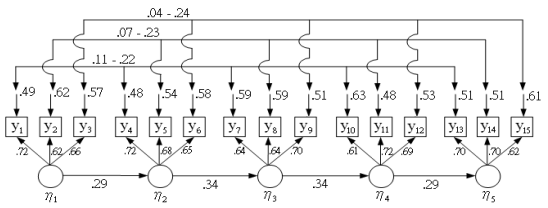
    Set value of uniqueness at .15

    (1-0.85)=0.15

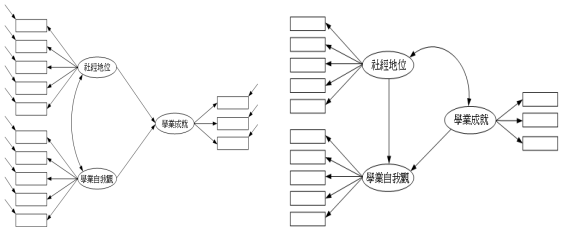
    Set value of lambda .922

    SQRT(.85)=.922

- 误差相关correlated uniqueness
  - 只在特殊设计 (重复测量multi-wave panel)，刻意容许误差相关
  - 在一般研究，通常不容许误差可以相关



- 为甚么要考虑等同模型equivalent models？
  - 以同样个数的参数（t），用不同组合产生许多不同模型，而其中再生协方差矩阵，完全相同
  - 换句话说，同样个数的参数（t）产生多个与样本数据有相同拟合程度、但结构不同的模型

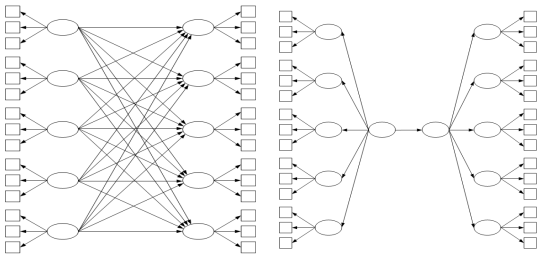


KT HAU SEM p. 187

- 结构方程是否验证变量间的因果关系causal relations？
  - 严格来说，非经设计用以探讨变量间因果效应的研究，都不能证明变量间是否真正存在因果关系。单从等同模型，已经可以举出拟合指数相同，但变量间效应相反的例子
  - 利用非实验设计non-experimental design:
    - 采用纵贯longitudinal研究数据，每个变量至少要有2次测量（2时段以上设计）
    - 使用多个指标multiple indicator以推算潜变量
    - 样本要够大并具代表性，使结果具有实质意义和普遍性
    - 考虑不同模型的意义，考虑指标误差项相关的意义

KT HAU SEM p. 188

- 合宜和错误的高阶因子(higher-order factor)
  - 不一定可以强将数个因子合并，并简化为高阶因子的关系
  - 例 :学生的性格(不可合并)如何影响学生成绩表现(或可合并)



KT HAU SEM p. 189