

Moving from 1st Generation technique to 2nd Generation analysis technique



Professor T. Ramayah
Room 118, Level 1,
School of Management,
Universiti Sains Malaysia,
11800 Minden,
Penang, Malaysia.
Tel: 604-653 3888 ext 3889
Fax: 604-657 7448
Email: ramayah@usm.my
ramayah@gmail.com

T. Ramayah Online - Windows Internet Explorer

http://www.ramayah.com/

File Edit View Favorites Tools Help

★ Favorites T. Ramayah Online Gmail - Inbox (30) - ramayah...

Page Safety Tools

T. RAMAYAH

academician & researcher

The image shows a large, empty lecture hall with rows of green upholstered seats and a wooden podium at the front. The room has a high ceiling with recessed lighting and wooden paneling on the walls.

start Micros... Vodafo... T. Ram... untile...

11:30 AM

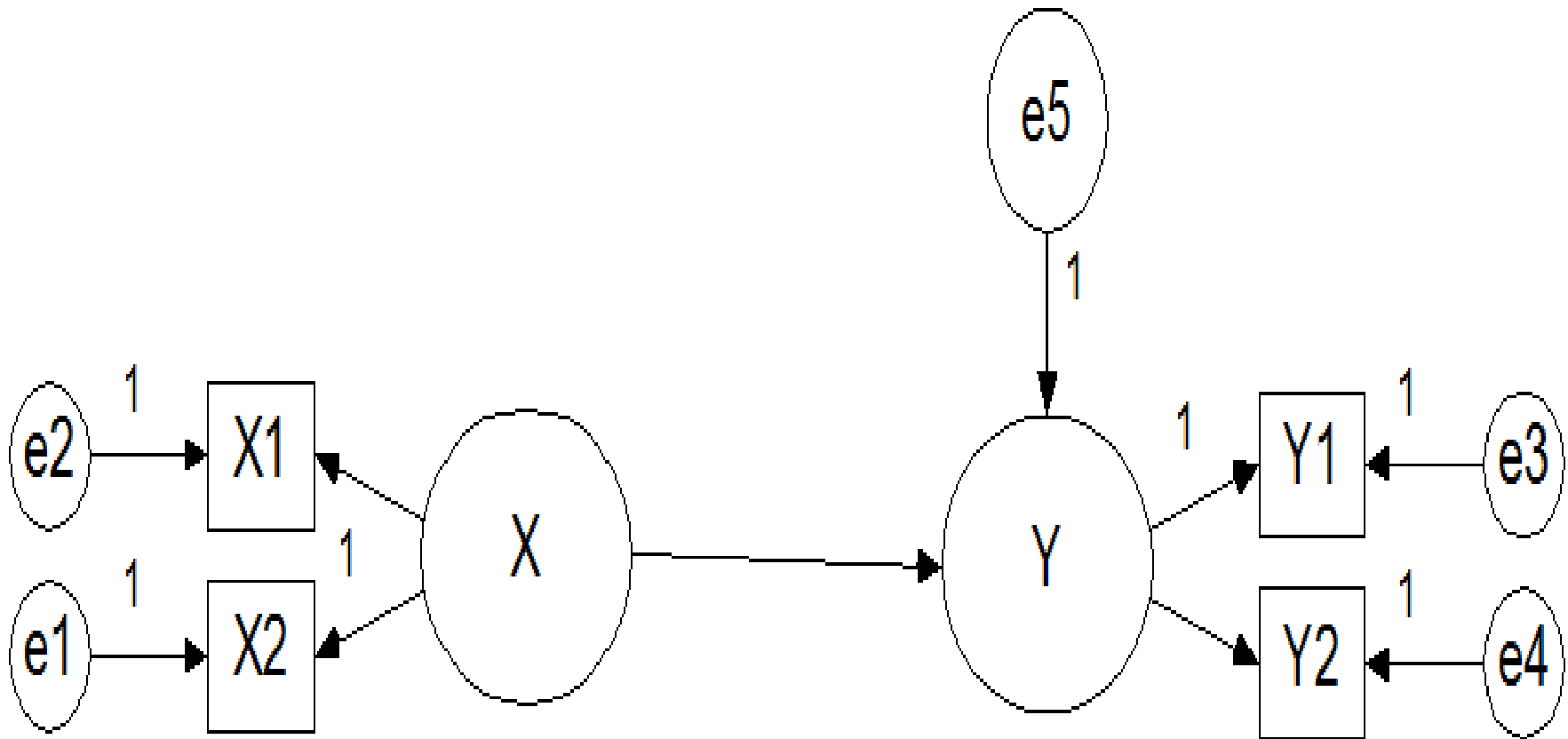
1st vs 2nd Generation Technique

	Primarily exploratory	Primarily confirmatory
1st Generation Techniques	<ul style="list-style-type: none">• multiple regression• logistic regression• analysis of variance• cluster analysis• exploratory factor analysis	<ul style="list-style-type: none">• correspondence analysis
2nd Generation Techniques	PLS-SEM	CB-SEM, including CFA

Structural Equation Modeling

- **Structural Equation Modeling . . .** is a family of statistical models that seek to explain the relationships among multiple variables.
- It examines the “structure” of interrelationships expressed in a series of equations, similar to a series of multiple regression equations.
- These equations depict all of the relationships among constructs (the dependent and independent variables) involved in the analysis.
- Constructs are unobservable or latent factors that are represented by multiple variables.
- Called 2nd Generation Techniques

Structural Equation Modeling



Distinguishing Features of SEM

- Compared to 1st Generation Techniques
 - It takes a confirmatory rather than exploratory
 - Traditional methods incapable of either assessing or correcting for measurement errors
 - Traditional methods use **observed** variables, SEM can use both unobserved (latent) and observed variables
 - Testing in one complete model

Components of Error

- Observed score comprises of 3 components (**Churchill, 1979**)
 - True score
 - Random error (ex; caused by the order of items in the questionnaire or respondent fatigue) (**Heeler & Ray, 1972**)
 - Systematic error such as method variance (ex; variance attributable to the measurement method rather than the construct of interest) (**Bagozzi et al., 1991**)

Structural Equation Modeling Defined

- Exogenous constructs are the latent, multi-item equivalent of independent variables. They use a variate (linear combination) of measures to represent the construct, which acts as an independent variable in the model.
 - Multiple measured variables (x) represent the exogenous constructs.
- Endogenous constructs are the latent, multi-item equivalent to dependent variables. These constructs are theoretically determined by factors within the model.
 - Multiple measured variables (y) represent the endogenous constructs.

SEM - Variations



- **CB-SEM (Covariance-based SEM)** – objective is to reproduce the theoretical covariance matrix, without focusing on explained variance.
- **PLS-SEM (Partial Least Squares SEM)** – objective is to maximize the explained variance of the endogenous latent constructs (dependent variables).

Selection

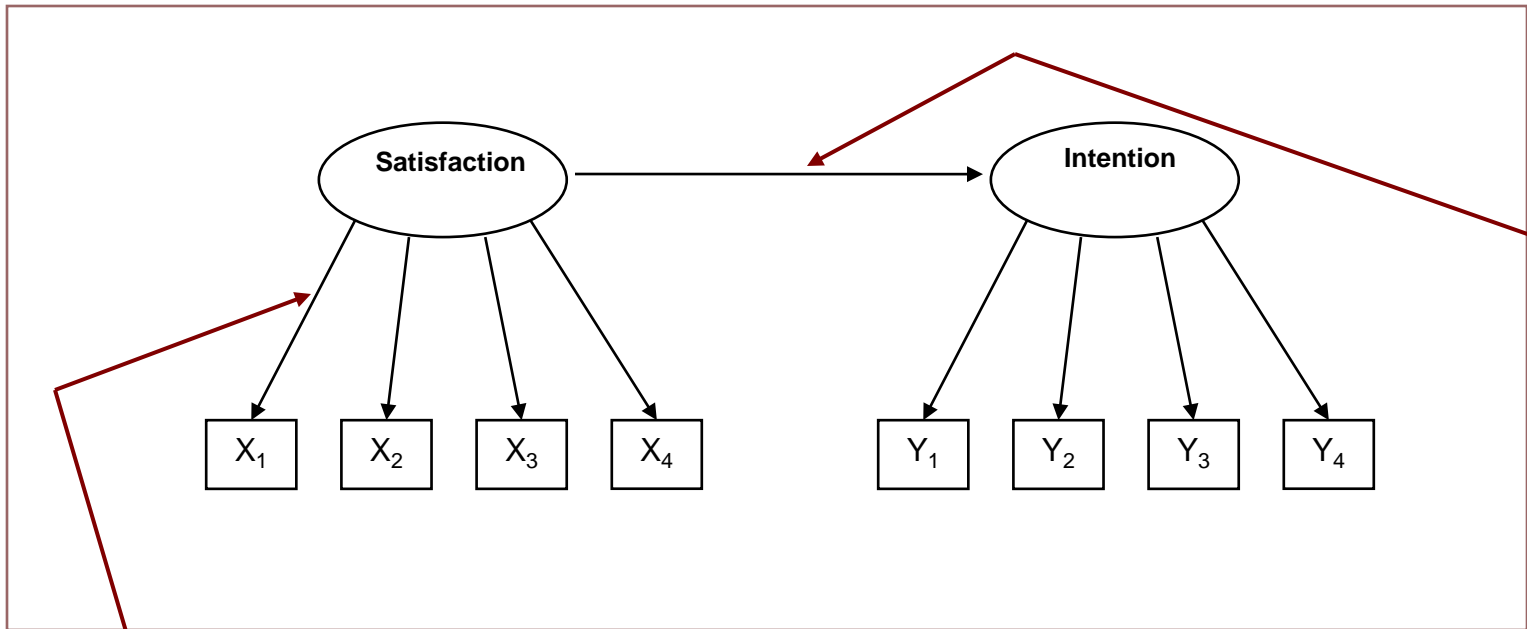
- The decision between these approaches is whether to use SEM for theory testing and development or for predictive applications (Anderson & Gerbing, 1988)
- In situations where prior theory is strong and further testing and development are the goal, covariance-based full-information estimation methods are more appropriate.

Two approaches to SEM

● Covariance based

- EQS, <http://www.mvsoft.com/>
- AMOS, <http://www-01.ibm.com/>
- SEPATH, <http://www.statsoft.com/>
- LISREL, <http://www.ssicentral.com/>
- MPLUS, <http://www.statmodel.com/>
- lavaan, <http://lavaan.ugent.be/>
- Onyx, <http://onyx.brandmaier.de/>

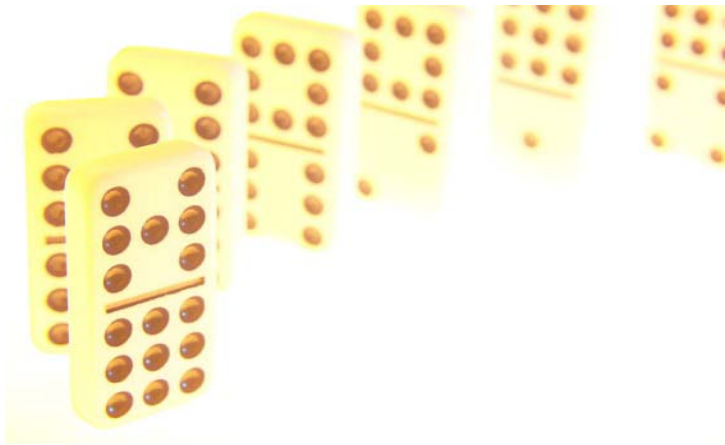
Two Latent Constructs and the Measured Variables



- Loadings represent the relationships from constructs to variables as in factor analysis.
- Path estimates represent the relationships between constructs as does B in regression analysis.

Establishing Causation – “Causal Modeling”

- Causal Inference – Hypothesizes a “cause-and-effect” relationship.



4 Types of Evidence

1. Covariation
2. Sequence
3. Nonspurious Covariance
4. Theoretical Support

Basics of SEM Estimation

- SEM explains the observed covariance among a set of measured variables:
 - It does so by estimating the observed covariance matrix with an estimated covariance matrix constructed based on the estimated relationships among variables.

Observed Covariance Matrix	Estimated Covariance Matrix
S	Σ_k

- The closer these are, the better the fit. When they are equal, the fit is perfect.

AMOS Data Input = observed sample covariances for HBAT 3-Construct model

Sample Covariances (Group number 1)

	AC1	AC2	AC3	AC4	OC4	OC3	OC2	OC1	EP4	EP3	EP2	EP1
AC1	1.937											
AC2	1.615	2.972										
AC3	1.364	1.687	2.009									
AC4	1.501	1.860	1.540	2.587								
OC4	.432	.601	.601	.804	4.206							
OC3	.327	.271	.302	.453	2.035	3.071						
OC2	.803	.794	.841	1.011	3.306	2.137	4.768					
OC1	.198	.452	.340	.524	2.506	1.961	2.884	6.360				
EP4	.264	.379	.247	.451	.936	.901	1.008	.803	1.936			
EP3	.345	.300	.286	.398	.792	.735	1.013	.412	1.235	1.777		
EP2	.391	.492	.378	.592	1.259	1.055	1.326	.938	1.469	1.331	2.644	
EP1	.357	.501	.406	.450	1.046	.757	1.005	.514	1.442	1.228	1.773	3.344

Covariances calculated for the sample – request Sample moments and look in Output under that subheading.

Variances are on the diagonal and covariances are off the diagonal.

Implied Covariances (Group number 1 - Default model)

	AC1	AC2	AC3	AC4	OC4	OC3	OC2	OC1	EP4	EP3	EP2	EP1
AC1	1.937											
AC2	1.619	2.972										
AC3	1.357	1.677	2.009									
AC4	1.502	1.857	1.556	2.587								
OC4	.599	.741	.621	.688	4.206							
OC3	.407	.503	.422	.467	2.025	3.071						
OC2	.660	.816	.684	.757	3.283	2.229	4.768					
OC1	.518	.641	.537	.595	2.578	1.751	2.838	6.360				
EP4	.333	.412	.345	.382	.990	.672	1.090	.856	1.936			
EP3	.299	.370	.310	.343	.889	.604	.979	.769	1.181	1.777		
EP2	.383	.473	.397	.439	1.138	.773	1.253	.984	1.512	1.358	2.644	
EP1	.367	.454	.380	.421	1.091	.741	1.201	.943	1.449	1.301	1.666	3.344

Covariances estimated by AMOS software – request Implied moments and look in Output under Estimates.

Residual Covariances (Group number 1 - Default model)

	AC1	AC2	AC3	AC4	OC4	OC3	OC2	OC1	EP4	EP3	EP2	EP1
AC1	.000											
AC2	-.003	.000										
AC3	.007	.010	.000									
AC4	-.001	.003	-.016	.000								
OC4	-.167	-.140	-.020	.116	.000							
OC3	-.080	-.233	-.119	-.014	.010	.000						
OC2	.143	-.022	.158	.254	.023	-.093	.000					
OC1	-.320	-.189	-.197	-.071	-.072	.211	.046	.000				
EP4	-.069	-.033	-.098	.069	-.054	.229	-.081	-.052	.000			
EP3	.046	-.070	-.024	.055	-.097	.131	.034	-.357	.054	.000		
EP2	.008	.019	-.018	.153	.121	.282	.073	-.046	-.043	-.026	.000	
EP1	-.010	.048	.026	.030	-.044	.016	-.196	-.429	-.007	-.073	.107	.000

Residuals = difference between observed and estimated covariances – request Residual moments.

A negative sign indicates the observed covariance (2.137) is smaller than the estimated covariance (2.229) by -.093.

Standardized Residual Covariances (Group number 1 - Default model)

	AC1	AC2	AC3	AC4	OC4	OC3	OC2	OC1	EP4	EP3	EP2	EP1
AC1	.000											
AC2	-.024	.000										
AC3	.057	.065	.000									
AC4	-.007	.020	-.116	.000								
OC4	-1.145	-.773	-.134	.689	.000							
OC3	-.645	-1.517	-.947	-.095	.048	.000						
OC2	.917	-.112	.993	1.413	.082	-.419	.000					
OC1	-1.803	-.857	-1.090	-.344	-.250	.885	.148	.000				
EP4	-.699	-.268	-.980	.609	-.354	1.809	-.504	-.290	.000			
EP3	.490	-.597	-.246	.505	-.674	1.087	.221	-2.066	.492	.000		
EP2	.071	.130	-.157	1.154	.687	1.908	.388	-.218	-.316	-.206	.000	
EP1	-.079	.298	.199	.199	-.227	.099	-.940	-1.819	-.047	-.528	.626	.000

Standardized Residuals – you look for patterns of larger residuals, generally => 4.0

Structural Equation Modeling

- No model should be developed for use with SEM without some underlying theory. Theory is needed to develop both the . . .
 - Measurement model specification.
 - Structural model specification.

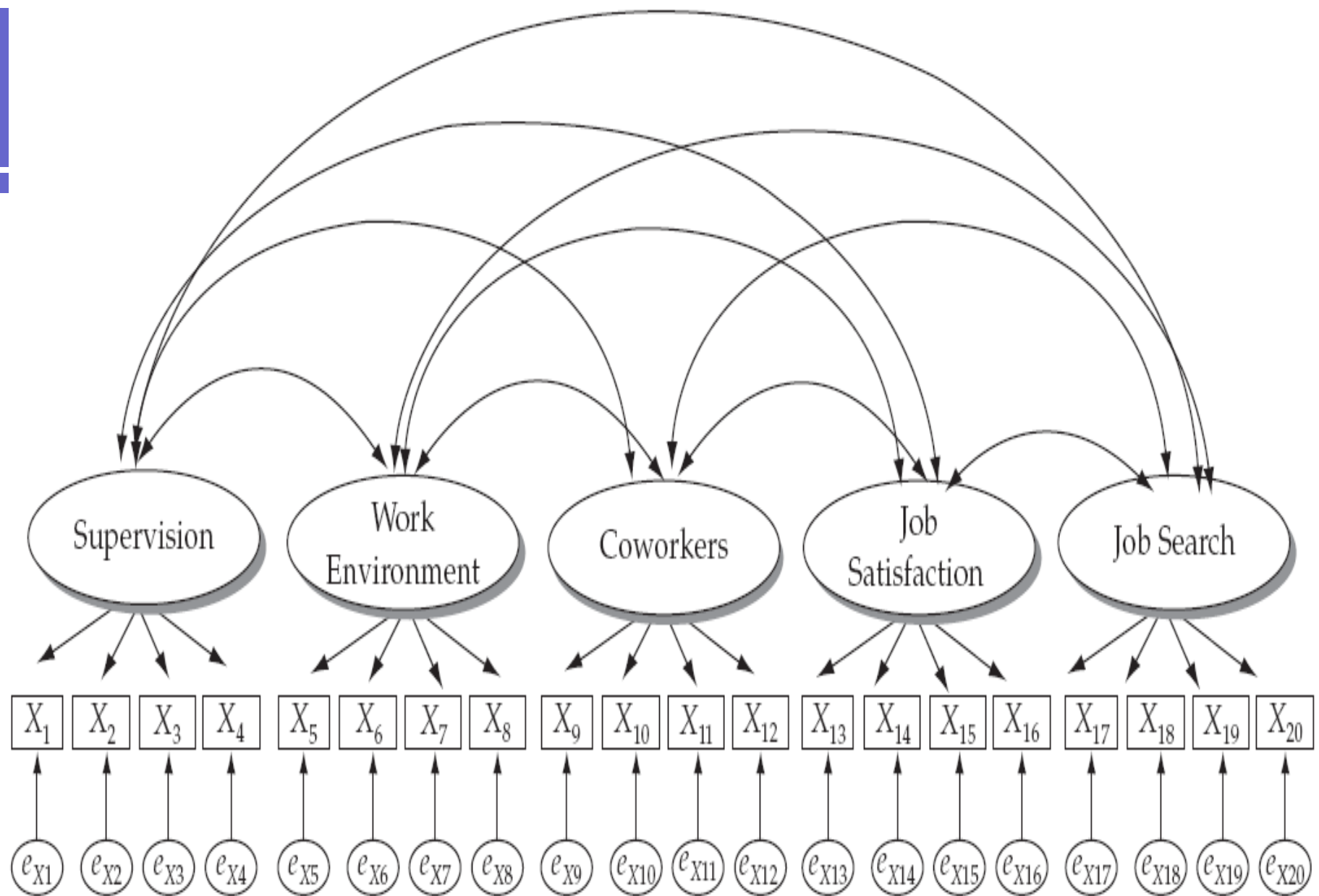


FIGURE 12-9 A Path Diagram Showing Hypothesized Measurement Model Specification (CFA Model)

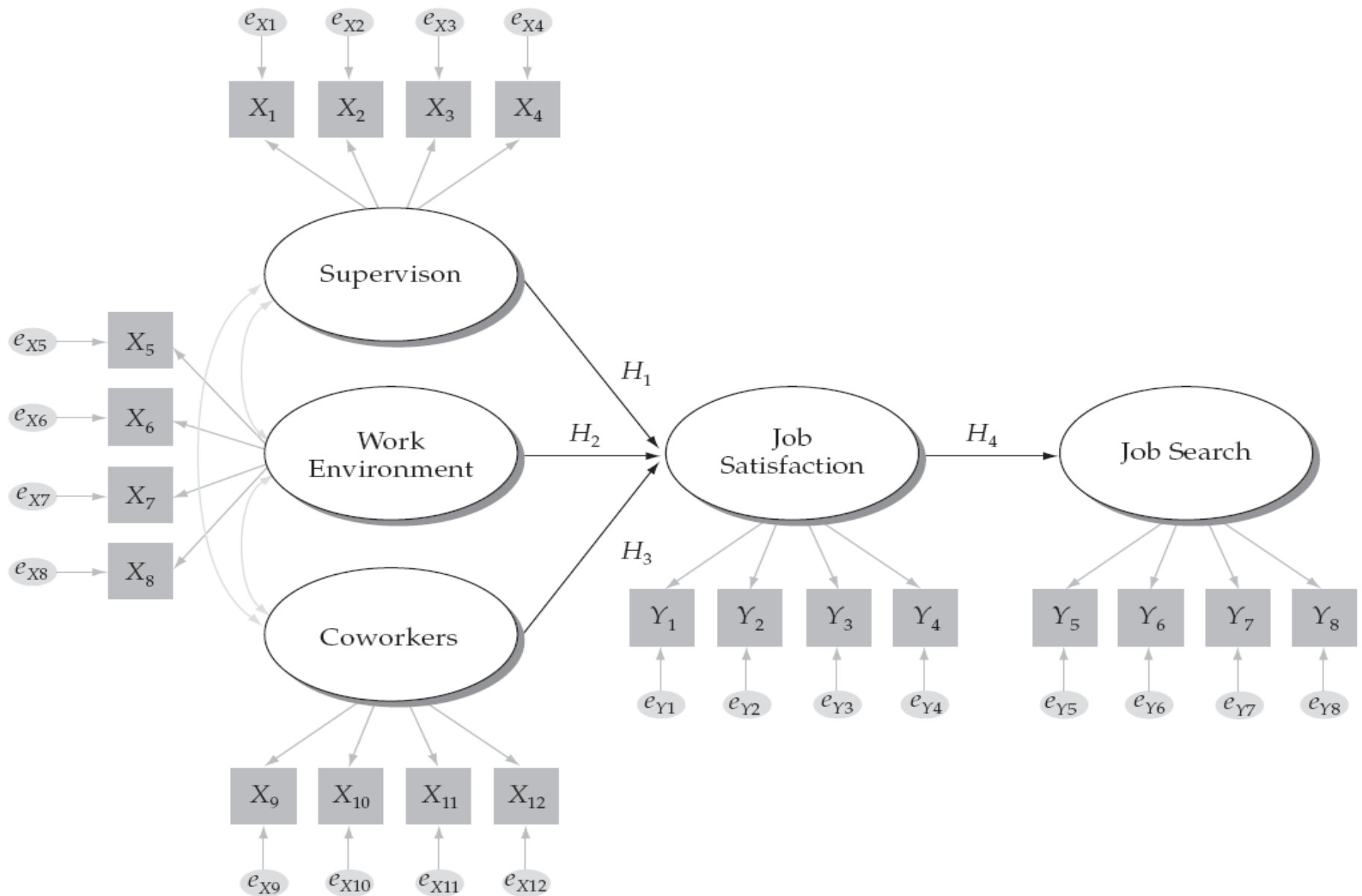


FIGURE 12-10 A Path Diagram Showing Specified Hypothesized Structural Relationships and Measurement Specification

Missing Value Imputation

- Traditional

- No replacement
- Mid point of the scale
- Random number
- Mean value of the other respondents
- Mean value of the other responses

- Current

- FIML
- EM
- MI

Missing Value Imputation

Missing Data Imputation.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Quality Control
ROC Curve...
Amos 19...

Missing Value Analysis

Quantitative Variables:
ID
V1
V2
V3
V4
V5
V6
V7
V8
V9
V10
V11
V12
V13
V14
V15

Categorical Variables:

Maximum Categories: 25
Case Labels:

Use All Variables

Patterns...
Descriptives...
Estimation
Listwise
Pairwise
☒ EM
Regression
Variables...
EM...
Regression...

Missing Value Analysis: EM

Distribution
☐ Normal
☐ Mixed normal
Mixture proportion:
Standard deviation ratio:
☐ Student's t
Degrees of freedom:
Maximum iterations: 25
☒ Save completed data
☐ Create a new dataset
Dataset name:
☐ Write a new data file
File... C:\Documents a...No Missing.sav
Continue Cancel Help

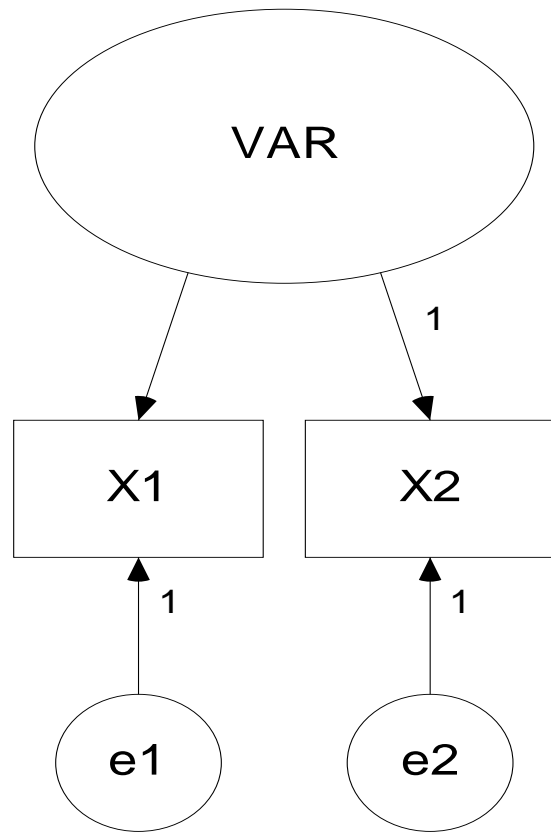
Missing Value Analysis: Save to File

Look in: DEAKIN FINAL EXERCISES & DATA
Outliers and Bootstrapping
Data 6.sav
CFA Data.sav
Data No Missing.sav
Data 1.sav
Data 2.sav
EFA.sav
Data 3.sav
Missing Data Imputation.sav
Data 4.sav
No Missing.sav
Data 5.sav
Self Exercise.sav
File name: No Missing.sav
Save as type: SPSS Statistics (*.sav)
Save
Cancel
Help
Store File To Repository...

IBM SPSS Statistics Processor is ready

start Doctorate Support Gr... DEAKIN FINAL EXERC... Microsoft PowerPoint... Missing Data Imputati... EN 11:29 AM

Under-identified Model – 2 items



Bits of Information = $\frac{1}{2} [p(p + 1)]$

Where p = number of measured items

S

X1

X2

X1

X2

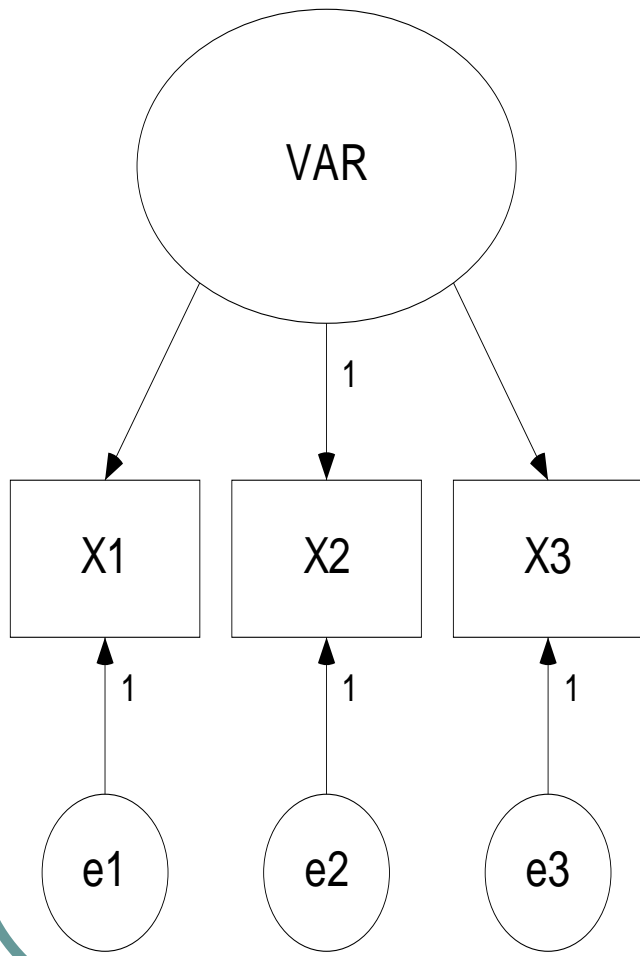
var(1)

cov(2,1)

cov(1,2)

var(2)

Just-identified Model – 3 items



S

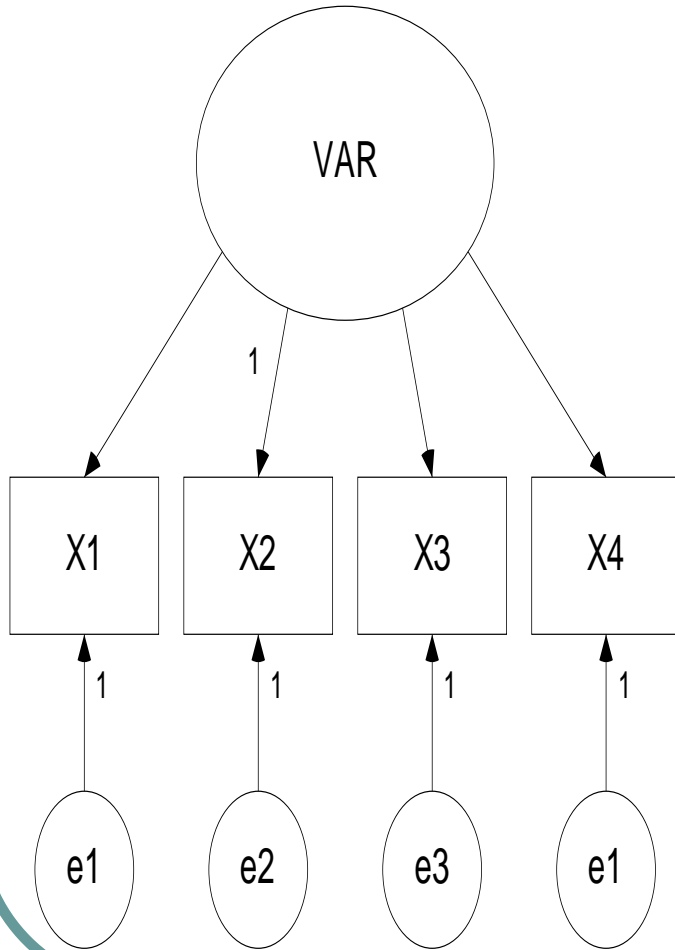
X1

X2

X3

X1	X2	X3
var(1)	cov(1,2)	cov(1,3)
cov(1,2)	var(2)	cov(2,3)
cov(1,3)	cov(2,3)	var(3)

Over-identified Model – 4 items



S

X1

X2

X3

X4

X1

X2

X3

X4

var(1) cov(1,2) cov(1,3) cov(1,4)

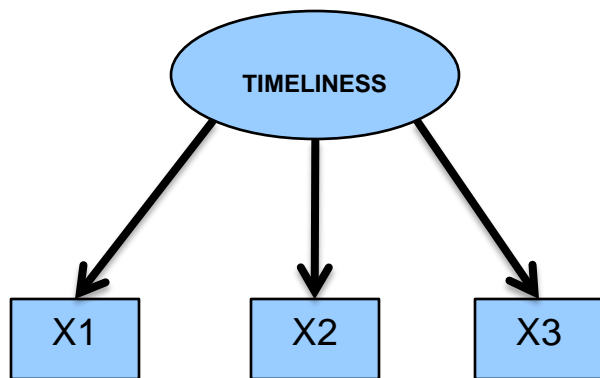
cov(1,2) **var(2)** cov(2,3) cov(2,4)

cov(1,3) **cov(2,3)** **var(3)** cov(3,4)

cov(1,4) **cov(2,4)** **cov(3,4)** **var(4)**

Indicators

- Reflective



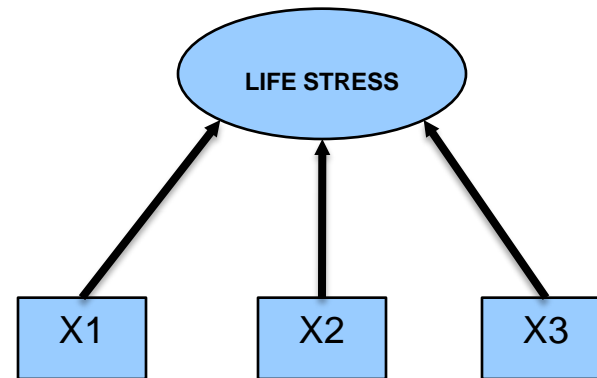
X1 = Accommodate last minute request

X2 = Punctuality in meeting deadlines

X3 = Speed of returning phone calls

- Indicators must be highly correlated (**Hulland, 1999**)

- Formative



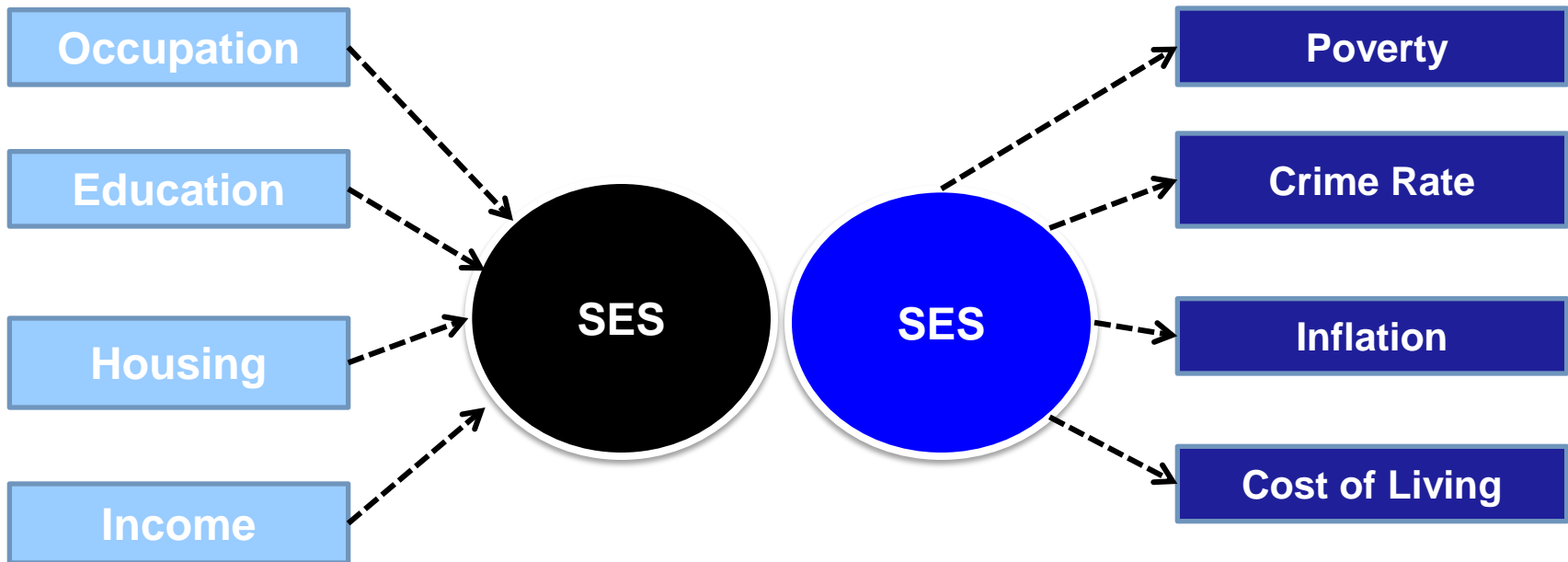
X1 = Job loss

X2 = Divorce

X3 = Recent accident

- Indicators can have +, - or 0 correlation (**Hulland, 1999**)


Example – Measuring SES



Problems in Specification

Reflective measurement is most commonly used but in many cases a formative measurement would be appropriate

	Should be reflective	Should be formative	Total
Modelled as reflective	947 (65%)	456 (31%) <small>(Type I error)</small>	1403 (96%)
Modelled as formative	17 (1%) <small>(Type II error)</small>	41 (3%)	58 (4%)
Total	964 (66%)	497 (34%)	1461 (100%)



32% of constructs have been measured incorrectly

Data bases are the Top 3 German- and Top 4 English-language journals:

- JARVIS/BURKE/PODSAKOFF (2003): Journal of Consumer Research, Journal of Marketing, Journal of Marketing Research, Marketing Science (1977 – 2000): N = 1,192
- FASSOTT (2006): Zeitschrift für betriebswirtschaftliche Forschung, Zeitschrift für Betriebswirtschaft, Die Betriebswirtschaft (X - 2003): N = 269

Absolute Fit Measures

Fit Indices	Acceptable Values	Source
Goodness-of-Fit Index (GFI)	≥ 0.90	Chau & Hu (2001)
Root Mean Square Error Approximation (RMSEA)	≤ 0.08	Brown and Cudeck (1993)
Root Mean Square Residual (RMR)	≤ 0.08	Brown and Cudeck (1993)
Standardized Root Mean Residual (SRMR)	≤ 0.08	Hu and Bentler (1999)
χ^2/df	≤ 3.0	Bagozzi & Yi (1988)

Incremental Fit Indices

Fit Indices	Acceptable Values	Source
Normed Fit Index (NFI)	≥ 0.90	Bentler and Bonnet (1980)
Non-normed Fit Index (NNFI) (TLI)	≥ 0.90	Bentler and Bonnet (1980)
Comparative Fit Index (CFI)	≥ 0.90	Bagozzi & Yi (1988)
Relative Fit Index (RFI)	≥ 0.90	Anderson and Gerbing (1988)

Parsimony Fit Indices

Fit Indices	Acceptable Values	Source
Adjusted Goodnes-of-Fit Index (AGFI)	≥ 0.80	Chau & Hu (2001)
Parsimony Normed fit Index (PNFI)	≥ 0.80	

Measurement Model and Construct Validity

- One of the biggest advantages of CFA/SEM is its ability to assess the **construct validity** of a proposed measurement theory. **Construct validity** . . . is the extent to which a set of measured items actually reflect the theoretical latent construct they are designed to measure.
- Construct validity is made up of two important components:
 1. Convergent validity – three approaches:
 - Factor loadings.
 - Variance extracted.
 - Reliability.
 2. Discriminant validity

Internal Consistency (Cronbach α)

$$\text{Cronbach's alpha : } \alpha = \left(\frac{N}{N-1} \right) * \left(1 - \frac{\sum_{i=1}^N \sigma_i^2}{\sigma_t^2} \right)$$

N = number of indicators assigned to the factor

σ_i^2 = variance of indicator i

σ_t^2 = variance of the sum of all assigned indicators' scores

j = flow index across all reflective measurement model

- Measures the reliability of indicators
- The value is between 0 and 1
- In early phase 0.7 acceptable, but in later phases values of 0.8 or 0.9 is more desirable (Nunnally, 1978)

Internal Consistency (Dhillon-Goldstein Rho)

$$\text{Composite reliability}(\rho) = \frac{(\sum_i \lambda_{ij})^2}{(\sum_i \lambda_{ij})^2 + \sum_i \text{var}(\varepsilon_{ij})}$$

λ_i = loadings of indicator i of a latent variable

ε_i = measurement error of indicator i

j = flow index across all reflective measurement model

- Measures the reliability of indicators
- The value is between 0 and 1
- Composite reliability should be 0.7 or higher to indicate adequate convergence or internal consistency (Gefen et al., 2000).

Average Variance Extracted (AVE)

$$AVE = \frac{\sum_i \lambda_i^2}{\sum_i \lambda_i^2 + \sum_i \text{var}(\epsilon_i)}$$

λ_i^2 = squared loadings of indicator i of a latent variable

$\text{var}(\epsilon_i)$ = squared measurement error of indicator i

- Comparable to the proportion of variance explained in factor analysis
- Value ranges from 0 and 1.
- AVE should exceed 0.5 to suggest adequate convergent validity (Bagozzi & Yi, 1988; Fornell & Larcker, 1981).

Discriminant Validity

- Fornell & Larcker (1981) criterion
 - A latent variable should explain better the variance of its own indicators than the variance of other latent variables
 - The AVE of a latent variable should be higher than the squared correlations between the latent variable and all other variables. (Chin, 2010; Chin 1998b; Fornell & Larcker, 1981).
- Cross loadings
 - The loadings of an indicator on its assigned latent variable should be higher than its loadings on all other latent variables.

Discriminant Validity

- The **square root of the Average Variance Extracted (AVE)** that exceeds the intercorrelations of the construct with the other constructs in the model to ensure discriminant validity (Chin, 2010; Chin 1998b; Fornell & Larcker, 1981).
- Example:

TABLE 8. Mean, standard deviation, intercorrelations of the latent variables for the first-order constructs.

Construct	Mean	SD	Ability	Benevolence	Integrity	Predictability	Trust	Continuance
<i>Ability</i>	5.465	1.186	0.925*					
<i>Benevolence</i>	5.745	1.044	0.715	0.850*				
<i>Integrity</i>	5.080	1.319	0.693	0.684	0.915*			
<i>Predictability</i>	5.595	1.187	0.681	0.612	0.665	0.915*		
<i>Trust</i>	5.378	1.242	0.799	0.798	0.698	0.680	0.912*	
<i>Continuance</i>	5.184	1.605	0.756	0.684	0.683	0.626	0.762	0.949*

*Square root of the AVE on the diagonal.

Reporting Measurement Model

Model Construct	Measurement Item	Loading	CR ^a	AVE ^b
Commitment	COMMIT1	0.686	0.856	0.601
	COMMIT2	0.767		
	COMMIT3	0.885		
	COMMIT4	0.751		
Communication	COMMUN1	0.842	0.873	0.696
	COMMUN2	0.831		
	COMMUN3	0.829		
Trust	TRUST1	0.580	0.759	0.527
	TRUST2	0.587		
	TRUST3	0.948		
Performance	PERFORM1	0.837	0.898	0.747
	PERFORM2	0.900		
	PERFORM2	0.853		

Specifying the Structural Model

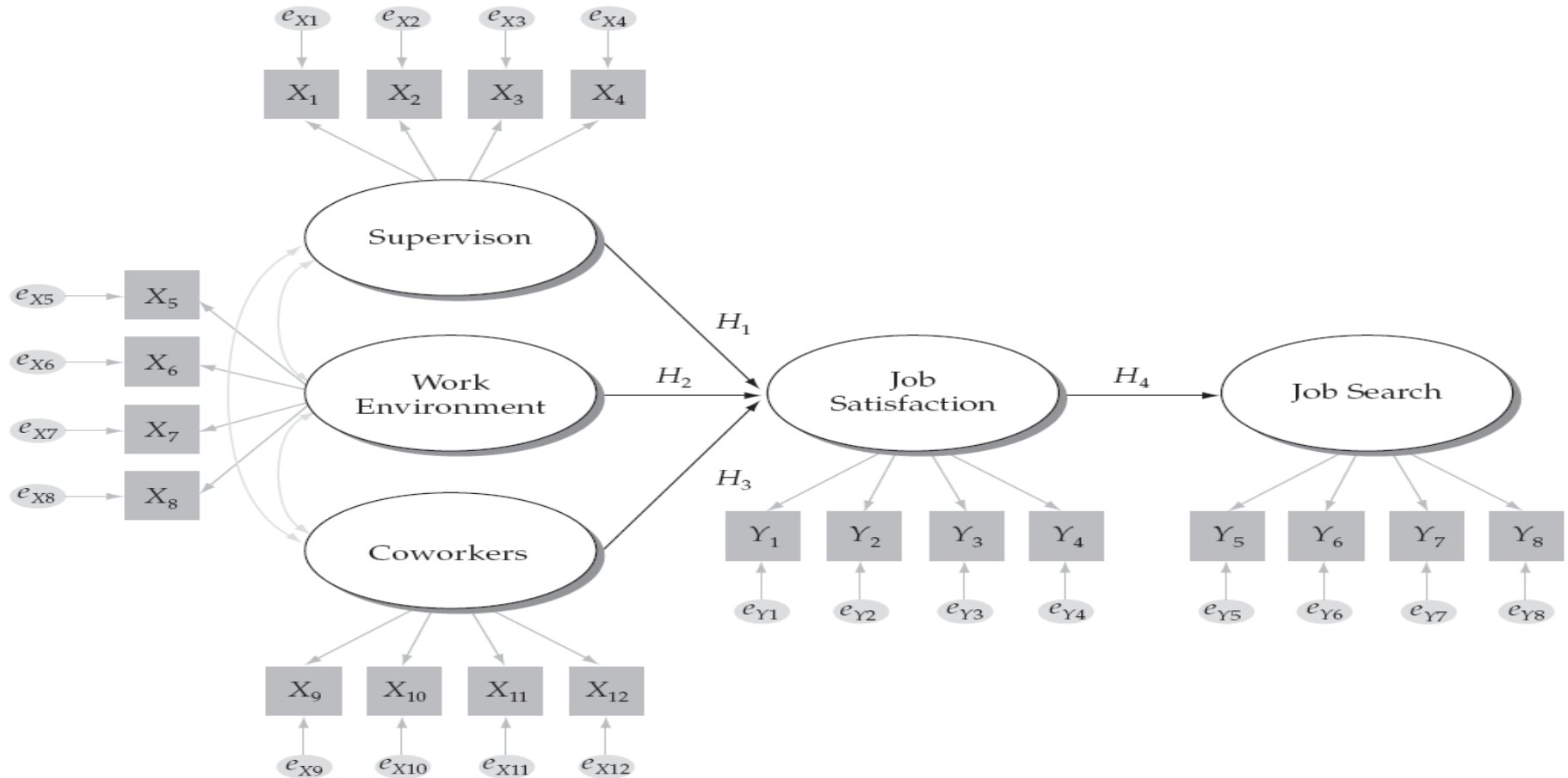


FIGURE 12-10 A Path Diagram Showing Specified Hypothesized Structural Relationships and Measurement Specification

Presenting Results

Table 4: Hypotheses testing

Hypothesis	Critical ratios (CR)	p-value	Decision
H1: System quality has a positive relationship with user satisfaction.	3.256	0.001	Supported
H2: Information quality has a positive relationship with user satisfaction.	5.399	0.000	Supported
H3: Service quality has a positive relationship with user satisfaction.	2.948	0.003	Supported
H4: User satisfaction is positively related to usage continuance.	5.069	0.000	Supported
H5: System quality is positively related to intention to use.	2.837	0.005	Supported
H6: Service quality is positively related to intention to use.	4.697	0.000	Supported

Modeling Strategy

- Confirmatory Modeling Strategy
 - Focus is on assessing the fit
- Competing Models Strategy
 - Focus on comparing the estimated model with other alternatives
- Model Development Strategy
 - Basic framework is provided
 - Improve the framework through modifications
 - Re-specification

Poor Practices

- Pursuit of fit
- Reducing number of items per construct
- Parceling of items
- Separate analysis for each construct
- Sample size
 - Representativeness
 - Generalizability

Two approaches to SEM

● Variance Based SEM

- Smart PLS, <http://www.smartpls.de/forum/>
- PLS Graph, <http://www.plsgraph.com/>
- WarpPLS, <http://www.scriptwarp.com/warppls/>
- Visual PLS, <http://fs.mis.kuas.edu.tw/~fred/vpls/start.htm>
- PLS-GUI, <http://www.rotman-baycrest.on.ca/index.php?section=84>
- SPAD-PLS, <http://spadsoft.com/content/blogcategory/15/34/>
- GeSCA, <http://www.sem-gesca.org/>



Google



Ramayah T

Gmail ▾



More ▾

7 of 45,038



COMPOSE

Trade Binary Options Join & Get Up To 50% Welcome Bonus - www.opteck.com/Malaysia - Trade Forex, Stocks & Mor...

Inbox (134)

Important

Sent Mail

Drafts (2)

Spam (14)



Cherish the moments v

Search, chat, or SMS

● Nur Muflihah
with my whole f...

🇲🇾 Azah Norman

🇲🇾 Doctorate SG

🇲🇾 nik maheran nik...

🇲🇾 normalini md ka...

Ned Kock wants to be friends on Facebook



Inbox x



Facebook

12:47 AM (8 hours ago)



to me ▾

facebook



Ned Kock wants to be friends with you on Facebook.

Works at Texas A&M International University · University of Waikato · Laredo, Texas

819 friends · 9 photos · 10 groups

Confirm Request

See All Requests

Choice

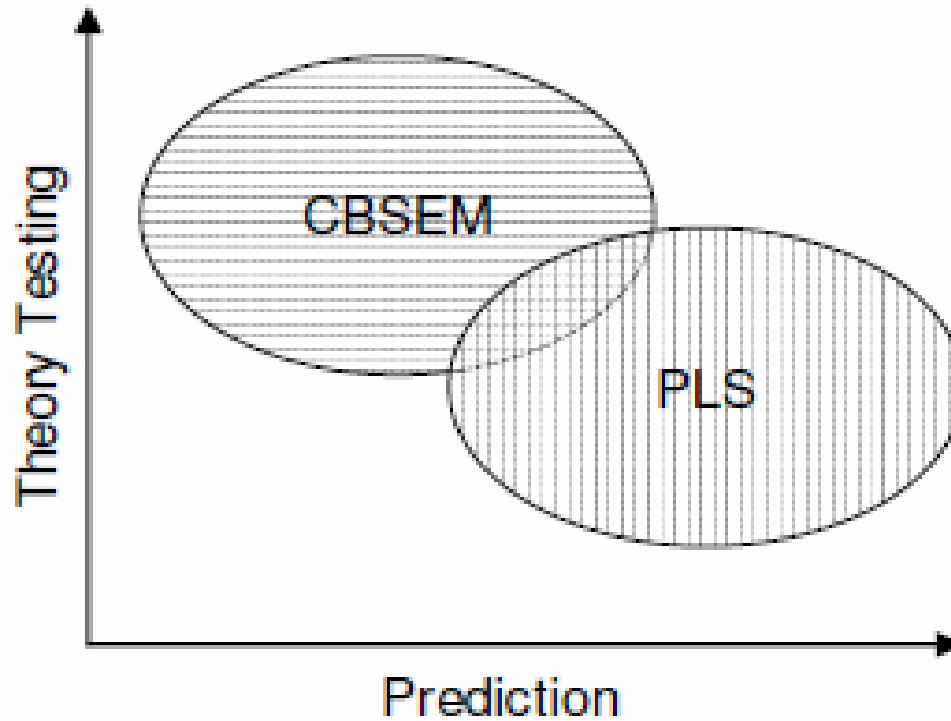


Figure 4: CBSEM vs. PLS (according to Henseler et al. 2009).

Why PLS?

- Like covariance based structural equation modeling (CBSEM), PLS is a latent variable modeling technique that incorporates multiple dependent constructs and explicitly recognizes measurement error (Karim, 2009)
- In general, two applications of PLS are possible (Chin, 1998a): It can either be used for theory confirmation or theory development. In the latter case, PLS is used to develop propositions by exploring the relationships between variables.

Reasons for using PLS

- Researchers' arguments for choosing PLS as the statistical means for testing structural equation models (Urbach & Ahleman, 2010) are as follows:
 - PLS makes fewer demands regarding sample size than other methods.
 - PLS does not require normal-distributed input data.
 - PLS can be applied to complex structural equation models with a large number of constructs.
 - PLS is able to handle both reflective and formative constructs.
 - PLS is better suited for theory development than for theory testing.
 - PLS is especially useful for prediction

Hair et al. (2013)

- PLS-SEM is advantageous when used with small sample sizes (**e.g., in terms of the robustness of estimations and statistical power; Reinartz et al., 2009**).
- However, some researchers abuse this advantage by relying on extremely small samples relative to the underlying population.
- All else being equal, the more heterogeneous the population in a structure is the more observations are needed to reach an acceptable sampling error level.

Choice

- Overall, PLS can be an adequate alternative to CBSEM if the problem has the following characteristics (Chin 1998b; Chin & Newsted 1999):
 - The phenomenon to be investigated is relatively new and measurement models need to be newly developed,
 - The structural equation model is complex with a large number of LVs and indicator variables,
 - Relationships between the indicators and LVs have to be modeled in different modes (i.e., formative and reflective measurement models),³
 - The conditions relating to sample size, independence, or normal distribution are not met, and/or
 - Prediction is more important than parameter estimation.

Incremental Study

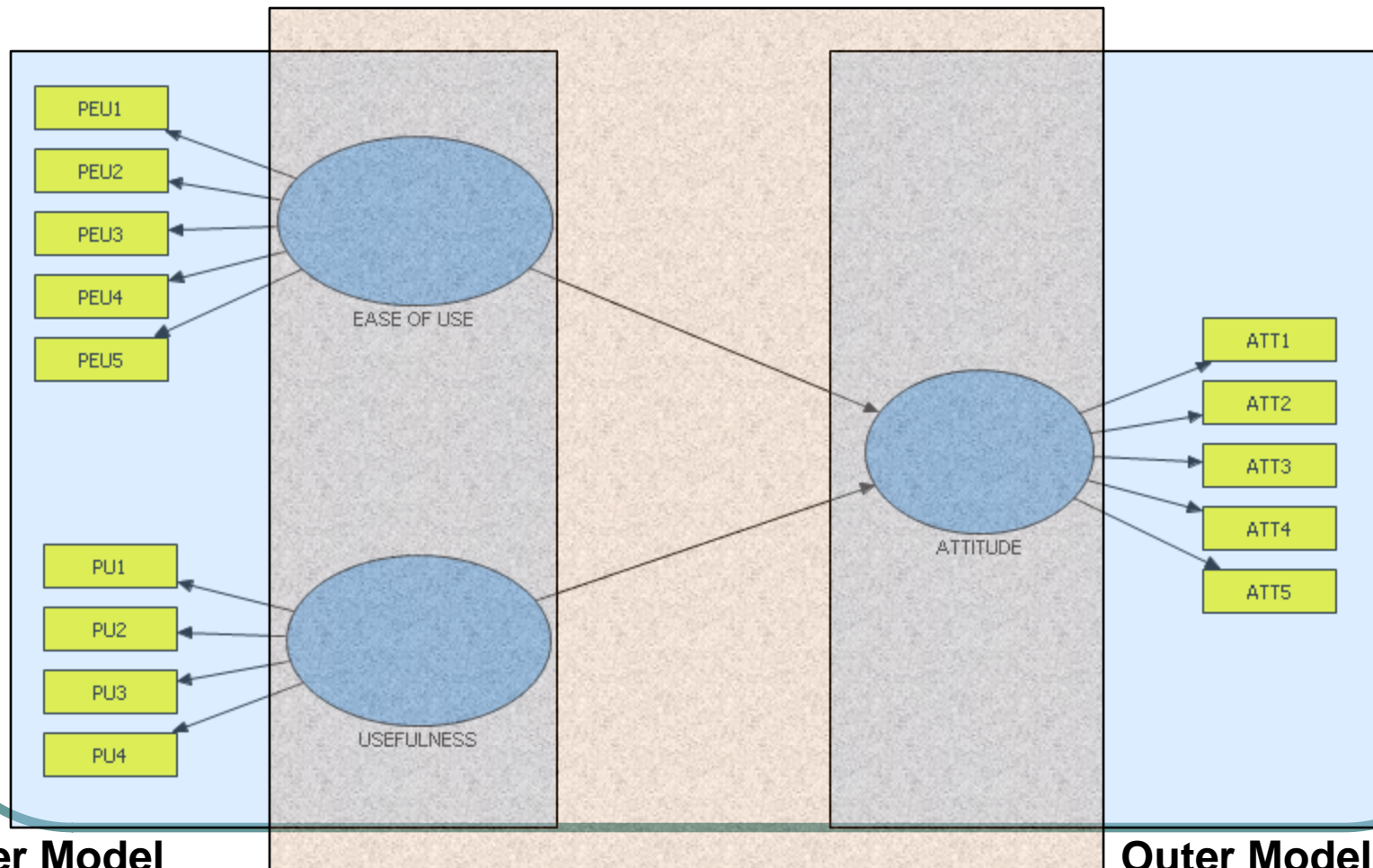
- For example, when the research has an interactive character. This is the case of an incremental study, which is initially based on a prior model but new measures and structural paths are then introduced into it.
- In this respect these statements are confirmed by the study of Reinartz et al. (2009): "**PLS is the preferable approach when researchers focus on prediction and theory development, our simulations show that PLS requires only about half as many observations to reach a given level of statistical power as does ML-based CBSEM**" (p. 334).

The 2 Step Approach

- A structural equation modeling process requires two steps:
 1. building and testing a measurement model, and
 2. building and testing a structural model.
- The measurement model serves to create a structural model including paths representing the hypothesized associations among the research constructs.

Modeling in PLS

Inner Model



**Outer Model
Exogenous**

**Outer Model
Endogenous**

Bootstrapping



Example: Bootstrapping

- Is there a correlation between IQ and a methodology re-examination result?

ID	IQ	MR
1	105	5.6
2	106	5
3	114	7.1
4	123	7.4
5	134	6.1
6	141	8.6

- $\text{Corr}(\text{IQ}, \text{MR}) = 0.733$ Is this significant?

Building the Bootstrap Samples

Sample 1

ID	IQ	MR
6	141	8.6
4	123	7.4
3	114	7.1
5	134	6.1
2	106	5.0
5	134	6.1

corr = 0.546

Sample 2

ID	IQ	MR
3	114	7.1
3	114	7.1
1	105	5.6
3	114	7.1
3	114	7.1
5	134	6.1

corr = -0.060

Sample 3

ID	IQ	MR
2	106	5.0
2	106	5.0
2	106	5.0
2	106	5.0
4	123	7.4
4	123	7.4

corr = 1.000

...

Sample 500

ID	IQ	MR
6	141	8.6
4	123	7.4
3	114	7.1
5	134	6.1
2	106	5.0
5	134	6.1

corr = 0.546

- Standard deviation of corr = 0.277

- $t = \frac{0.733}{0.277} = 2.646$

- Comparison

- $t_{0.05, 499} = 1.965$
 - $t_{0.01, 499} = 2.586$

Extending the Life of SPSS

SPSS, SAS, and Mplus Macros and Code - Andrew F. Hayes, Ph.D. - Mozilla Firefox

File Edit View History Bookmarks Tools Help


Inbox - Outlook Web App, light version x Inbox (130) - ramayah@gmail.com - ... x SPSS, SAS, and Mplus Macros and Co... x

afhayes.com/spss-sas-and-mplus-macros-and-code.html

Most Visited Getting Started Suggested Sites Web Slice Gallery

Preacher & Hayes

Andrew F. Hayes, Ph.D.



Home

My C.V.

Teaching and Speaking

Introduction to Mediation, Moderation, and Conditional Process Analysis

Statistical Methods for Communication Science

Sage Sourcebook of Advanced Data Analysis Methods for Communication Research

SPSS, SAS, and Mplus Macros and Code


Mediation and

Recommend Triantoro Safaria, Hiram Ting and 648 others recommend this.

Tweet 31

PROCESS can found found on the web page for Hayes (2013)

My Macros and Code for SPSS and SAS



One of my professional pleasures is writing specialized code for popular statistical programs that will accomplish things that the programs can't otherwise do. On this page you will find information about many of the macros I have written. Most of these are described in various publications, and I recommend you read the corresponding publication before using the macro.

Notice to SPSS 18 users: I get many emails from users of SPSS 18 who have had trouble getting my

8:57 AM 11/4/2013

Dialog Boxes & Macros

Data2.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Quality Control
ROC Curve...
IBM SPSS Amos...

Automatic Linear Modeling...
Linear...
Curve Estimation...
Partial Least Squares...
Preacher and Hayes (2008) Multiple Mediation (INDIRECT)
Preacher and Hayes (2004) Simple Mediation Procedure (SOBEL)
Binary Logistic...
Multinomial Logistic...
PROCESS, by Andrew F. Hayes (<http://www.afhayes.com>)
Ordinal...
Probit...
Nonlinear...
Weight Estimation...
2-Stage Least Squares...
Optimal Scaling (CATREG)...

Visible: 29 of 29 Variables

	ID	Team	sd3	sd4	sd5	sd6	sd7	sd8	sd9	sd10	sd11	sd12	sd13	lp1	lp2
1	1	1.0	4	7	4	4	4	4	4	4	4	4	4	4	2
2	2	1.0	7	7	7	5	6	6	5	4	6	6	4	4	5
3	3	1.0	4	4	4	4	4	5	5	5	6	4	5	5	5
4	4	2.0	6	6	6	6	6	3	4	4	4	5	5	7	6
5	5	2.0	4	3	2	4	4	4	4	4	4	4	4	7	7
6	6	2.0	5	4	3	4	4	4	4	4	4	4	4	6	6
7	7	2.0	5	5	5	4	3	5	3	5	3	3	5	4	4
8	8	2.0	4	4	2	2	6	6	6	6	5	5	5	5	5
9	9	2.0	5	5	5	4	3	3	4	4	4	4	5	4	4
10	10	3.0	7	6	6	5	6	4	6	4	6	6	5	5	6
11	11	3.0	7	6	6	6	6	6	6	6	6	6	6	5	5
12	12	3.0	5	4	4	3	4	5	6	6	6	6	7	7	7
13	13	3.0	7	6	6	4	4	5	5	5	5	5	3	7	6
14	14	3.0	5	6	5	5	5	5	5	6	6	5	6	6	6
15	15	3.0	5	4	3	4	3	4	3	4	3	3	6	6	6
16	16	4.0	5	4	3	4	5	5	5	5	5	4	5	5	5
17	17	4.0	6	5	4	4	6	5	4	6	5	4	3	4	4
18	18	4.0	4	5	3	4	6	6	5	5	5	5	5	6	4
19	19	4.0	6	5	3	5	5	5	5	5	5	5	5	6	4
20	20	4.0	5	4	4	5	4	4	4	4	4	3	4	6	5
21	21	4.0	5	3	6	4	5	4	5	4	5	3	5	5	5
22	22	5.00	4	5	2	6	5	5	5	5	5	4	6	6	6
23	23	5.00	5	5	7	5	5	6	5	4	6	6	4	5	6
24	24	5.00	4	3	6	3	5	4	4	3	4	5	5	5	5
25	25	5.00	6	6	5	6	6	6	3	3	4	4	4	3	4
26	26	5.00	1	6	6	3	3	3	5	7	7	6	6	7	4
27	27	5.00	3	4	5	4	4	2	4	4	1	5	1	5	4
28	28	5.00	4	6	7	4	4	6	6	5	4	6	6	5	4
29	29	6.00	4	6	7	7	6	7	2	4	5	5	4	5	5
30	30	6.00	6	6	7	6	6	6	5	4	4	4	5	6	6
31	31	6.00	6	4	6	4	5	6	6	3	5	6	6	7	6
32	32	6.00	7	4	6	7	7	5	5	6	6	5	6	7	7
33	33	6.00	6	6	6	6	6	4	3	7	7	7	5	7	6
34	34	7.00	4	5	5	4	5	5	4	5	4	4	4	4	4
35	35	7.00	5	5	4	5	5	4	3	6	6	7	5	3	4
36	36	7.00	4	5	4	4	6	4	3	3	3	4	5	3	5
37	37	7.00	5	4	5	5	5	4	4	5	4	4	4	6	5

Data View Variable View

IBM SPSS Statistics Processor is ready

8:58 AM 11/4/2013

Simple Mediation

Data2.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help



																		Visible: 29 of 29 Variables		
	ID	Team	sd1	sd2	sd3	sd4	sd5	sd6	sd7	sd8	sd9	sd10	sd11	sd12	sd13	lp1	lp2			
1	1	1.00	4	4	7	4	4	4	4	4	4	4	4	4	4	4	2			
2	2	1.00	7	7	7	7	5	6	6	5	5	4	6	6	4	4	5			
3	3	1.00	3	4	4	4	4	4	4	5	5	5	6	4	5	5	5			
4	4	2.00								3	4	4	4	5	5	7	6			
5	5	2.00								2	4	4	4	4	4	7	7			
6	6	2.00								3	4	4	4	4	4	6	6			
7	7	2.00								5	4	3	5	3	3	5	4			
8	8	2.00								2	2	6	6	6	5	5	5			
9	9	2.00								5	4	3	3	4	4	5	4			
10	10	3.00								6	5	6	4	6	5	5	6			
11	11	3.00								6	6	6	6	6	6	5	5			
12	12	3.00								4	3	4	5	6	6	7	7			
13	13	3.00								6	4	4	5	5	3	7	6			
14	14	3.00								5	5	5	6	6	5	6	6			
15	15	3.00								3	4	3	4	3	3	6	6			
16	16	4.00								3	4	5	5	5	4	5	5			
17	17	4.00								4	4	6	5	4	3	4	4			
18	18	4.00								3	4	6	6	5	5	5	6			
19	19	4.00								3	5	5	5	5	5	6	4			
20	20	4.00								4	5	4	4	3	4	6	5			
21	21	4.00								6	4	5	4	5	3	5	5			
22	22	5.00								2	6	5	5	5	4	6	6			
23	23	5.00								5	4	5	6	6	4	5	6			
24	24	5.00								4	3	4	3	5	5	5	5			
25	25	5.00								3	3	4	4	4	4	3	4			
26	26	5.00								3	5	7	7	6	6	7	4			
27	27	5.00								4	4	1	5	5	1	5	4			
28	28	5.00								6	6	5	4	6	6	5	4			
29	29	6.00								4	5	5	4	5	4	5	5			
30	30	6.00								5	5	4	4	4	5	6	6			
31	31	6.00								6	3	5	6	6	6	7	6			
32	32	6.00	7	4	6	6	7	7	5	5	6	6	5	5	6	7	7			
33	33	6.00	6	6	6	6	6	6	4	3	7	7	7	7	5	7	6			
34	34	7.00	4	5	5	4	4	5	5	4	5	4	4	4	4	4	4			
35	35	7.00	5	5	4	5	5	5	4	3	6	6	6	7	5	3	4			
36	36	7.00	4	5	5	4	4	6	4	3	3	3	4	5	3	4	5			
37	37	7.00	5	4	5	5	5	5	4	4	5	5	4	4	4	6	5			

Preacher and Hayes (2004) Simple Mediation Procedure (SOBEL)

Variables:

- sd2
- sd3
- sd4
- sd5
- sd6
- sd7
- sd8
- sd9
- sd10
- sd11
- sd12
- sd13
- lp1
- lp2
- lp3
- lp4
- lp5
- lp6
- lp7
- lp8
- lp9
- lp10
- lp11
- lp12
- age2

Dependent Variable (Y)

Proposed Mediator (M)

Independent Variable (X)

Sobel test standard error

Second order

Bootstrap samples

0

Show effect sizes

OK Paste Reset Cancel Help

Data View Variable View

IBM SPSS Statistics Processor is ready

8:59 AM
11/4/2013

Multiple Mediation

IBM SPSS Statistics Data Editor window showing a dataset named "Data2.sav" with 29 variables. The variables include ID, Team, and 13 standardized variables (sd1 to sd13), along with two latent variables (lp1 and lp2). The data is displayed in a table with 37 rows.

The "Preacher and Hayes (2008) Multiple Mediation Procedure" dialog box is open, showing the following settings:

- Dependent Variable (Y):
- Proposed Mediator(s) (M):
- Independent Variable (X):
- Covariate(s):
- Contrast Indirect Effects: ☐
- Bootstrapping: ☒ Number of samples: 1000
- Percentile: ☐
- Bias Corrected (BC): ☒
- BC and Accelerated: ☐
- Confidence Intervals: 95
- Normal theory tests: ☐

The dialog box also includes buttons for OK, Paste, Reset, Cancel, and Help.

The bottom status bar indicates "IBM SPSS Statistics Processor is ready" and the date/time "9:00 AM 11/4/2013".

Multiple Mediation

SPSS, SAS, and Mplus Macros and Code - Andrew F. Hayes, Ph.D. - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Inbox - Outlook Web App, light version x Inbox (130) - ramayah@gmail.com - ... x

SPSS, SAS, and Mplus Macros and Co... x +

afhayes.com/spss-sas-and-mplus-macros-and-code.html

☆ ▾ ↻

Preacher & Hayes

🔍 🏠

Most Visited Getting Started Suggested Sites Web Slice Gallery

MEDIATE

MEDIATE for SPSS is an alternative to PROCESS for implementing the kind of analysis described in

Hayes, A. F., & Preacher, K. J. (2012). Statistical mediation analysis with a multicategorical independent variable. *Manuscript submitted for publication.* [PDF coming soon]

MEDIATE conducts mediation analysis (single and multiple mediators) with either continuous, dichotomous, or multicategorical independent variables. It is similar in functionality to [INDIRECT](#) but offers additional features that accommodate multiple independent variables simultaneously and that simplify the coding of multicategorical independent variables. When analyzing the effect of a multicategorical independent variable, the user can produce the requisite $k - 1$ variables coding group (where k is the number of groups) manually and enter them as independent variables or have MEDIATE automatically generate the variables using either indicator, effect, sequential coding, or Helmert coding. It offers tests of relative direct and indirect effects, including percentile bootstrap and Monte Carlo confidence intervals for indirect effects. It also automatically conducts a test of homogeneity of regression (i.e., interaction between X and M in the model of Y).

Please read the [download instructions at the top of this page.](#)

SPSS version

Documentation: [mediate.pdf](#)

Macro: [mediate.sps](#)

There is no SAS version of MEDIATE.

MODMED

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Assessing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, 42, 185-227. [PDF]



Process Models

Data2.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help



																			Visible: 29 of 29 Variables	
	ID	Team	sd1	sd2	sd3	sd4	sd5	sd6	sd7	sd8	sd9	sd10	sd11	sd12	sd13	lp1	lp2			
1	1	1.00	4	4	7	4	4	4	4	4	4	4	4	4	4	4	2			
2	2	1.00	7	7	7	7	5	6	6	5	5	4	6	6	4	4	5			
3	3	1.00										5	6	4	5	5	5			
4	4	2.00										4	4	5	5	7	6			
5	5	2.00										4	4	4	4	7	7			
6	6	2.00										4	4	4	4	6	6			
7	7	2.00										3	5	3	3	5	4			
8	8	2.00										6	6	6	5	5	5			
9	9	2.00										3	3	4	4	5	4			
10	10	3.00										6	4	6	5	5	6			
11	11	3.00										6	6	6	6	5	5			
12	12	3.00										4	5	6	6	7	7			
13	13	3.00										4	5	5	3	7	6			
14	14	3.00										5	6	6	5	6	6			
15	15	3.00										3	4	3	3	6	6			
16	16	4.00										5	5	5	4	5	5			
17	17	4.00										6	5	4	3	4	4			
18	18	4.00										6	6	5	5	5	6			
19	19	4.00										5	5	5	5	6	4			
20	20	4.00										4	4	3	4	6	5			
21	21	4.00										5	4	5	3	5	5			
22	22	5.00										5	5	5	4	6	6			
23	23	5.00										5	6	6	4	5	6			
24	24	5.00										4	3	5	5	5	5			
25	25	5.00										4	4	4	4	3	4			
26	26	5.00										7	7	6	6	7	4			
27	27	5.00										1	5	5	1	5	4			
28	28	5.00										5	4	6	6	5	4			
29	29	6.00										5	4	5	4	5	5			
30	30	6.00										4	4	4	5	6	6			
31	31	6.00										5	6	6	6	7	6			
32	32	6.00										6	5	5	6	7	7			
33	33	6.00										7	7	7	5	7	6			
34	34	7.00										4	4	4	4	4	4			
35	35	7.00	5	5	4	5	5	5	4	3	6	6	6	7	5	3	4			
36	36	7.00	4	5	5	4	4	6	4	3	3	3	4	5	3	4	5			
37	37	7.00	5	4	5	5	5	5	4	4	5	5	4	4	4	6	5			

PROCESS Procedure for SPSS, written by Andrew F. Hayes (www.afhayes.com)

Data File Variables

- ID
- Team
- Section 1 [sd1]
- sd2
- sd3
- sd4
- sd5
- sd6
- sd7
- sd8

Model Number

4

Bootstrapping for indirect effects

Bootstrap Samples

1000

Bootstrap CI method

☐ Percentile

☒ Bias Corrected

Confidence level for confidence intervals

95%

Covariate(s) in model(s) of...

☒ ...both M and Y

☐ ...M only

☐ ...Y only

Outcome Variable (Y)

Independent Variable (X)

M Variable(s)

Covariate(s)

Proposed Moderator W

Proposed Moderator Z

Proposed Moderator V

Proposed Moderator Q

OK Paste Reset Cancel

About Options Conditioning

Data View Variable View

IBM SPSS Statistics Processor is ready

9:00 AM
11/4/2013

Workshop

Mediation and Moderation Seminar | Statistical Workshops in Philadelphia | Statistics Training Course and Class | Statistical Horizons - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Inbox - Outlook Web App, light version x Inbox (131) - ramayah@gmail.com - ... x Mediation and Moderation Analysis ... x Mediation and Moderation Seminar | ... x +

www.statisticalhorizons.com/seminars/public-seminars/medandmod13



Preacher & Hayes



Most Visited Getting Started Suggested Sites Web Slice Gallery



Home

About

Resources

FAQs

Seminars

Instructors

Blog

Contact Us

Public Seminars

On-Site Seminars

Mediation and Moderation

A 5-Day Seminar Taught by **Andrew Hayes, Ph.D.** and **Kristopher Preacher, Ph.D.**

[Read 13 reviews of this course](#)

This seminar focuses on two topics in causal analysis that are closely related and often confused. Suppose we have three variables, X , M and Y . We say that M is a **mediator** of the effect of X on Y if X carries its influence on Y at least partly by influencing M , which then influences Y . This is also known as an **indirect effect** of X on Y through M . On the other hand, we say that M **moderates** the effect of X on Y if that effect varies in size, sign, or strength as a function of M . This is also known as **interaction**.

Although these concepts are fairly simple, the statistical issues that arise in estimating and testing mediation and moderation effects turn out to be rather complex and subtle. **Andrew Hayes** and **Kristopher Preacher** have been among the leading contributors to the literature on these methods. They have developed powerful new methods for estimating mediation and moderation effects and special software tools that can be used with SAS or SPSS.

In this seminar, you will learn about the underlying principles and the practical applications of these methods. The seminar is divided roughly into three parts:

1. Partitioning effects into direct and indirect components, and how to quantify and test hypotheses about indirect effects.

REGISTER NOW

SEMINAR INFORMATION

Monday July 15, 2013 9:00 AM -
Friday July 19, 2013 5:00 PM (Eastern Time)

The Hub Commerce Square
2001 Market Street – Kyoto Room
Philadelphia, Pennsylvania 19103
United States

[View Map](#)

CONTACT INFORMATION

Phone: 610-642-1941
Fax: 419-818-1220
Email: info@statisticalhorizons.com

PAYMENT INSTRUCTIONS

PayPal and all major credit cards are accepted. The fee of \$1695 includes all course materials. If registration is completed by June 17, the fee is reduced to \$1495.

Thank you for listening



Persevere **1**

Success **2**