Slide 1

Email: jkanglim@unimelb.edu.au
Office: Room 1110 Redmond Barry Building
Website: http://jeromyanglim.googlepages.com/

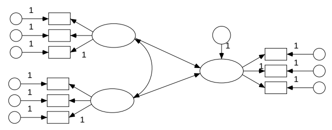
Appointments: For appointments regarding course or with the application of statistics to your thesis, just send me an email

Structural Equation Modelling

325-711 Research Methods

2007

Lecturer: Jeromy Anglim



Description: This seminar on structural equation modelling (SEM) is designed as an introduction to the technique. SEM is used to test 'complex' relationships between observed (measured) and unobserved (latent) variables and also relationships between two or more latent variables. The module assumes participants are familiar with introductory statistics but have little or no knowledge of SEM. It will consider the appropriate use of SEM, fundamentals underlying SEM and the use of SEM notation (inc. path diagrammes). Further, the module will introduce three models fundamental to SEM: confirmatory factor analysis, causal models and measurement models. It will introduce model specification and assessment of model fit. Amos will be used to demonstrate relevant concepts.

Slide 2

Overview

- Introduction
  - Overview of SEM
  - Prerequisite knowledge
- Path diagrams & Amos
- Worked Example
  - Data preparation
  - Measurement Models and SEM
  - Estimation, Fit, & Parameter estimates
- Advanced Topics

Slide 3

Readings & Resources

– Hair, J., Black, B. Babin, B., Anderson, R. and Tatham, R. (2006). Multivariate Data Analysis (6. th. edition). Upper Saddle River, NJ: Prentice-Hall.
  - Chapter 10 Structural Equation Modeling: An Introduction.
  - Chapter 11: SEM Confirmatory Factor Analysis
- Boomsma, A. (2000). Reporting analyses of covariance structures. Structural Equation Modeling, 7(3), 461-483.
- McDonald, R. P., & Ringo Ho, M. (2002). Principles and practice in reporting structural equation analyses. Psychological Methods, 7(1), 64-82.

- Amos Tutorials
  - http://www.utexas.edu/its-archive/rc/tutorials/stat/amos/
  - http://www.amosdevelopment.com/download/
  - http://amosdevelopment.com/video/index.htm
- Resources
  - Kline – Principles and Practice of Structural Equation Modelling
  - Garson - http://www2.chass.ncsu.edu/garson/pa765/structur.htm

Boomsma (2000): This article is very useful for providing a checklist of all the things you should report when writing up a structural equation model for a journal article or thesis. The emphasis is on providing all the information required to understand and evaluate the quality of the proposed model.
Hair et al: Chapter 10 to 12 of the 6th edition provide a good overview of structural equation modelling, for those not wanting to deal too much with formulas.
Kline: see website for many helpful overheads
http://www.guilford.com/cgi-bin/cartscript.cgi?page=etc/sem_data/kline_data.html&dir=research&cart_id=305142.19124
http://www.amosdevelopment.com/download/ - This has the student version of Amos and a comprehensive user's guide. If you are planning on using Amos as your tool, I would advice working through the examples in the user's guide.

Slide 4

Why do SEM?

- General
  - Estimates relationships between latent variables
  - Allows for explicit tests of competing models
  - Explore direct, indirect and total effects
  - Explore multivariate relationships in an integrated manner
- Specific Examples
  - **Scale development:** confirmatory factor analysis of a scale
  - **Mediation:** Test plausibility of simple and complex mediational models
  - **Between group model stability**

Slide 5

SEM in a nutshell

- A graphical representation of the pattern of correlations between a set of variables
- A good model is:
  - Parsimonious
  - Theoretically justifiable
  - Reproduces the underlying correlation matrix based on the constraints imposed
- It's the general linear model with latent variables

Slide 6

A word of warning

- Prerequisites
  - Good understanding of correlation, regression, factor analysis, reliability analysis, variance, z-scores is required
- Word of warning
  - SEM is a powerful tool, but If you don't understand what it's doing, learn what it does mean or don't use it
- Making it work
  - If you are serious about it, you probably want to do a complete course on SEM
  - OR systematically work through some good textbooks doing the examples

Some students adopt SEM, because it "looks good", or their supervisor told them it's what is expected of a PhD, but do not go through the necessary training or revision of the prerequisites above and SEM in general to make it work. The result is a few months of fiddling around with SEM with no additional interpretive value.

Slide 7

Refresher of the prerequisites

- Correlation & covariance
  - Form, Direction, Degree
- Exploratory Factor Analysis
- Reliability
- Regression
  - Regression coefficients (standardised and unstandardised)
  - r-squared
  - Confidence intervals and p values
  - Threats to valid inferences (assumptions)
- Causal Inference

Correlation & Covariance: Correlation describes the linear relationship between two variables (i.e., the "form"). Correlations range from -1 to +1. The direction of the relation is indicated by the sign and the degree of the relationship is indicated by the absolute size of the correlation.
Covariance is an unstandardised form of correlation. If covariance is a positive number, there is a positive relationship between the variables, and if it is a negative number, there is a negative relationship between the variables.
Correlation matrices and covariance matrices describe the pairwise relationships between a set of variables.
Exploratory Factor Analysis (EFA): EFA is a tool for assessing the factors that underlie a set of variables. It is frequently used to assess which items should be grouped together to form a scale. It introduces the distinction between the observed variables (e.g., questions on a test) and latent variables (the underlying factor that we are usually interested in). This is extended within the SEM context when performing confirmatory factor analysis.
Multiple Linear Regression: This involves predicting an outcome variable from one or more predictor variables. The predictor variables are weighted in order to form a composite variable that aims to maximise prediction of the outcome variable. Regression coefficients are used to indicate the expected increase in the outcome variable for an increase of one on the predictor variable holding all other predictor variables constant. Standardised regression coefficients are frequently used to indicate the relative importance of predictors. Structural equation modelling has all these elements. Sometimes we are interested in assessing the relative importance of different latent predictor variables in explaining variation in a latent outcome variable.
Causal inference: Causal inference is strongest in controlled experiments involving random allocation of subjects to conditions. Longitudinal data and other quasi-experimental designs can also provide evidence for causal claims, although the threats to such an influence are typically stronger than in experiments. Finally, correlational cross-sectional designs which represent the majority of applications of SEM designs make provide the weakest form of evidence of causality. It is for this reason that we need to be cautious in the interpretation of directional arrows in structural equation modelling.

Slide 8

Variable types

- Exogenous versus endogenous
- Independent versus dependent
- Mediators
- Latent versus observed

Exogenous variables have causes that are assumed to be external to the model. Exogenous variables can only have double headed arrows (i.e., correlation) going into them.
Endogenous variables are predicted by other variables in the model. Endogenous variables will have a directed arrow entering into them (i.e., prediction) both from the substantive predictors and a residual term that represents the variance not explained by the predictors.
Latent variables are not measured directly in a study. They are assumed to bring about the observed responses.
Observed variables are directly measured in a study.

Slide 9

Taxonomy of Models

- Path analysis
- Confirmatory Factor Analysis
  - Measurement model
- Structural Equation Model

- Exploratory to confirmatory continuum
  - Use of exploratory factor analysis
  - Competing models approach
  - Post hoc modification
  - Replication

Path Analysis: A model with only observed variables
Confirmatory Factor Analysis: A model with no directed arrow going into a latent variable
Structural Equation Model: A model with at least one directed arrow going into a latent variable

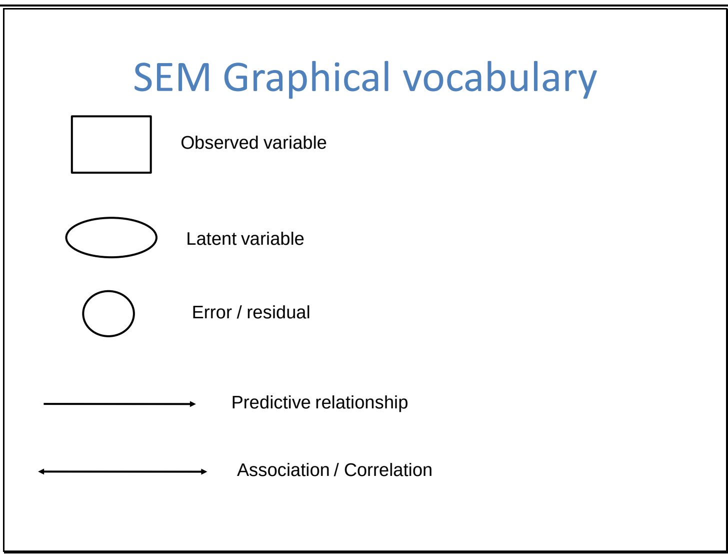
Slide 10

Software

- LISREL
- SPSS's Amos
- EQS
- Mplus
- Mx
- sem package in R
- SAS PROC CALIS

Software differs in terms of the way models are specified. Some allow for direct drawing of path diagrams, which is often easier to learn if you are less mathematically inclined. New versions are released on a regular basis.
LISREL: The original SEM software.
AMOS: Today's talk focuses on Amos.
R: A free version of SEM software for doing basic analyses. http://personality-project.org/revelle/syllabi/454/454.syllabus.pdf
http://socserv.mcmaster.ca/jfox/Courses/Oxford-2006/index.html

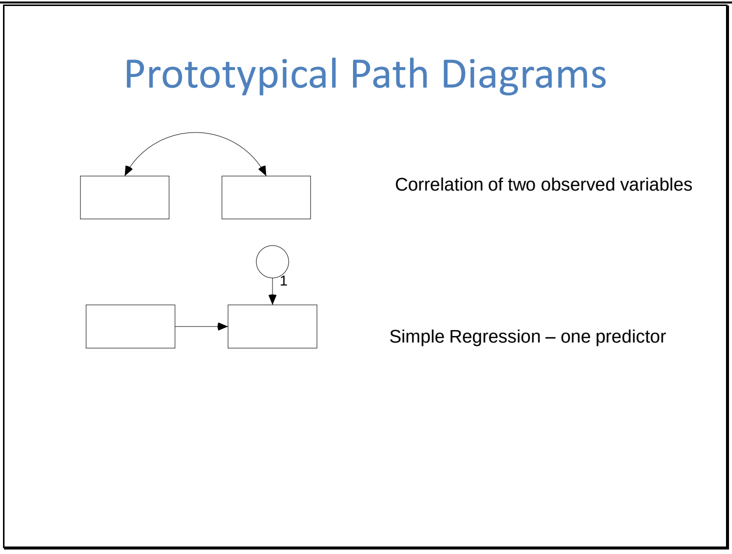
Slide 11



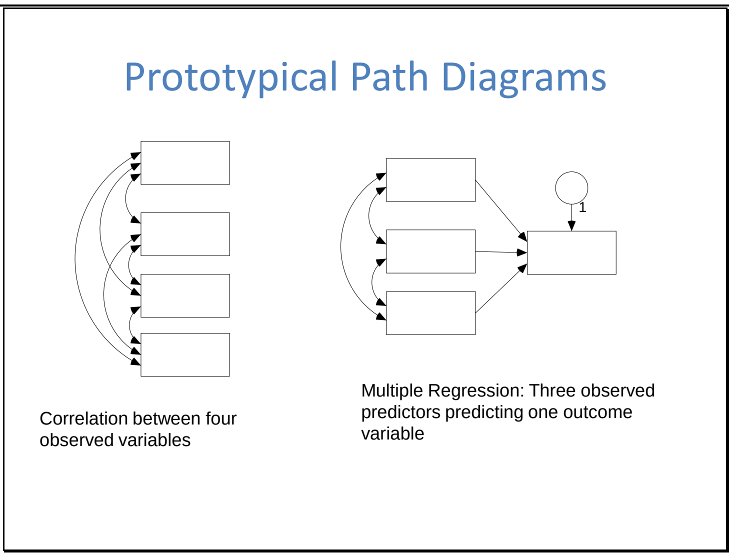
Structural equation models can be specified in several formats. Some involve the specification of multiple matrices. Amos introduced a way of specifying models in terms of path diagrams. These path diagrams follow a set of standard conventions. It is an important skill to be able to convert theoretical hypotheses and the data into a path diagram.
Rectangle: Observed variables, such as items from a questionnaire.

Ellipse: Latent variables that are estimated from observed variables
Error: Error in predicting a variable. Every variable that has a directed arrow coming in will have an error associated with it.
Single-headed arrow: Relationships that are predictive.
Double headed arrow: Correlation

Slide 12

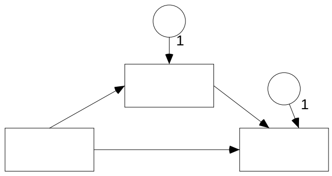


Slide 13

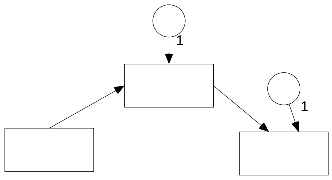


Slide 14

Prototypical Path Diagrams



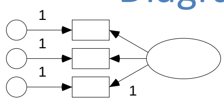
Partial Mediation with observed variables



Full mediation with observed variables

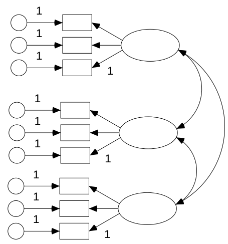
Slide 15

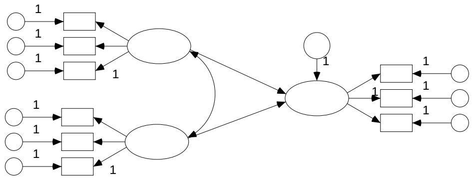
Prototypical Path Diagrams



Latent factor with three indicator variables

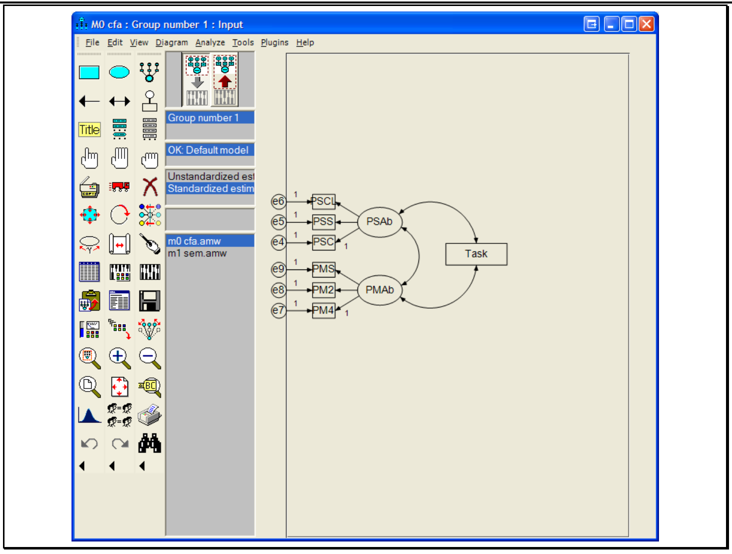
Confirmatory factor analysis with three correlated latent factors and three indicators per latent factor





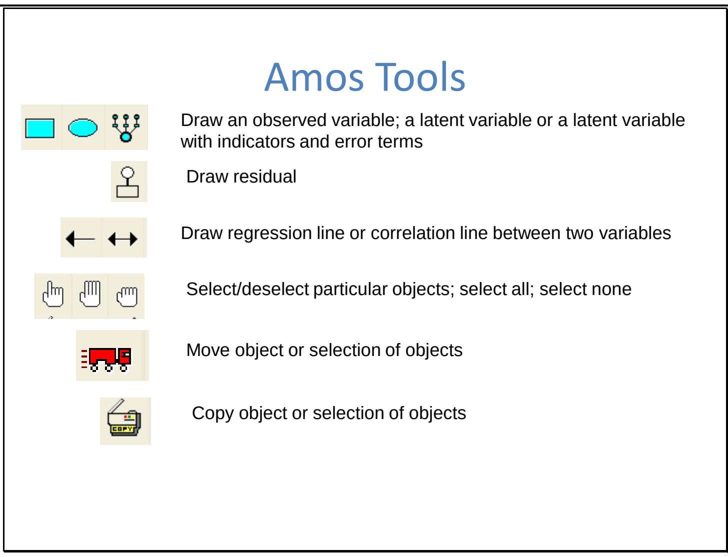
Structural equation model: One latent variable predicted by two latent predictor variables; each latent variable has three indicators each

Slide 16



The Amos user interface consists of a range of drawing tools in a toolbar, a set of menus, and a model drawing area.

Slide 17

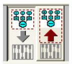


Slide 18

Amos Tools

Set analysis properties including method of estimation, output options, and bootstrapping options

Run the model



Left downward arrow: model specification mode
Right upward arrow: output mode

View tabular output

- Several other useful tools
  - Important to get to know the different drawing tools
  - Increase efficiency; Increase attractiveness and interpretability
- Macros / Plugins
  - Draw Covariances
  - SRMR
  - Name unobserved variables

Amos has a user interface that facilitates exploration. It is well worth exploring what each of the buttons on the tool bar does.
Additional tips: Holding your mouse over a tool will highlights its name. Preserve symmetries tool is particularly useful for making it easier to move or copy diagrams and factors.
Plugins
Draw covariances: This speeds up the process of drawing the covariances between a set of variables. This is typically required in confirmatory factor analysis. To use the tool, highlight the desired variables and run the plug in.
SRMR: This plugin is a way of getting the SRMR statistic for the model.
Name unobserved variables: This tool speeds up the process of assigning names to unobserved variables using a coding system of e1, e2, etc. This plugin is usually run after any specific latent variable names have been given.

Slide 19

Model Testing

- Define theorised model
- Structure dataset
  - Missing values
  - Normality & outliers assessment
- Assess measurement model/s; then structural model/s
  - Estimation
  - Fit
  - Interpretation
  - Optionally modify and refine models
- Relate back to theory

Slide 20

Simple Working Example

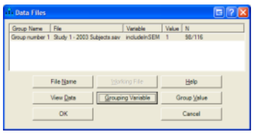
- 3 tests assumed to measure psychomotor ability
  - Lower scores indicate more ability
  - Simple Reaction Time; 2 Choice Reaction Time; 4 Choice Reaction Time
- 3 tests assumed to measures perceptual speed ability (higher scores indicate better performance)
  - Higher scores indicate more ability
  - Clerical Speed Test; Number Sort Test; Number Comparison Test
- A measure of task performance
  - Average number of seconds to complete text editing task
  - lower scores indicate better performance
- Sample Size = 98
  - i.e., a little small, but large correlations and a small number of parameters being estimated

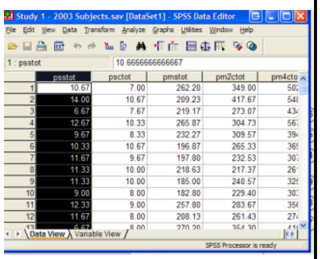
See Hair et al for a further discussion of sample size considerations.

Slide 21

Data File Input

- Raw Data





Correlation or Covariance Matrix

```
MATRIX DATA
variables= rowtype_ tep_overall pscltot  psstot psctot pmstot pm2ctot pm4ctot
/format=list lower diag /contents= N corr sd.
BEGIN DATA.
N 98 98 98 98 98 98 98
CORR     1        -0.446   -0.274   -0.426   0.514    0.578    0.626
CORR     -0.446   1        0.435    0.518    -0.348   -0.405   -0.406
CORR     -0.274   0.435    1        0.61     -0.26    -0.285   -0.235
CORR     -0.426   0.518    0.61     1        -0.279   -0.321   -0.316
CORR     0.514    -0.348   -0.26    -0.279   1        0.708    0.687
CORR     0.578    -0.405   -0.285   -0.321   0.708    1        0.847
CORR     0.626    -0.406   -0.235   -0.316   0.687    0.847    1
sd 12.99 4.17     2.62     1.99     38.59    51.34    92.24
END DATA.
```

To link the data file into Amos: File >> Data files; then click “File Name”

Raw Data: Amos can handle a wide range of file formats including SPSS, Excel, plain text, and other formats.

Correlation or covariance matrix: It is also possible to analyse a correlation matrix without having the raw data. The above syntax provides a template that you can use to enter data in this format. This can be useful if you wish to analyse a correlation matrix published in a journal article. It is also a method for analysing different measures of association such as the polychoric correlation. Some options in Amos require there to be raw data.

Slide 22

Exploratory Factor Analysis

Pattern Matrix[a]

	Factor 1	Factor 2
pm2ctot PMA: 2 Choice RT Average (ms)	.933	.004
pm4ctot PMA: 4 Choice RT (ms)	.918	.019
pmstot PMA: Simple RT Average (ms)	.744	-.031
psctot PSA: Number Comparison (Average Problems Solved (Correct - Incorrect) per minute)	.044	.866
psstot PSA: Number Sort (Average Problems Solved (Correct - Incorrect) per minute)	.043	.739
pscltot PSA: Clerical Speed Total (Average Problems Solved [Correct - Incorrect] per minute)	-.205	.524

Extraction Method: Maximum Likelihood.
Rotation Method: Promax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

CFA

X	0
X	0
X	0
0	X
0	X
0	X

Factor Correlation Matrix

Factor	1	2
1	1.000	-.457
2	-.457	1.000

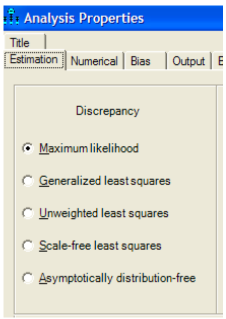
Extraction Method: Maximum Likelihood.
Rotation Method: Promax with Kaiser Normalization.

Exploratory Factor Analysis (EFA): A 2 factor exploratory factor analysis was performed using maximum likelihood extraction and an oblique rotation (promax; kappa = 4). It suggests that the loadings shows something like simple structure in accordance with the factor structure proposed. The correlation between the two factors is estimated to be -.46.
Confirmatory factor analysis (CFA): EFA allows all the loadings to freely vary. In contrast, CFA constrains certain loadings to be zero, typically only allows the item's loadings on its main factor to freely vary. To the extent that the EFA shows large cross loadings (e.g., loadings on non-main factors above .3), measures of fit are likely to be poorer on the CFA.

Slide 23

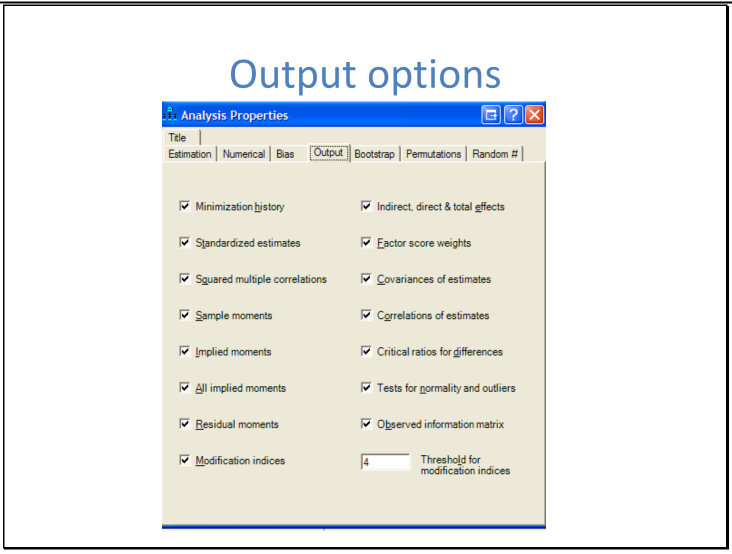
Estimation

- Maximum Likelihood
- Generalised least squares
- Others



Maximum Likelihood: This is the most frequently used
Iterative procedure that seeks to minimise a discrepancy between model and sample covariance matrices.

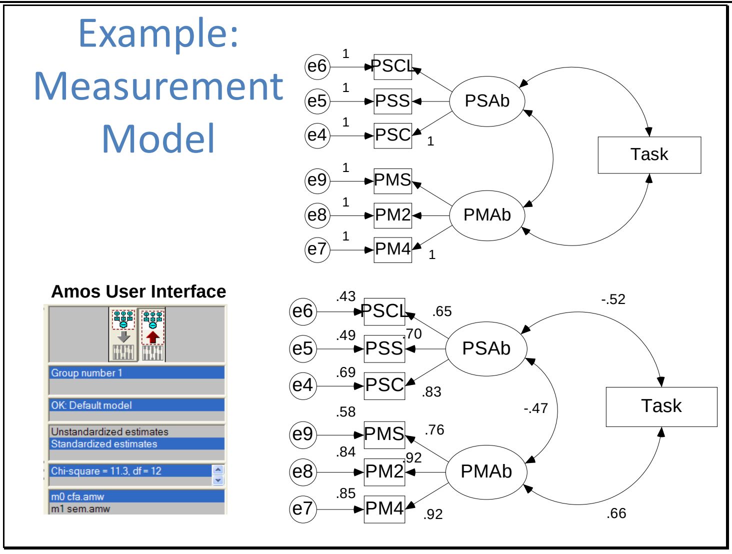
Slide 24



It's often easiest just to ask for everything. You can then navigate to the output that you want to see.
Standardized estimates: This shows the standardised regression coefficients and the correlations on the plot and in the Amos output.
Squared multiple correlations: This shows the r-squared for each variable that has a directional arrow coming into it.
Sample moments: This shows the correlations, covariance, variance and optionally also the means for the observed variables.
Implied moments: This shows the correlations, covariance, variance and optionally also the means that have been estimated based on the proposed model and the application of the estimation method. If the proposed model is good, the sample and implied moments should be very similar.
All implied moments: This also shows correlations, covariance, variance and optionally also the means, for latent variables as well as observed variables. One of the main benefits of SEM is that you can examine relationships between latent variables. As such this output can be particularly interesting.
Residual moments: This shows the difference between the sample and the model covariance matrix. This can be used to assess the effectiveness of the model in explaining the sample covariance. The standardised residual covariances are particularly useful in assessing areas where the model may be failing.
Modification Indices: These are only available if you have raw data with no missing data in the working data file (i.e., there was no missing data or missing data has been replaced). Modification indices indicate the improvement in fit that will result in the inclusion of a particular relationship in the model. Instead of showing all possible modifications, setting a threshold for modification indices reduces the display of modification indices to a smaller set.
Indirect, direct and total effects: This output is relevant when running a structural equation model with some variables that are mediators or partial mediators. In this case some of the effect of certain variables will be direct and some of the effect will be mediated by another variable. The accompanying output quantifies these effects.
Tests for normality and outliers: This provides statistical tests for normality, which can be used to guide decisions about transformations. In particular, there is a test of multivariate

normality which is not readily available in SPSS. Measures of Mahalanobis distance are also provided for each case indicating cases that are in some way unusual either in a univariate or multivariate sense. It is often worth following up on these cases to assess whether they should be removed from analyses.

Slide 25



Task performance is an observed variable and the two abilities (PSAb – Perceptual Speed Ability; and PMAb – Psychomotor Abiltiy) are latent variables.
The output displayed above represents standardised output (i.e., the numbers represent correlations and standardised regression coefficients).
Correlations between variables are shown with numbers next to double headed arrows. The two latent abilities are correlated .47. Both abilities correlate fairly highly with task performance.
The directed arrows from the abilities to the observed variables indicate the loadings of the variable on the proposed latent factor. Thus, a one standard deviation increase on perceptual speed ability (PSAb) is associated with a .65 standard deviation increase on the Clerical Speed Test (PSCL). Remember that in simple regression, a standardised regression coefficient is the same as the correlation. Thus, we could also say that PSAb correlates .65 with PSCL.
Each of the six indicator variables for the two latent ability variables has an associated error variable. This reflects the assumption that each observed test variable is partially predicted by the latent factor it is trying to measure and the rest is error. The values to the left of the six indicator variables reflect the variance explained by the latent ability factor (i.e., the r-squared). In the context of factor analysis, this is referred to as the communality. We can see that for PSCL with only one predictor, this is just the squared loading (.65 * .65 = .42; .01 differenc is due to rounding)
Note: That one of the loadings for each factor is constrained to one. This is to ensure that the model is identifiable by setting a scale for the latent factor. Amos sets this value by default, but there may be times where you wish to set it yourself or that you wish to change the item which is assigned the value of 1.
Amos User Interface: From top to bottom of the Amos display:

- Once the model has been successfully run, the red upward arrow will appear. Clicking on this shows the output on the path diagram.
- This is only relevant when there are multiple groups, such as when doing a multiple groups confirmatory factor analysis.
- Amos allows the specification of multiple nested models which can then be directly compared in terms of fit. In the current example there is only one model
- There is a choice between standardised and unstandardised estimates being displayed on the path diagram. Generally, standardised estimates are easier to interpret.
- a basic summary of the chi-square statistic is provided. If this is expanded it will give further information about model estimation.
- Multiple models can be saved in the one directly and can be quickly brought up.

Slide 26

Normality

Variable	min	max	skew	c.r.	kurtosis	c.r.
tep_overall	14.8	78.6	.7	3.0	.2	.3
pmstot	169.9	415.2	1.8	7.3	5.5	11.0
pm2ctot	193.1	417.7	.7	2.7	-.3	-.6
pm4ctot	259.7	696.6	1.0	4.2	1.0	2.1
pscltot	9.3	32.2	-.1	-.3	.1	.3
psstot	4.3	16.7	.0	.2	-.6	-1.1
psctot	5.0	15.0	.6	2.4	.2	.4
Multivariate					13.5	5.9

The critical ratio represents skewness (or kurtosis) divided by the standard error of skewness (or kurtosis). It is interpreted as one would interpret a z-score. Values greater than 2, 2.5 or 3 are often used to indicate statistically significant skew or kurtosis. In the present situation it would appear that the psychomotor tests in particular are showing some positive skew. Analyses could be repeated after transforming these variables to remove the skewness.

Slide 27

Sample Moments

Sample Covariance Matrix

	Task	PMS	PM2	PM4	PSCL	PSS	PSC
Task	167.2						
PMS	255.4	1474.8					
PM2	381.9	1388.6	2609.6				
PM4	742.5	2421.7	3970.1	8422.9			
PSCL	-23.9	-55.5	-85.8	-154.8	17.2		
PSS	-9.3	-26.1	-38.1	-56.4	4.7	6.8	
PSC	-10.9	-21.3	-32.5	-57.6	4.3	3.2	3.9

Sample Correlation Matrix

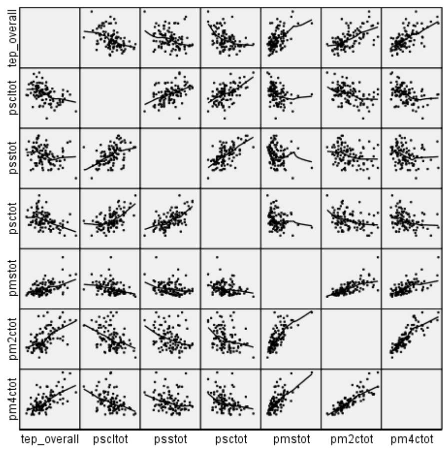
	Task	PMS	PM2	PM4	PSCL	PSS	PSC
Task	1.00						
PMS	.51	1.00					
PM2	.58	.71	1.00				
PM4	.63	.69	.85	1.00			
PSCL	-.45	-.35	-.40	-.41	1.00		
PSS	-.27	-.26	-.29	-.24	.43	1.00	
PSC	-.43	-.28	-.32	-.32	.52	.61	1.00

The tables above show the sample correlation and covariance matrices. This is essentially the raw data that the structural equation model is aiming to represent. The number of cells shown in the matrix reflects the number of distinct sample moments. The diagonal on the covariance matrix represents the variance of each variable.

In the present example, all the variables appear to be at least moderately correlated. The psychomotor tests appear particularly highly correlated with each other. The perceptual speed tests are also fairly highly intercorrelated, but not as much as the psychomotor tests. All the tests are correlated with task completion time, although the psychomotor tests appear to show a slightly stronger relationship.

Slide 28

Scatter Matrix



SEM can be thought of as a tool for modelling the correlations between variables. Thus, it is important that the correlation coefficient is an appropriate summary statistic of the relationship between the variables. It is important to spend some time getting to know the

data and assessing whether a linear relationship adequately describes the data and assessing whether there are any outliers.
The above graph was obtained in SPSS, by going Graphs >> Scatter >> Matrix Scatter. The resulting graph was double clicked and the points were selected and the point size was reduced to 1 and the Elements >> Fit line at total >> loess option was selected.
Above we see that in general a linear relationship seems to represent the relationship between the variables fairly well. However, there are a few bivariate outliers.

Slide 29

Parameters

- Number of Sample moments $= \frac{k^2}{2} + \frac{k}{2}$
  - Where k is the number of variables
  - Equivalent to the number of variances and unique covariances
- Identification

Identification is a relatively technical matter which determines whether a model can be estimated. A general strategy is to have at least 3 or four indicators per latent variable. Hair et al (p. 783) provides further discussion of this important consideration.

Slide 30

Notes for Model

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments:	28
Number of distinct parameters to be estimated:	16
Degrees of freedom (28 - 16):	12

Result (Default model)

Minimum was achieved
Chi-square = 11.27
Degrees of freedom = 12
Probability level = .51

The model had 7 variables. This leads to 7*7/2 + 7/2 = 28 sample moments. These reflect 21 measures of covariance and 7 measures of variances.

Estimated parameters were:

9 parameters: the variance for the 7 observed variables and the 2 latent variables

3 parameters: the covariance for the 3 latent variables

4 parameters: four of the six loadings (two of the loadings are constrained to equal one)

Slide 31

Model Parameter Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
psctot	<---	PSAb	1				
psstot	<---	PSAb	1.106	0.175	6.312	***	par_1
pscltot	<---	PSAb	1.645	0.311	5.289	***	par_2
pm4ctot	<---	PMAb	1				
pm2ctot	<---	PMAb	0.554	0.041	13.536	***	par_3
pmstot	<---	PMAb	0.345	0.036	9.529	***	par_4

In addition to the estimates displayed on the path diagrams, Amos also provides the same information in tabular form. The tabular output also displays standard errors and tests of statistical significance. The Critical ratio is the estimate divided by the standard error. Values greater than 2 tend to indicate an estimate that is statistically significantly different from zero at the .05 level.

Output can be obtained for the unstandardised and standardised regression coefficients, covariances and correlations and the squared multiple correlations.

Slide 32

Assessing fit

- Chi square: (called CMIN in Amos)
- RMSEA : Root Mean Square Error of Approximation
- SRMR: Standardised Root Mean Residual
- CFI: Comparative Fit Index
- Many more:
  - Chi Square / df
  - Tucker Lewis Index
  - AIC, BIC, etc.
  - And more

Chi-square: Sample size minus 1 multiplied by discrepancy between sample covariance matrix and the model covariance matrix. Smaller chi-square values indicate better fit. Ideally, the chi-square would be non-significant indicating no significant discrepancy between model and data. However, in most applications this will not hold. Beyond having a poor model, chi-square will increase with larger samples and non-normally distributed data. Both of these issues has nothing to do with the appropriateness of the proposed model. Also, slight discrepancies between model and the data may result in a statistically significant chi square.
RMSEA: rule of thumb suggests <.05 is good; <.08 is acceptable; Software also provides confidence intervals
SRMR: rule of thumb suggests < .08 is good
CFI: rule of thumb suggests >.90 is good; >.95 is very good
See Appendix C of the Amos 7 User's guide for a relatively user friendly overview of the different fit measures reported by Amos:
http://amosdevelopment.com/download/Amos%207.0%20User's%20Guide.pdf
See chapter 10 of Hair et al (p.745 onwards 6^{th} edition) for a relatively non-technical overview.

Slide 33

Assessing Fit

- Measures
  - Some measures of fit reward more parsimonious models (i.e., greater df)
  - Chi-square punishes models with larger sample sizes
  - Some measures of fit provide confidence intervals (e.g., RMSEA)
- General Strategy
  - Report a selection of fit measures that capture different aspects of the data
  - Indicate rules of thumb chosen for interpretation and base on statistical literature (e.g., simulations, reviews, etc)
  - Rules of thumb vary with factors such as sample size & number of indicator variables
  - Compare fit relative to other competing models

Slide 34

Assessing Fit

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	16	11.271	12	0.506	0.939
Saturated model	28	0	0		
Independence model	7	359.252	21	0	17.107

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	0.969	0.945	1.002	1.004	1
Saturated model	1		1		1
Independence model	0	0	0	0	0

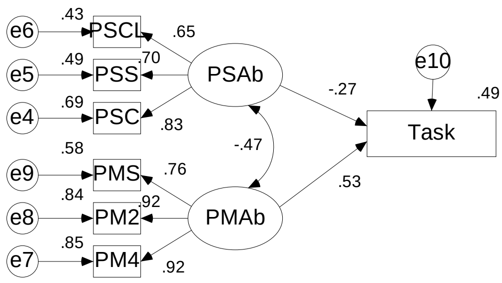
RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	0	0	0.098	0.702
Independence model	0.407	0.371	0.445	0

Default model: this is the model that was proposed.
Saturated model: all variances and covariances are allowed to freely vary.
Independence model: All covariances are constrained to be zero.
In the present example the proposed model (“Default model”) had a non-significant chi square (chi-square (df=12, n=98) = 11.2, ns.).
CFI and RMSEA were both very good.
The small sample size and small number of indicators to latent factors would be a potential qualification to these otherwise very good findings. The effects of the small sample size is evident in the large 90% confidence intervals around RMSEA with the upper limit suggesting a possibly poor fit (.098).

Slide 35

Structural Model



In this case the structural model is no different in terms of fit to the measurement model. The structural model does provide an estimate of the r-squared for task performance (49 % of variance explained).

Slide 36

Reporting

- Theoretical Rationale
- Sample correlation matrix with means and standard deviations
- Use a selection of fit measures
- Indicate the reasons for particular interpretive cut-offs used
- Report model coefficients
- Examine reporting in similar studies
- Examine guides in the literature on reporting recommendations

Slide 37

Further Issues about indicators

- Reflective indicators
  - Items are manifestation of underlying factor
  - E.g., personality, performance, job satisfaction
- Formative indicators
  - Factor is a composite of the indicators
  - e.g., SES; life stress scales; strategies in performance
- Item Parcelling

Many tests use many items to represent a single latent variable. The rationale is that a greater number of items increases the reliability. Specifying a large number of items can lead to problems in model estimation and issues with unidimensionality of scales. One somewhat controversial strategy for resolving this involves item parcelling. This is a technique for combining multiple items in a scale into parcels and then using the parcels as the indicators in the structural equation model. Several methods exist for parcelling including random assignment, grouping based on exploratory factor analysis loadings and grouping based on conceptual grounds.

Slide 38

More Advanced Applications

- Multiple Group Testing
- Latent Growth Curve Analysis
- Bootstrapped Standard Errors
- Bayesian analysis
- Ordinal and categorical variables
- Analysis of polychoric correlations
- Multi-Method Multi-Trait models