Five things you need to know before learning Structural Equation Modeling

by [guest](https://www.theanalysisfactor.com/author/guest/) [1 Comment](https://www.theanalysisfactor.com/five-things-you-need-to-know-before-learning-structural-equation-modeling/#comments)

*By Manolo Romero Escobar*

If you already know the principles of general linear modeling (GLM) you are on the right path to understand Structural Equation Modeling (SEM).

As you could see from my [previous post,](https://www.theanalysisfactor.com/why-need-knowledge-of-multiple-regression-to-understand-sem/) SEM offers the flexibility of adding paths between predictors in a way that would take you several GLM models and still leave you with unanswered questions.

It also helps you use latent variables (as you will see in future posts).

GLM is just one of the pieces of the puzzle to fit SEM to your data. You also need to have an understanding of:

**1. Substantive knowledge**

You must understand your research domain.

No consultant or statistician would master the constructs and theories as well as you do.

The numbers don’t remember where they come from, as Lord said. But you need to know where you want to take them to.

At the end of the day, the winning model among equally good fitting ones may be the one that *makes more sense in terms of the theory behind it*.

**2. Interpretation of coefficients in multiple regression**

Yes, it is part of GLM, so you already know this.

However, it doesn’t hurt to brush up on how to interpret [standardized](https://www.theanalysisfactor.com/how-to-get-standardized-regression-coefficients/) and [unstandardized regression coefficients](https://www.theanalysisfactor.com/video-interpreting-linear-regression-coefficients-walk-through-output/). You’ll use both often.

**3. Statistical tests and significance**

The technical aspect, at least. Mainly that if you divide any statistic by its standard error you get a ratio that can be interpreted as a t-test.

Software will provide you with the estimate of the statistic and their standard errors, and their p-values. It’s better if you have an idea of where those p-values come from. Also, go back to your notes on intro to stats about [chi-square tests](https://www.theanalysisfactor.com/difference-between-chi-square-test-and-mcnemar-test/).

**4. Data screening**

You may want to jump the gun on running SEM on your 1000-case data set. You had better stop and check for multivariate [normality](https://www.theanalysisfactor.com/anatomy-of-a-normal-probability-plot/).

Easier said than done when you have 20 indicators. But even a simple sanity check on the data can save you much frustration in the long run.

Always scan for wrongly coded [missing values](http://thecraftofstatisticalanalysis.com/approaches-missing-data) and [out of range cases](https://www.theanalysisfactor.com/outliers-to-drop-or-not-to-drop/).

A quick look at the matrix of correlations can also give you hints on where potential problems you will find when fitting your model.

**5. The effect of sample size**

You will find many rules of thumb about how many cases by factors you need for different SEM models.

The truth is, until you have collected the data you will not know if it is sufficient. The reason being that well behaved data makes estimation easier even in small [sample sizes](https://www.theanalysisfactor.com/resources/by-topic/effect-size-statistics-power-and-sample-size-calculations/).