

INTRODUCTION TO SEM AND CFA ANALYSIS USING AMOS



LONG BUNTENG

For Post-Graduate Studies
Volume 2

Table of Contents:

1. Structural Equation Modelling **3**
2. The Concepts of Latent Constructs in Research **5**
3. The Minimum Sample Size in SEM **6**
4. Measurement Vs Structural Model **7**
5. Sample Amos Model **8**
6. Full Structural Model Analysis **9**
7. Interpreting the Output **11**
8. Modification Indices **12**
9. SEM Result Reports **13**
10. Report Structural Model Fit Parameters **14**
11. Reporting Hypothesized Relationships **15**
12. Confirmatory Factor Analysis (CFA) **16**
13. How CFA is different from EFA? **19**
14. Understanding Model Fit **21**
15. Improving Model Fit Modification Indices **22**
16. Improving Model Fit Standardized Residual Covariances **25**
17. Process of Assessing Model Fit **26**

Structural Equation Modelling

Structural equation modelling (SEM) is a set of statistical techniques used to measure and analyze the relationships of observed and latent variables.

A Structural Equation Model comprises of Measurement and Structural Model

Key Terminologies

Constructs: measure theoretical concepts that are abstract, complex, and cannot be directly observed by means of (multiple) items. Constructs are represented in path models as circles or ovals and are also referred to as latent variables.

Indicators: these are directly measured observations (raw data), also referred to as either items or manifest variables, which are represented in path models as rectangles. They are also available data (e.g., responses to survey questions or collected from company databases) used in measurement models to measure the latent variables.

Factor loadings: the bivariate correlations between a construct and the indicators. They determine an item's absolute contribution to its assigned construct. Loadings are of primary interest in the evaluation of

reflective measurement models but are also interpreted when formative measures are involved.

Latent variables: elements of a structural model that are used to represent theoretical concepts in statistical models. A latent variable that only explains other latent variables (only outgoing relationships in the structural model) is called exogenous, while latent variables with at least one incoming relationship in the structural model are called endogenous. Also see Constructs.

Critical t value: the cutoff or criterion on which the significance of a coefficient is determined. If the empirical t value is larger than the critical t value, the null hypothesis of no effect is rejected.

The Concept of Latent Constructs in Research

In science and social science researches, most of the times the researchers are dealing with latent constructs.

As has been said earlier, these constructs are measured using a set of items in a questionnaire. Since the Ordinary Least Squares (OLS) procedures could not entertain latent constructs, the researchers need to employ SEM for the analysis.

Using SEM, the researcher could model the relationship among these constructs together with their respective items in the model and analyze them simultaneously.

In this case, at least two measurement models involved – one for independent construct and the other one is for dependent construct.

The theorized link between measurement model for independent construct and measurement model for dependent construct is called a structural model.

Thus, instead of modeling the OLS regression and analyze using ANOVA, the researcher is working with the Structure Equation Modeling (SEM) and analyzed using software's like AMOS.

The Minimum Sample Size in SEM

There are endless debates in the literatures as to how many respondents should be obtained in order to employ SEM.

However, there is no clear-cut answers to it since every research differs (among other things) in term of the population characteristics, and the number of constructs employed in a model.

Hair et al. (2010), offer the following suggestion for minimum sample size depending on the model complexity and basic measurement model characteristics.

Sample Size in SEM

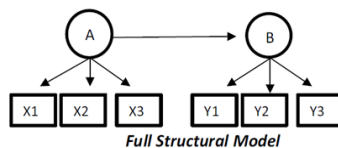
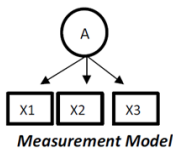
Model Characteristics (Number of latent constructs and items)	Minimum Sample Required
1. Five or less latent constructs. Each latent construct has more than three measuring items	100 sample
2. Seven or less latent constructs. Each construct has more than three items	150 sample
3. Seven or less latent constructs. Some constructs have less than three items (just identified model)	300 sample
4. More than seven latent constructs. Some constructs have less than three items (just identified model)	500 sample

Measurement vs Structural Model

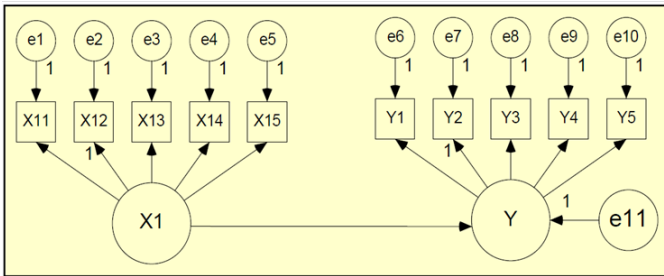
The measurement model in SEM is where the researcher is going to assess the validity of the indicators for each construct. After showing the validity of the measurement model, the researcher can proceed to the structural model.

The structural model is concerned with the influence and significance between constructs. The term “full structural model” means that the measurement and structural relationships of each construct are included in the model testing.

Parameters—the term “parameter” indicates the size and nature of the relationship between two objects in a model. Parameters can be fixed to a constant or can be estimated freely from the data. A parameter estimate will take place on the measurement level with indicators and error terms as well as on the structural level between constructs.



Sample AMOS Model



Note: X1 and Y are latent constructs. In Amos syntax, latent constructs are represented by the ellipses. The latent construct X1 is measured using items X11 to X15, while latent construct Y is measured using items Y1 to Y5. The measured items are represented by rectangles in the model. Sometimes (in literature) the measured items are called latent indicators.

Key: X1 = Exogenous construct, while X11 to X15 is a set of 5 items to measure latent construct X1

In the Amos diagram, e1 to e5 are errors in measurement for items X11 to X15

Y = Endogenous construct, while Y1 to Y5 is a set of 5 items to measure latent construct Y

In the Amos diagram, e6 to e10 are errors in measurement for items Y1 to Y5

Full Structural Model Analysis

A full structural model assesses the relationships between constructs but also includes the measurement indicators for each variable.

A full structural model will allow you to account for the measurement error in a construct's indicators while also assessing the relationships between constructs.

You will initially draw all the constructs and indicators like you did in the CFA, then you will start including the direct paths between constructs. This is a more robust model and will account for each indicator individually.

Unlike a composite variable path model, each indicator of a construct is included along with its effect on other constructs.

Just like in the path model, you need to include error terms for each dependent construct and to make sure all error terms are labeled. In the CFA, all constructs were considered independent constructs, but that is not the case in a full structural model.

A construct is considered an independent if it has a structural relationship that influences another construct and is not being influenced by any other construct in the model.

In this example, Authentic Leadership and Ethical Leadership are the only independent variables. Since

these two constructs are considered independent, a covariance needs to be added between them.

After adding the structural relationships and labeling all error terms, you are ready to run the analysis. A full structural analysis will give you not only the structural relationships between constructs but also the measurement properties or factor loadings for each construct.

Interpreting the Output

- After running the analysis, let's go to the output.
- In the "Estimates" link, the results will be presented just like the one you got with the CFA analysis.
- Here are the unstandardized regression weights and significance for your structural relationships.
- Here are the standardized regression weights for your relationships.
- With a simple path analysis, you want to report the standardized weights so that you can compare the strength of relationships.

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	PLabel
LST <- AU	.019	.118	.161	par_14
LST <- BET	.657	.196	3.354	*** par_15
AT1 <- AU	1.000			
AT1 <- AU	.723	.065	11.211	*** par_1
AT1 <- AU	.938	.063	14.966	*** par_2
AT1 <- AU	.902	.066	13.685	*** par_3
BET <- BET	1.000			
BET <- BET	1.212	.083	14.682	*** par_4
BET <- BET	1.149	.081	14.199	*** par_5
BET <- BET	1.091	.072	15.137	*** par_6
BET <- BET	1.235	.080	15.723	*** par_7
BET <- BET	1.165	.079	14.775	*** par_8
LST <- LST	1.000			
LST <- LST	1.107	.059	18.831	*** par_9
LST <- LST	1.023	.056	18.199	*** par_10
LST <- LST	.883	.059	14.953	*** par_11
AT4 <- AU	.569	.061	9.365	*** par_13

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
LST <- AU	.019
LST <- BET	.569
AT1 <- AU	.807
AT1 <- AU	.596
AT1 <- AU	.766
AT1 <- AU	.710
BET <- BET	.714
BET <- BET	.812
BET <- BET	.786
BET <- BET	.836
BET <- BET	.869
BET <- BET	.816
LST <- LST	.819
LST <- LST	.873
LST <- LST	.849
LST <- LST	.732
AT4 <- AU	.569

Analyzing the Model Fit

All our relationships were significant, so let's next look at the model fit.

You can see the model fit is very good across all the different model fit indices.

You can see how much better the model fits the data when a full structural model is used compared to a path model. The ability to model measurement error on the indicator level allows for greater explanation of the variance in the model.

Hence, we have a model that fits the observed data.

CMIN

Model	NPAR	CMIN	DF	PCMIN/DF
Default model	14	121.118	14	.000
Unrestricted model	120	.000	6	
Independence model	15	1338.458	.000	92.146

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.086	.915	.857	.494
Unrestricted model	.000	1.000		
Independence model	.628	.242	.154	.212

Baseline Comparisons

Model	NPAR	RFI	IFI	TLI	CFI
Default model	14	.915	.918	.918	.918
Unrestricted model	1.000	1.000	1.000	1.000	1.000
Independence model	.000	.000	.000	.000	.000

Paradigmatic Adjusted Measures

Model	PRATIO	PNFI	PNFI
Default model	.819	.765	.761
Unrestricted model	.000	.000	.000
Independence model	1.000	.000	.000

NCF

Model	NCF	LO IN	RM IN
Default model	119.818	99.174	118.256
Unrestricted model	.000	.000	.000
Independence model	120.658	1106.751	1487.686

FMN

Model	FMN	FMN	FMN
Default model	.844	.769	.765
Unrestricted model	.000	.000	.000
Independence model	1.000	.000	.000

RMSEA

Model	RMSEA	LO IN	RM IN
Default model	.048	.037	.039
Unrestricted model	.000	.000	.000
Independence model	.399	.399	.399

Modification Indices

Next, let's look at the modification indices. With the full structural model, we should have already assessed the measurement model when we performed a CFA; thus, our focus is not on adjusting/covarying error terms with constructs.

At this point, the modifications that concern us are the regression weights between constructs. As you can see, there are no modifications that are substantial or worthy of consideration.

The modification indices will suggest unacceptable alterations such as indicators having structural relationships with other indicators. Even if the modification index is high, these are nonsensical suggestions and should not be considered.

Regression Weights: (Group number 1 - Default model)

			MLPar Change
LS1 <--	AT3	18.228	.159
LS2 <--	LS1	4.264	.067
LS3 <--	LS1	4.344	-.068
LS3 <--	AT3	9.063	-.095
LS4 <--	AT4	7.084	.096
BE1 <--	LS3	4.101	.068
BE1 <--	LS4	5.192	.080
BE1 <--	AT3	4.483	-.072
BE2 <--	AT4	4.393	-.070
BE4 <--	AT1	4.788	-.071
BE6 <--	AT3	11.321	.131
AT2 <--	AT3	9.515	.118
AT2 <--	AT4	8.007	.117
AT3 <--	LST	8.718	-.172
AT3 <--	LS2	6.310	-.115
AT3 <--	LS3	16.549	-.178
AT3 <--	LS4	7.188	-.122
AT3 <--	BE1	7.113	-.116
AT3 <--	AT2	5.418	.102
AT3 <--	AT4	7.443	.131
AT4 <--	LS4	4.625	.095
AT4 <--	AT3	6.283	.109
AT5 <--	BET	5.438	.124
AT5 <--	LST	5.062	.109
AT5 <--	LS3	8.324	.105
AT5 <--	BE1	8.858	.107
AT5 <--	BE2	7.675	.098
AT5 <--	BE3	4.857	.087
AT5 <--	BE5	4.447	.072
AT5 <--	AT4	6.351	-.100

SEM Results- Reports

With the structural model, it is a good idea to present standardized regression weights, t-values, model fit statistics, and R² values. I like to present R² values of the dependent constructs so the reader can see how much of the variance is being explained with my independent variables.

I also like to state if my hypotheses were supported or not. I feel like there is no need to present the measurement indicators again because you should have presented this with a confirmatory factor analysis prior to performing the structural analysis.

Ultimately, you want to say a lot in a little amount of space. Presenting your results in a table should be done in a manner where the reader can see all the necessary information to determine the final results of your SEM test. This is not the only way to present structural results, but it is a good template.

Structural Model Results are reported in two steps.

First, The Fit Indices are reported for the model

Second, the hypothesized relationships are assessed for significance and results reported accordingly. Following two sections present a sample of reporting full structural model.

Report Structural Model Fit Parameters

A structural equation model generated through AMOS was used to test the relationships. A good-fitting model was accepted if the value of the CMIN/df, the goodness-of-fit (GFI) indices (Hair et al., 2010); the Tucker and Lewis (1973) index (TLI); the Confirmatory fit index (CFI) (Bentler, 1990) is ≥ 0.90 (Hair et al., 2010). In addition, an adequate-fitting model was accepted if the AMOS computed value of the standardized root mean square residual (RMR) < 0.05 , and the root mean square error approximation (RMSEA) is between 0.05 and 0.08 (Hair et al., 2010). The fit indices for the model shown in Table 1 fell within the acceptable range: CMIN/df = , the goodness-of-fit (GFI) = , TLI = , CFI = , SRMR = , and RMSEA = .

The squared multiple correlation was _____ for life satisfaction, this shows that _____ variance in life satisfaction is accounted by _____ and _____.

Reporting Hypothesized Relationships

The study assessed the impact of authentic and ethical leadership on life satisfaction. The impact of Authentic Leadership on life satisfaction was positive and significant ($b = \underline{\hspace{1cm}}$, $t = \underline{\hspace{1cm}}$, $p < \underline{\hspace{1cm}}$), supporting H1. The impact of Ethical Leadership on self-efficacy was positive and significant ($b = \underline{\hspace{1cm}}$, $t = \underline{\hspace{1cm}}$, $p < \underline{\hspace{1cm}}$), supporting H2. Model fit indices and Hypotheses results are presented in Table.

Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis (CFA) is a statistical technique that analyzes how well your indicators measure your unobserved constructs and if your unobserved constructs are uniquely different from one another.

In a CFA, an unobservable construct is often referred to as a “factor”.

So, when we use the term “factor”, it represents an unobservable construct we are trying to measure.

In a diagram format, an unobserved variable is represented by a circle or oval.

The indicators that measure the unobserved variable will have single headed arrows coming from the unobserved construct to each of the indicators.

Each indicator is represented by a square or a rectangle.

The single arrow line directly from the factor, or unobservable construct, to the indicator represents the influence that is reflected from the unobservable construct to its indicators.

Statistical estimates of these direct effects are called “factor loadings” and are interpreted as regression

coefficients that may be in unstandardized or standardized form.

Indicators that “reflect” an unobservable construct or factor are called simply reflective indicators.

With most indicators, a level of error is present. Each measurement error of an indicator represents “unique variance”, or variance not explained by the factor level.

Measurement errors are represented by two types of unique variance: random error and systematic error.

Random error occurs when the measurement of a concept is unpredictably fluctuating due to unforeseen or random influences. This type of error is considered “noise” because there is often no consistency to the measurement error.

Systematic error is different in that the error is systematic in nature. The results of a measurement can be consistently high or low due to a systematic influence in the measurement of a concept.

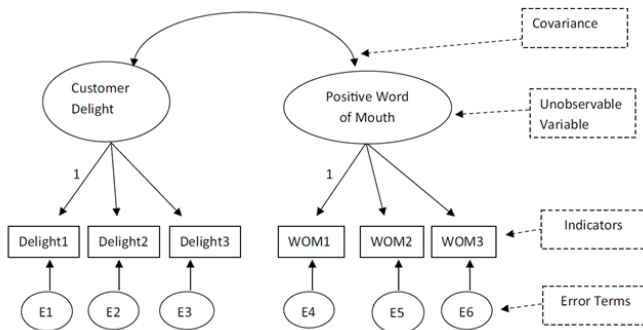
For instance, taking a survey online as compared to a paper and pencil survey may systematically influence the way respondents answer questions about a construct. With systematic error, this is often classified as a “bias” compared to the “noise” of random error. Note that if an unobserved variable has only one indicator, then the error term for the single indicator variable is constrained to have a mean of 0 and a variance of 0. (You must assume the indicator has no measurement error—which is often a very flawed assumption.)

In a CFA, you also want to account for “unmeasured covariance”. To do this, a two-headed arrow is drawn between all independent unobservable variables.

In a CFA, the analysis will treat all unobserved variables as exogenous or independent variables.

Thus, you will draw a covariance double headed arrow between each unobservable construct in a CFA.

If you don’t do this, AMOS will give you an error message stating that a covariance should be drawn between all unobserved constructs in your CFA.



Source: Collier, J. E. (2020). Applied structural equation modeling using AMOS: Basic to advanced techniques. Routledge.

How CFA is different from EFA?

An exploratory factor analysis (EFA) is useful in data reduction of a large number of indicators and can be quite helpful in seeing if indicators are measuring more than one construct.

EFAs are typically the first step in determining if an indicator is measuring a construct. In an EFA, the researcher is not denoting which indicators are measuring a construct.

The analysis simply tries to let every indicator load on every construct. This is where “cross loading” can occur, where an indicator is actually loading strongly on more than one construct. This is a concern that your indicator is not a strong measure of the specified construct.

In contrast, a CFA does not allow an indicator to load on more than one construct. The researcher prior to the analysis will specify what the indicators are for each construct, and those indicators can load only on that specific construct.

Other ways that an EFA and CFA are different is an EFA is typically conducted with correlation matrices, which can be problematic in comparing parameters across samples.

CFA uses a covariance matrix and is more adept at handling comparisons across samples.

An EFA will also consider data rotation that is often done to achieve a better loading of indicators on a

construct or sometimes a reduction of cross loading with other constructs.

A CFA does not worry about rotation because you are denoting the specific items that are loading on a construct.

Understanding Model Fit

The researcher first specifies a model based on theory, then determines how to measure constructs, collects data, and then inputs the data into the SEM software package.

The package fits the data to the specified model and produces the results, which include overall model fit statistics and parameter estimates.

Model fit is the feasibility of the model with the data.

Model Fit Criteria

Measure	Name	Description	Good fit if:
X ²	Model Chi Square	Assess overall fit and the discrepancy between the sample and fitted covariance matrices. Sensitive to sample size. H0: The model fits perfectly.	p-value > 0.05
(A)GFI	(Adjusted) Goodness of Fit	GFI is the proportion of variance accounted for by the estimated population covariance. Analogous to R ² . AGFI favors parsimony.	GFI ≥ 0.95 AGFI ≥ 0.90
(N)NFI TLI	(Non)Normed Fit Index Tucker Lewis index	An NFI of .95, indicates the model of interest improves the fit by 95% relative to the null model. NNFI is preferable for smaller samples. Sometimes the NNFI is called the Tucker Lewis index (TLI)	NFI ≥ 0.95 NNFI ≥ 0.95
CFI	Comparative Fit Index	A revised form of NFI. Not very sensitive to sample size. Compares the fit of a target model to the fit of an independent, or null, model.	CFI ≥ .90
RMSEA	Root Mean Square Error of Approximation	A parsimony-adjusted index. Values closer to 0 represent a good fit.	RMSEA < 0.08
(S)RMR	(Standardized) Root Mean Square Residual	The square-root of the difference between the residuals of the sample covariance matrix and the hypothesized model. If items vary in range (i.e. some items are 1-5, others 1-7) then RMR is hard to interpret, better to use SRMR.	SRMR < 0.08
AVE (CFA only)	Average Value Explained	The average of the R ² s for items within a factor	AVE > .5

Source: Principles and Practice of Structural Equation Modeling. Rex B. Kline. 2005

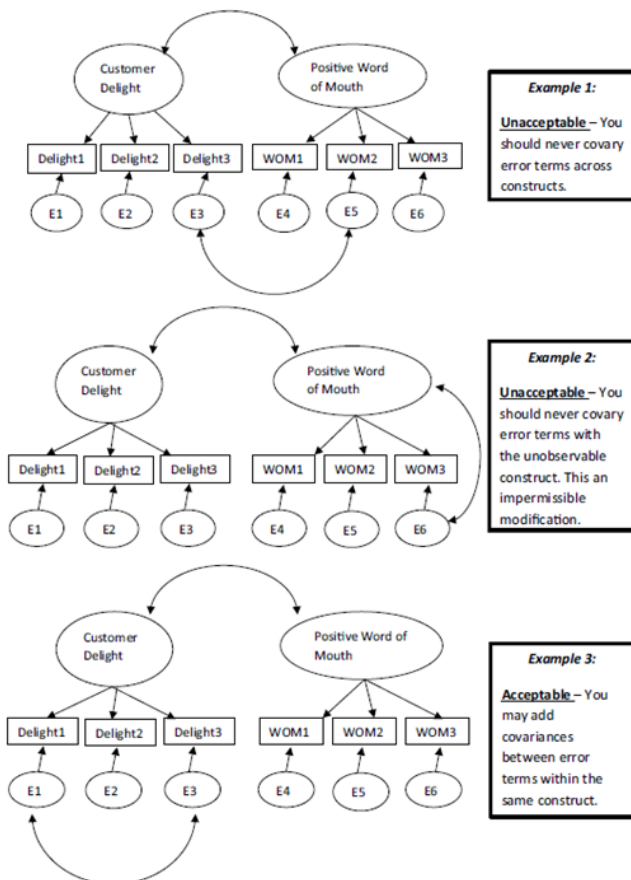
Improving Model Fit

Modification Indices

Check for outer loadings, loadings that are too low shall be removed. Start by removing loadings less than .50 and re-run the model when an item removed to assess the model fit.

Check for modification indices. However, be cautious, not every suggested covariance shall be drawn.

Assess Standardized Residual Covariance (Estimates Section)



Source: Applied structural equation modeling using AMOS: Basic to advanced techniques by Collier, J. E. (2020).

In instances where an unobserved variable has multiple indicators, you might want to correlate the error terms if it is theoretically justifiable and helps to explain the variance within the construct.

Frequently, indicators within a construct are very similar to one another and there is redundancy between the

indicators. How big does the modification indices value need to be to add another covariance?

– Consider any valid modification suggestion if the chi-square value is not changing by at least 10. A chi-square value of 10 at the expense of one degree of freedom is almost significant at a .001 level.

I look for instances where the change is extreme. It is not uncommon to have a modification suggestion greater than 20 in large models.

An equal concern for multiple high modification indices within a construct is the possibility that this shared unexplained error is the result of a method bias in how the data was collected.

Improving Model Fit

Standardized Residual Covariances

Standardized residual covariances indicate the standardized differences between the proposed covariances based on the model and the observed covariance matrix computed based on the collected data.

Significant values of standardized covariance indicate significant differences in covariance between proposed model based computed covariance and observed covariance and affects the overall goodness of fit.

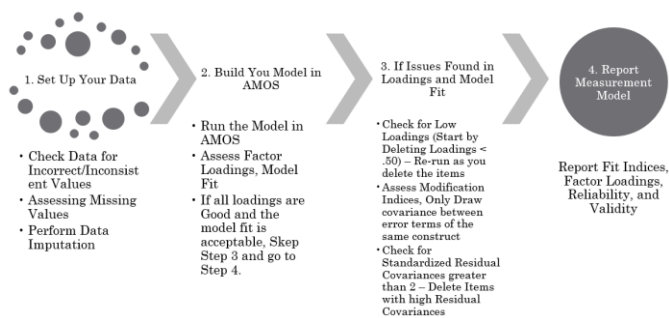
Although, the standardized residual covariance are like modification indices, they can directly be addressed through deletion of the concerned item. Therefore, the significant standardized residuals may be treated only after taking care of modification indices.

Look for standardized residual covariance estimates greater than 2 to identify particular covariances which seem associated with significant discrepancies of the given model from a perfect model.

Be Cautious: Do not Delete an item which is central to your construct and can affect Content Validity

Make sure you have at least 3 items per Construct.

Process of Assessing Model Fit



The End of the Book