Confirmatory Factor Analysis Part 3: Higher-Order Factors

- Today's topics:
  - Higher-Order Factor Models
  - The Problem of Equivalent Models
  - Wrapping Up…
  - Diagrams for HFS Example Models
  - See Brown p. 145-6, 155-6 for a description of what all should be reported from a CFA analysis…

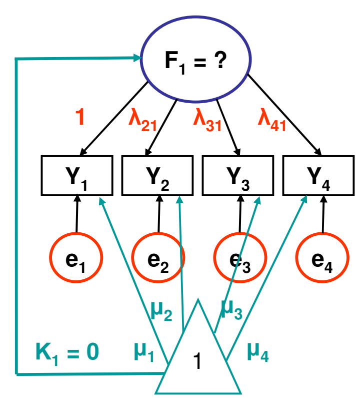
Sequence of Steps in Confirmatory Factor Analysis

1. Specify your measurement model(s)
   - How many factors, which items load on which factors, and whether your need any method factors or error covariances
   - For models with large numbers of items, you should pry start by modeling each construct in a separate analysis
2. Assess model fit, per factor, if possible (if 4+ indicators)
   - Global model fit: Does a one-factor model adequately fit each set of indicators thought to measure the same latent construct?
   - Local model fit: Do the normalized residuals suggest any points of 'strain'? Any items not loading significantly that you might drop? Any *justifiable* additional cross-loadings or error correlations?
3. Once your single-factor measurement models are in good condition, it's time to consider the (higher-order) structural model

Higher-Order Factor Models

- Question: What kind of higher-order factor structure best accounts for the **pattern of covariance among the measurement model factors**?
  - In other words, what should the **structural model** look like?
  - Your decide whether the measurement model factors should be allowed to correlate or not (presuming the correlation is < 1 if so)
  - Just like for the items, there may be a common factor that accounts for the covariance among measurement model *factors*
  - **A higher-order factor** would be suggested by similar magnitude of correlations across the measurement model factors
    - A higher-order factor would suggest that a 'total score' (across subscales) is a reasonable thing to construct
    - Can 'rescue' a multi-dimensional construct that was supposed to be unidimensional (i.e., maybe can still fit a single higher-order factor)

Suggested Measurement Model Scaling to fit Higher-Order Factors



"Marker Item" for item loadings
→ Fix 1 item loading to 1
→ Estimate factor variance

Factor variance will become "factor variance leftover" (so it can't be a fixed quantity).

"Z-Score" for item intercepts
→ Fix factor mean to 0
→ Estimate all intercepts

All the factor means will be 0, and the item intercepts will just be their means (→ plausible values).

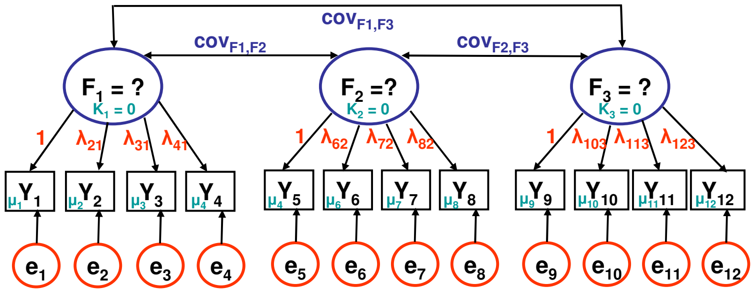
Identifying a 3-Factor Structural Model
Option 1: 3 Correlated Factors

Measurement Model for Items: (over-identified)

9 loadings, 12 intercepts
12 item error variances
df = 90 – 33 = 57

Structural Model for Factors: (just-identified, no factor means)

3 factor variances
3 factor covariances
df = 6 – 6 = 0



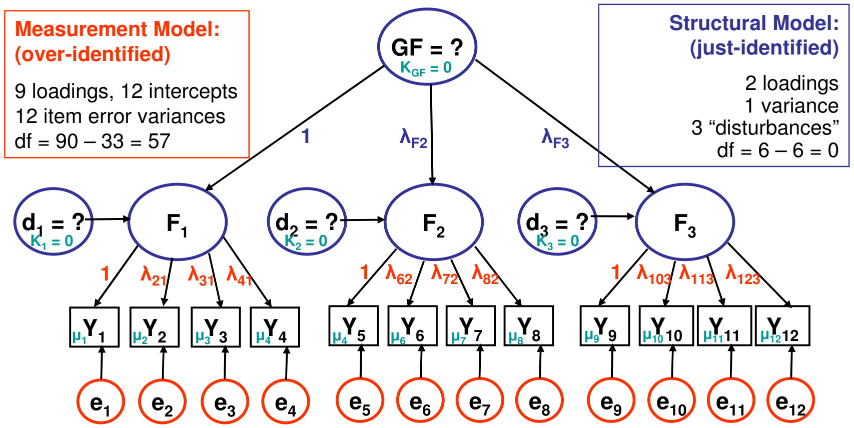
Option 2: 3 Factor "Indicators"
(Marker Approach)

Measurement Model: (over-identified)

9 loadings, 12 intercepts
12 item error variances
df = 90 – 33 = 57

Structural Model: (just-identified)

2 loadings
1 variance
3 "disturbances"
df = 6 – 6 = 0



Note: Factor error variances are called "disturbances" (d's instead of e's)
Unlike other programs, Mplus doesn't differentiate factor variances from disturbances
Predicted factor means are called "factor intercepts"

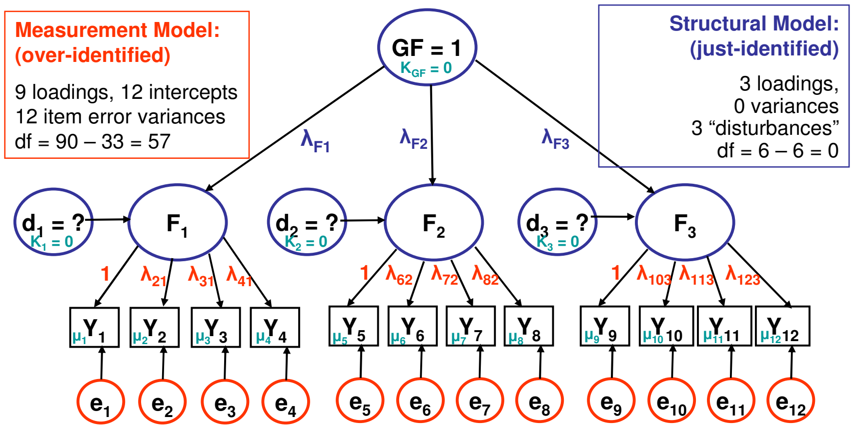
Option 3: 3 Factor "Indicators"
(Z-Score GF Approach)

Measurement Model: (over-identified)

9 loadings, 12 intercepts
12 item error variances
df = 90 – 33 = 57

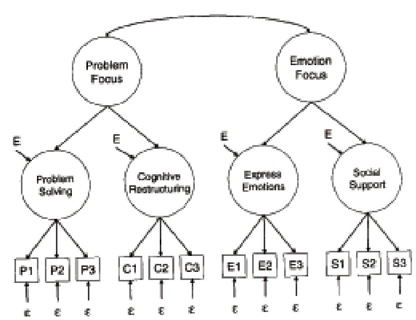
Structural Model: (just-identified)

3 loadings,
0 variances
3 "disturbances"
df = 6 – 6 = 0



All 3 options will yield equivalently fitting (but slightly rearranged) models.

Structural Model Identification: 2 Factor "Indicators"



Measurement Model for Items:

8 loadings, 12 intercepts
12 item error variances
df = 90 – 32 = 58

Structural Model for Factors:

2 factor loadings
2 factor variances
1 factor covariance
4 factor disturbances
df = 10 – 9 = 1

This model will not be structurally identified without the covariance between the higher-order factors.

Structural Model Identification

- Same principles as measurement model identification:
  - Total possible number of structural parameters depends on number of **factor variances and covariances** (NOT items)
  - **2 measurement model factors → Under-identified**
    - They can be correlated, which would be just-identified… that's it
  - **3 measurement model factors → Just-identified**
    - They can all be correlated OR a single higher-order factor can be fit
    - Some variance estimate per factor (so, 3 total) in either option
    - Factor variances and covariances will be perfectly reproduced
  - **4 measurement model factors → Can be over-identified**
    - They can all be correlated (6 correlations required; just-identified)
    - They can have a higher-order factor (4 loadings; over-identified)
    - Some variance estimate per factor (so, 4 total) in either option
    - **The fit of the higher-order factor can now be tested**

Examples of Structural Model Hypothesis Testing

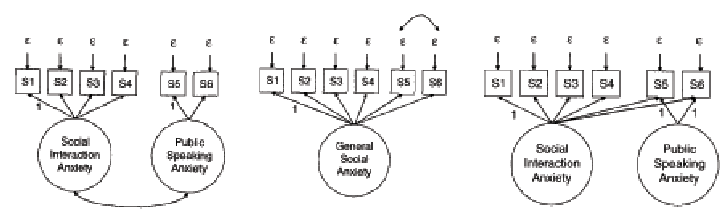
- Do I have a higher-order factor of my subscale factors?
  - If 4 or more subscale factors: Compare fit of alternative models
    - Baseline: All 6 factor covariances estimated freely
      Alternative: 1 higher-order factor instead (so 2 df left over)
  - If 3 (or fewer) subscale factors: CANNOT BE DETERMINED
    - Baseline and alternative models will fit equivalently
- Do I have 1 factor or 2 factors?
  - Use z-score approach to factor identification (both variances = 1)
    - Baseline: Two factors, estimated covariance (so is correlation)
      Alternative: Two factors, covariance fixed to 1 (so correlation = 1)
- Do I have "content factors" or "method factors"?
  - Stay tuned for an example of this scenario

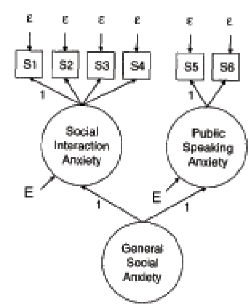
Equivalency across Models

- Remember, the purpose of a measurement model is to reproduce the observed matrix of variances and covariances of the items
- This means that models that generate the same predicted covariance matrix are equivalent models
- This will often not be comforting, but it is the truth…
- Here's an example: These models make very different theoretical statements, but they will nevertheless fit the same.

- Generally speaking, the fewer df left over (i.e., the more complicated the model), the more equivalent alternative solutions there are.

More Equivalent Models…

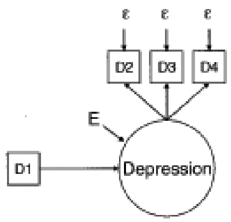


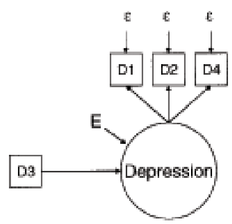


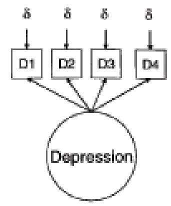
The idea here is that there is a general factor of "Social Interaction Anxiety", in part measured by two items dealing with public speaking.

The extra relationship between the two public speaking items can be modeled in many different, yet equivalent ways…

Even More Equivalent Models…







Right: One can think these 4 items as "effects" (indicators) of depression…

Left: One can think of any one item as "causing" depression and the others as "effects" of depression…

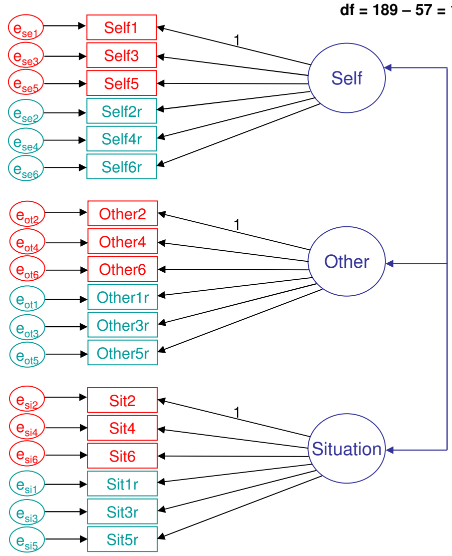
Point of the story: CFA/SEM cannot give you TRUTH. Contrary to what it's sometimes called, it is not really "causal modeling".

Wrapping Up…

- Fitting measurement and structural models are two separate issues:
  - **Measurement model**: Do my lower-order factors account for the *observed covariances among my ITEMS*?
  - **Structural model**: Do higher-order factors account for the *estimated covariances among my measurement model FACTORS*?
    - Higher-order factors support the idea of constructing a 'total score'
- Figure out the measurement models FIRST, then structural models
  - Recommend fitting measurement models separately per factor, then bringing them together once you have each factor settled
  - This will help to pinpoint the source of misfit in complex models
- Keep in mind that structural models may not be 'unique'
  - Mathematically equivalent models can make very different theoretical statements, so there's no real way to choose between them if so…

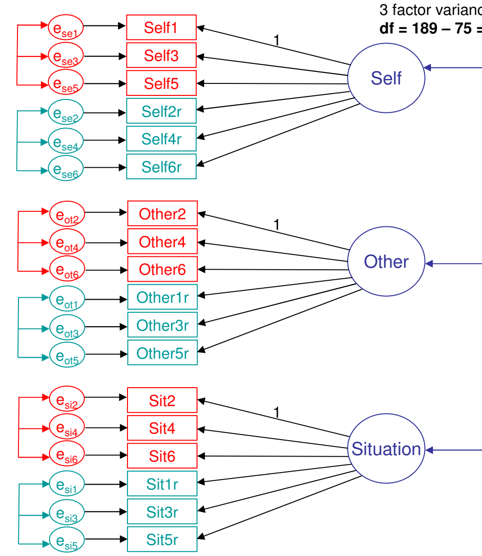
Model 1. Three correlated factors of Self, Other, and Situation (ignoring wording variance for now)

Total possible df for 18 items = 189
15 item loadings, 18 intercepts, 18 error vars
3 factor variances, 3 factor covariances
df = 189 – 57 = 132



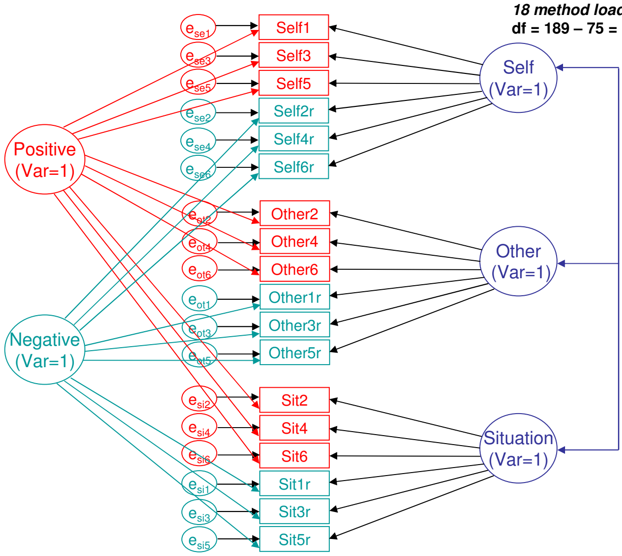
Model 2. Three correlated factors of Self, Other, and Situation (with error correlations among items of same valence)

Total possible df for 18 items = 189
15 item loadings, 18 intercepts, 18 error vars
18 error covariances
3 factor variances, 3 factor covariances
df = 189 – 75 = 114



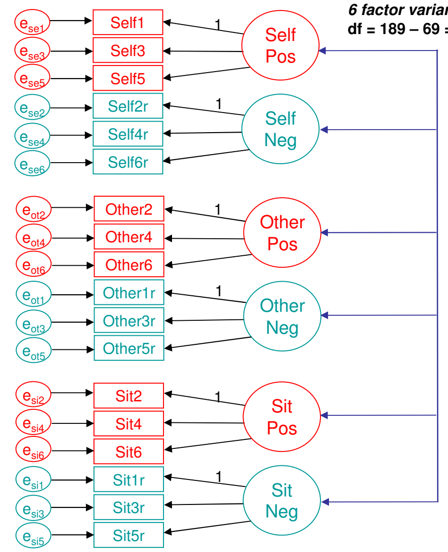
Model 3. Three correlated factors of Self, Other, and Situation
Two uncorrelated wording factors

Total possible df for 18 items = 189
18 item loadings, 18 intercepts, 18 error vars
0 factor variances, 3 factor covariances
18 method loadings
df = 189 – 75 = 114



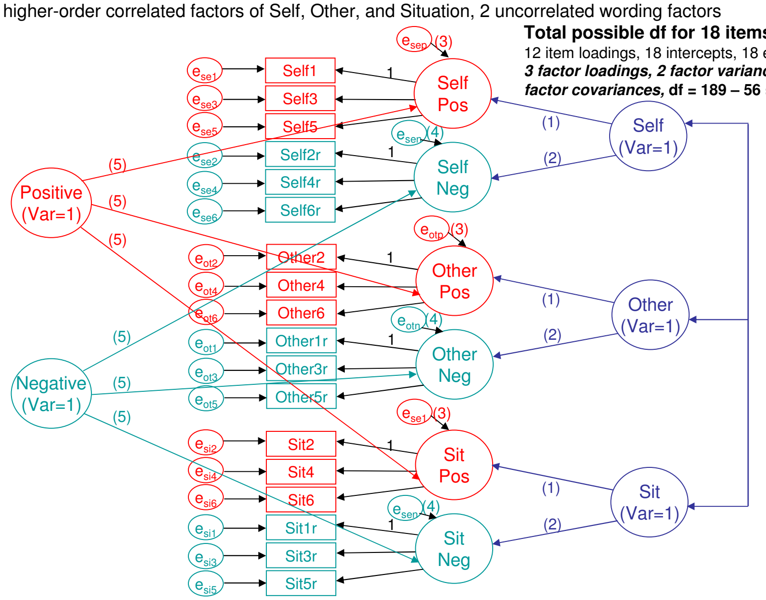
Model 4. Six lower-order factors for positive and negative self, other, and situation forgiveness,
allowed to be freely correlated

Total possible df for 18 items = 189
12 item loadings, 18 intercepts, 18 error vars
6 factor variances, 15 factor covariances
df = 189 – 69 = 120



Model 5. Six lower-order factors for positive and negative self, other, and situation forgiveness,
3 higher-order correlated factors of Self, Other, and Situation, 2 uncorrelated wording factors

Total possible df for 18 items = 189
12 item loadings, 18 intercepts, 18 error vars
***3 factor loadings, 2 factor variances, 3 factor covariances*, df = 189 – 56 = 133**



Model 6. Forgiveness and Unforgiveness as constructs; self, other, situation as methods factors

Total possible df for 18 items = 189
12 item loadings, 18 intercepts, 18 error vars
***4 factor loadings, 3 factor variances*, 3 factor covariances, df = 189 – 56 = 133**