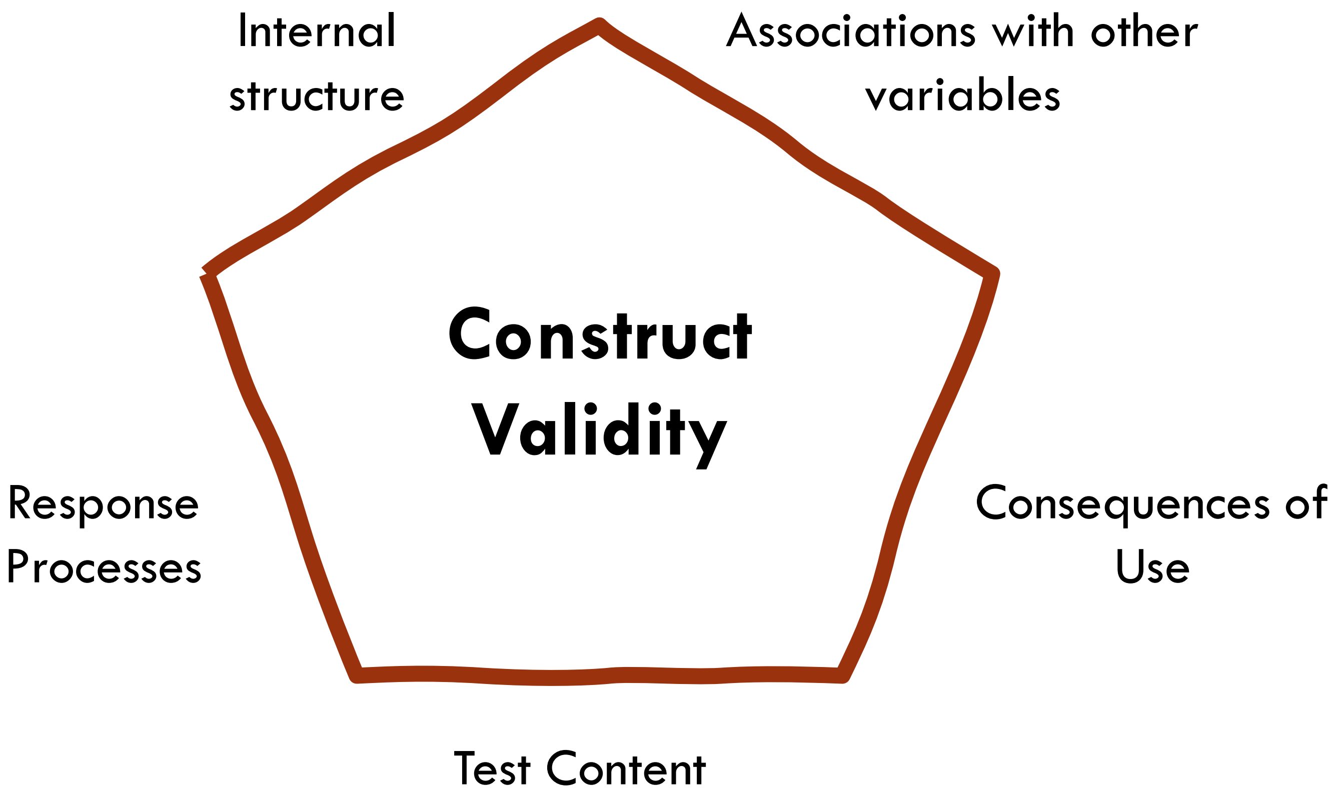
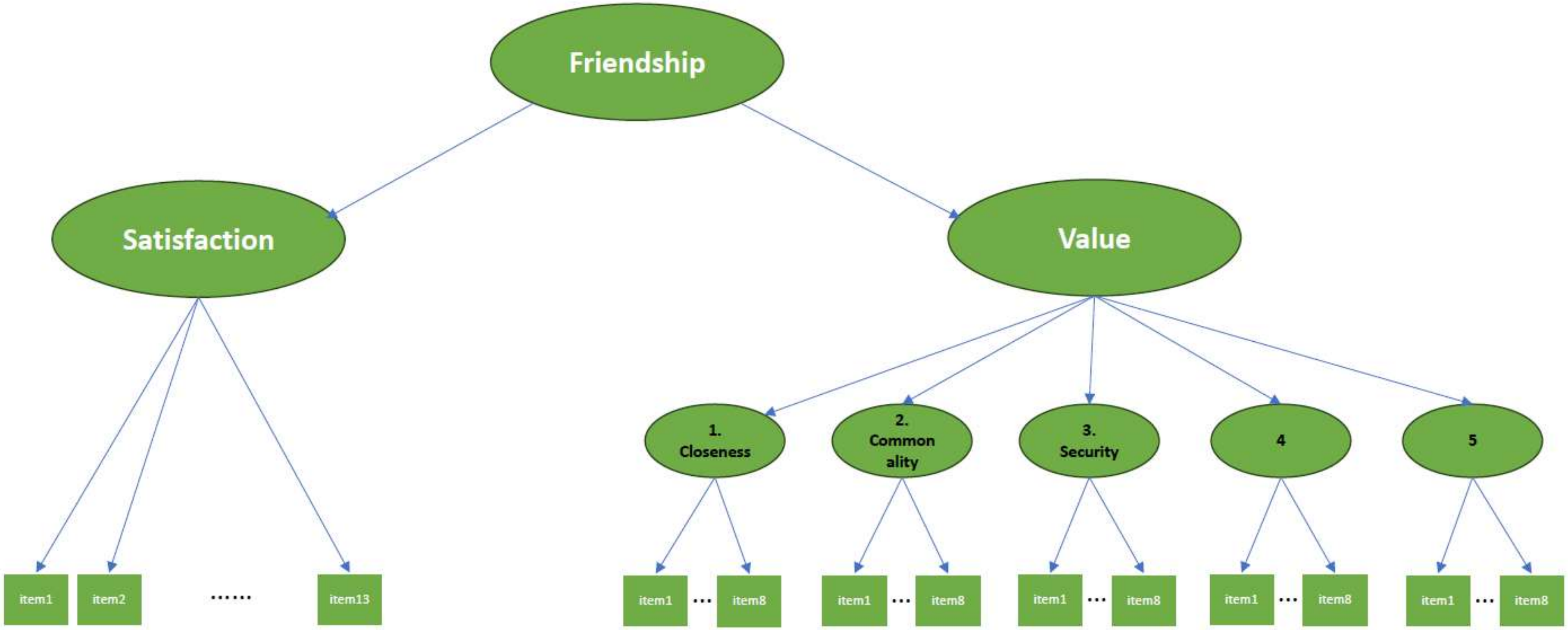
LATENT VARIABLE MODELS

6: More complex CFAs

Validity


Internal structure
Associations with other variables
Construct Validity
Response Processes
Consequences of Use
Test Content

Friendship questionnaire design



Example data exercise

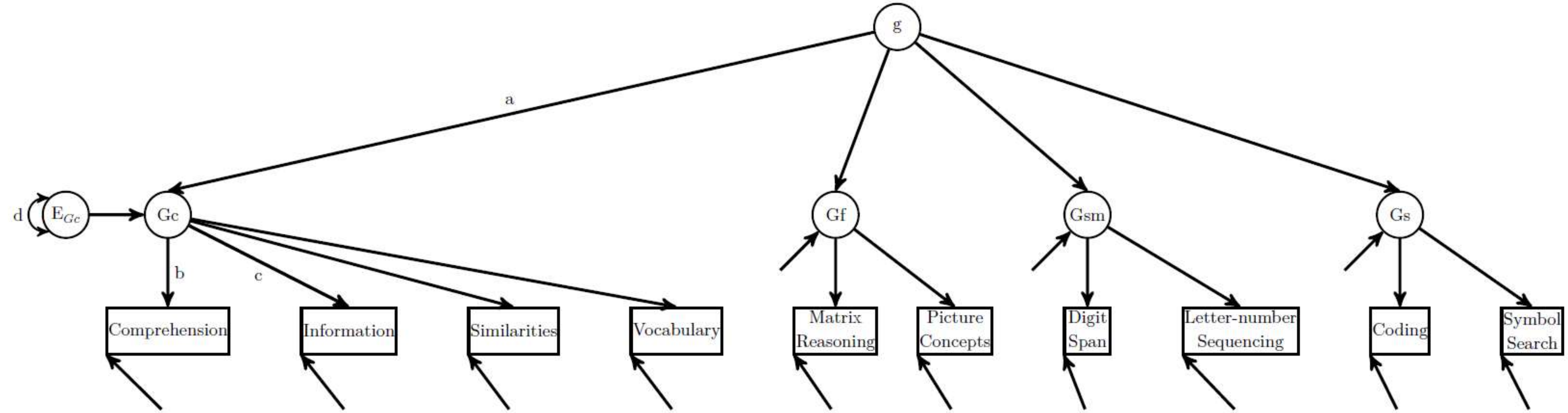
- Load data in R:

```
item_dat <- readRDS(file = "Item data.Rda")
```

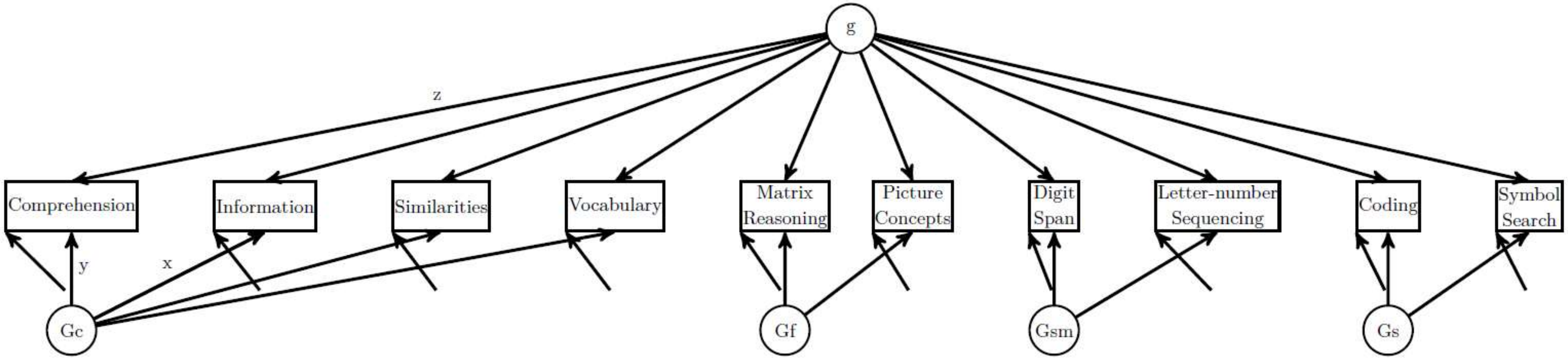
- Define the model in lavaan, fit it to the dataset, and evaluate

Hierarchical CFAs

- Second-order models (or even higher order)
- Bifactor models
- Rule of thumb: Bifactor always fits best
  - Second-order is a specific case of bifactor model, so more restrictive
  - Second-order model is more parsimonious (has less freely estimated parameters)

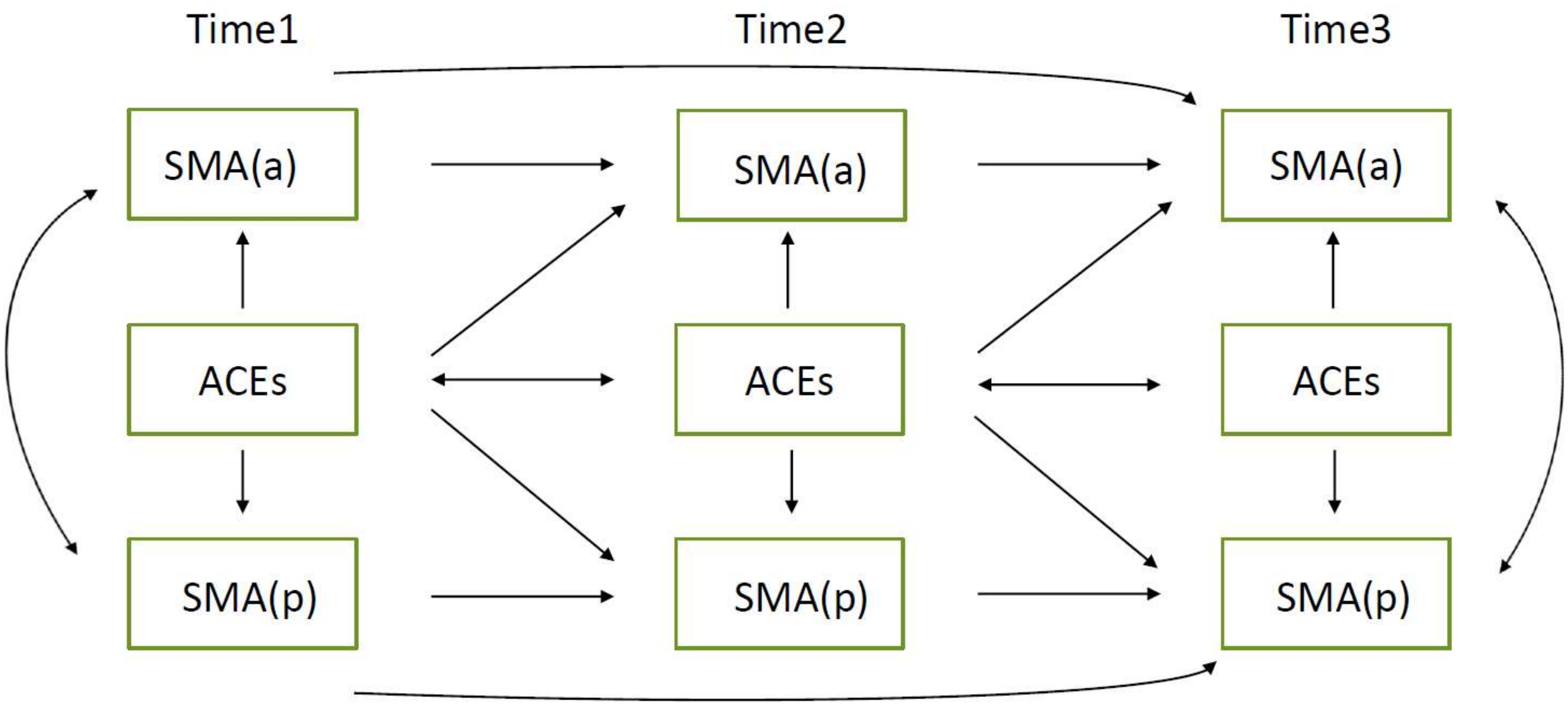


(a) Higher-order model.



(b) Bi-factor model.

Beaujean book, Figure 9.1 (see chapter 9)


Time1
Time2
Time3
SMA(a)
SMA(a)
SMA(a)
ACEs
ACEs
ACEs
SMA(p)
SMA(p)
SMA(p)

Remove items because loadings are too low?

- No!
  - CFA fit measures should not be the main determinant of validity! *You* define the construct.
  - Check item content when loadings are low:
    - Can improve the item?
    - Might this item simply have low variance, explaining the low loadings?
  - If you want to remove an item: Make sure content validity is not harmed. It is in the questionnaire for a reason.

Missing data?

- With (R)ML estimation, can use full-information maximum likelihood (FIML). Add to call:
  `missing = "fiml"`
- With DWLS(MV) estimation, can use pairwise-complete observations. Add to call:
  `missing = "pairwise"`
- Multiple imputation is the gold standard. Outside the scope of this basic course. In fact, FIML (and pairwise complete) often perform equally well
  - Schafer, J.L. and Graham, J.W. (2002). Missing Data: Our View of the State of the Art. *Psychological Methods, 7*(2), 147-177.

Missing data

- Missing Completely At Random (MCAR)
  - As name implies, missingness is completely random (i..e, not associated with any variables in the model.
  - Nothing systematic that makes some values more likely to be missing than others.
- Missing At Random (MAR)
  - There is a systematic relationship between the missingness and the *observed* data, but *not* the missing data.
  - Whether an observation is missing is independent from the missing value itself, but it may be dependent on other observed variables in the model
  - E.g., dataset with gender and weight. Gender has no missings, weight has missings
    - If women are more likely to have weight missing, missing is MAR
    - If people with higher weight are more likely to have missing weight, missing is MNAR
- Missing Not At Random (MNAR)
  - Missingness and missing values are dependent
  - Note: what is MNAR in one model may be MAR in another model
    - E.g., if women are more likely not to report their weight than men, and gender is not included in the model, missing MNAR. If gender is included in the model, missing is MAR
    - Thus, can make MNAR into MAR by adding variables to the model ('auxiliary variables' approach)

Missing data

- Missing Completely At Random (MCAR)
  - As name implies, missingness is completely random (with any variables in the model.
  - Nothing systematic that makes some v e missing than others.
- Missing At Random (MAR)
  - There is a systematic relationship betv and the *observed* data, but *not* the missing data.
  - Whether an observation is missing is independent from the missing value itself, but it may be dependent on other observed variables in the model
  - E.g., dataset with gender and weight. Gender has no missings, weight has missings
    - If women are more likely to have weight missing, missing is MAR
    - If people with higher weight are more likely ht, missing is MNAR
- Missing Not At Random (MNAR)
  - Missingness and missing values are dependent
  - Note: what is MNAR in one model may be MAR in another model
    - E.g., if women are more likely not to report their weight than men, and gender is not included in the model, missing MNAR. If gender is included in the model, missing is MAR
    - Thus, can make MNAR into MAR by adding variables to the model ('auxiliary variables' approach)

Unlikely in practice

Can be dealt with in analysis

Problematic

Missing data

- Listwise deletion
  - Best avoided
  - Only not problematic if data are MCAR and very little missingness
    - Although power is reduced, parameter estimates are unbiased
- Analysis based on pairwise complete observations
  - Unbiased parameter estimates when missing data are MAR
  - Can add 'auxiliary variable(s)' (i.e., variable(s) that contain information about missing values, but were not part of the original model) to turn MNAR into MAR
- Use full information maximum likelihood (FIML) estimation
  - Yields unbiased parameter estimates when missing data are MAR
  - Can add 'auxiliary variable(s)' (i.e., variable(s) that contain information about missing values, but were not part of the original model) to turn MNAR into MAR
- Multiple imputation (MI)
  - Outside scope of this course
  - Both FIML and MI perform similarly well (i.e., parameter estimates are unbiased) when data are MCAR or MAR (Schafer & Graham, 2002)