What is Structural Equation Modeling?

1. Confirmatory analysis – implies you have an idea before hand of what you want to test
   1. Usually discussed as causal – although this view is challenged
2. How it works?
   1. Regression on crack
   2. Modeled in a picture
3. Goodness of fit – way to test if the model is “publishable”
   1. Relationship to null hypothesis testing:
      1. Good fit (small numbers/large numbers) – reject null
      2. Bad fit – retain null
4. It’s good stuff:
   1. Confirmatory – makes you think about theory before you start (although hypothesis testing with EFA is not *impossible* it’s just different).
   2. Error variance estimation – allows you to control for error variance, examine it, instead of it just being a big pain
   3. Latent variables – you can incorporate unobserved variables into your analysis.
5. Definitions:
   1. Latent variables – unobserved variables, circles on the AMOS diagrams
      1. Sometimes called factors
      2. Example: IQ is a factor that causes what?
      3. Cannot be measured directly
   2. Observed variables – things you measured directly, boxes on the AMOS diagram
      1. Way you measure the latent variables directly
   3. Exogenous latent variables – they are the *cause* of an effect
      1. Influence the values of other latent variables
      2. The changes in these variables thought to be caused by external factors
   4. Endogenous – more like dependent variables, they are the effects of other variables (exogenous latent variables)
      1. Changes in these variables explained by the model
6. Diagram this!
7. Factor analysis
   1. EFA – the link between observed and latent is unknown or untested
      1. For example, you might think you have a 3 part model, but you have to test it first
      2. You figure out how your scale (or other stuff) works, which things group together and based on that grouping, you draw a conclusion on the underlying factors (tfear example)
      3. This process usually is motivated by theory, but you usually start by exploring the items and their variance.
   2. CFA – you have an expected relationship between the observed and latent variables based on theory/past research
   3. Measurement model – when you are using only measured observed variables and their relationship to factors (exogenous latent variables).
8. Fully latent model
   1. This model is even more specific than the information above.
   2. You can predict the causal direction between two latent factors.
   3. Called the full model because it has two parts:
      1. Links between latent factors and observed factors
      2. Links between latent factors
   4. Recursive model – specifies the direction – only one direction
   5. Nonrecursive model – allows for feed back, goes both directions
9. Basic statistical mumbo jumbo
   1. Process:
      1. You have a model
      2. You collect the data based on your observed variables
      3. You test if the data and your model play nice ☺
   2. *We can never do this perfectly…people are weird*.
      1. The weird stuff is called the “residuals” or error.
      2. Data = model + residuals
         1. Data = actual scores on observed variables
         2. Model – the structure you are trying to test
         3. Residual – the stuff you cannot account for.
   3. Ways to go about SEM:
      1. Have a theory > test the data > bust > cry = strictly confirmatory
      2. Have multiple theories > test the data > pick the one that works best = alternative models
      3. Have a theory > test the data > eeww messy > test different paths > test stuff again > new pretty model reinterpret theory = model generating
10. The way it looks:
    1. Circles = latent factors
       1. Endogenous do have an arrow head going into them.
       2. Exogenous do *not* have an arrow head going into them.
       3. Also error variances are circles.
    2. Rectangles = observed variables
    3. Single head arrows – represent a causal direction (y = x + error)
    4. Double head arrows – represent correlation or covariance, relationship does not have a predicted direction
    5. (ps the picture on page 9 does not make sense to me)
11. Path diagrams
    1. The visual picture of a model
    2. Talk about pictures here
    3. Measurement error – error as part of an observed variable
       1. Random error measurement
       2. Error uniqueness – usually called systematic error variance
       3. Residual – error term for latent variables
       4. Since error is not really measured (in a sense), they are circles on the diagram
    4. One way arrows are structural regression coefficients
       1. Usually interpreted in the same was a regular B regression coefficients
    5. One way error arrows represent the impact of measurement error on the observed variables
       1. You want these values to be low
       2. Sort of like the unaccounted for variance in a regression equation
    6. Two way arrows represent covariance or correlation between two variables
       1. Think about this like you would a regular correlation number
12. Structural equations
    1. Since this is regression, you can think about it like a regression equation
    2. Especially important for factor scores (in multigroup models!)
13. Measurement and structural model
    1. Measurement model is the relationship between the latent and observed variables
    2. Structural model is the relationship between the latent variables.