Multivariate Statistics and Methodology with R

Structural equation modeling

Aja Murray; Aja.Murray@ed.ac.uk

This week

- Techniques
  - *Full structural equation modeling (SEM)*
- Functions
  - *sem() from the lavaan package*
  - *omega() from the psych package*
- Reading
  - *http://lavaan.ugent.be/tutorial/tutorial.pdf (sections 5 and 6)*

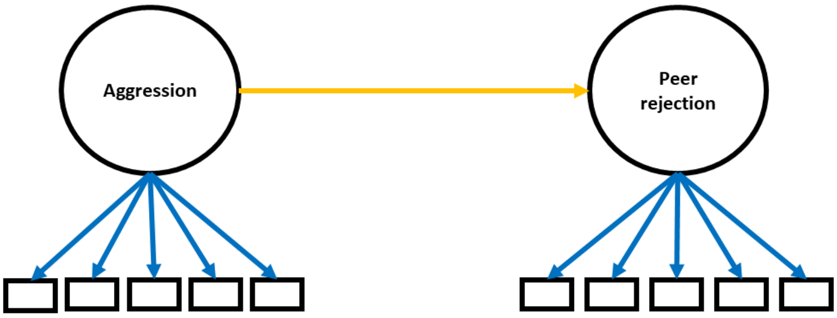
Learning outcomes

- Understand the potential benefits of using SEM over path analysis
- Estimate internal consistency values in R
- Specify, estimate, and interpret SEM models in R

SEM: bringing CFA and path analysis together

- We previously talked about how we can test latent variable models for constructs using CFA
  - *e.g., a two-factor model of aggression*
- We separately talked about how we can use path analysis to test sets of regression models
  - *e.g., a model to test whether peer rejection mediates the association between aggression and depression*
- SEM combines CFA and path analysis

A structural equation model



- SEM models regression paths between latent variables (the '**structural**' part of the model)
- The latent variables are from CFA measurement models (the '**measurement**' part of the model)

Why use a SEM model?

- Our measures of psychological constructs have imperfect **reliability**
- This means that scores have a degree of **measurement error** associated with them
- When we want to evaluate the relations between constructs, measurement error gets in they way
- Specifically, we are liable to underestimate the strength of relations between constructs when there is measurement error
- This is called **attenuation due to unreliability**
- Lower reliability measures lead to greater attenuation
- SEM, however, allows disattenuated estimates to be obtained

A brief detour into reliability

- Reliability theory suggests that:

$$ObservedScore = TrueScore + Error$$

- We can try to estimate how much of the variance in observed scores is due to error variance based on consistency of scores:
    - *across repeated administration of a measure over time (e.g., two weeks apart) (**test-retest reliability**)*
    - *across parallel forms of a test (**alternate forms reliability**)*
        - used to try and avoid practice effects
    - *across different raters (**inter-rater reliability**)*
        - e.g., teacher versus parent reports of aggression
    - *across items within a test (**internal consistency**)*

Internal consistency reliability

- Concerns correlations between items within same scale
- Logic is that if a measure is reliable, items within the measure should be correlated because they all reflect the construct well
- Traditional method was **split-half** reliability
  - *Divide test in two and correlate scores across the two halves*
  - *However, many possible ways to divide a test in two…*

Cronbach’s alpha

$$\alpha$$

- **Cronbach’s alpha** is a generalisation of split half-reliability
- Can be roughly interpreted as a measure of average correlation between all possible two-way splits of a measure
- Ranges from 0 to 1
- Values > .70 considered acceptable
- Most popular measure of reliability
- **However**, it assumes that all items are equally strongly correlated with underlying construct
  - *i.e., assumes equal factor loadings*
- Rarely true and is a big limitation of Cronbach’s alpha

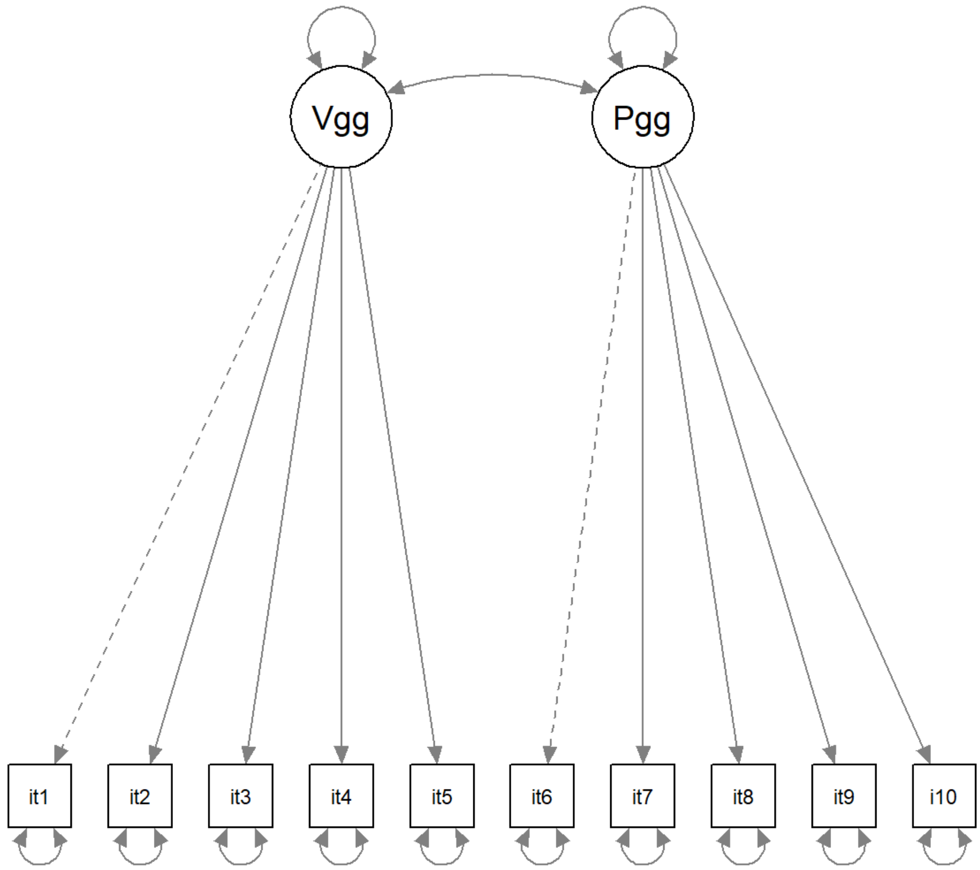
Omega

$$\omega$$

- **Omega** is an alternative measure of internal consistency reliability
- Based on the loadings from a factor analysis
- It is an estimate of the variance in the sum of all items (‘total score’) attributable to the latent factor(s)
- Ranges from 0 to 1
- Values > .70 considered acceptable
- Does not assume that the loadings are equal for all items
  - *This makes it a better measure of internal consistency than Cronbach’s alpha*
- **Use omega rather than Cronbach’s alpha to assess internal consistency**

Alpha and omega in R

- We could compute alpha and omega for our aggression data from the PCA, EFA and CFA lectures
- Recall:
  - *we had 10 aggression items*
  - *We determined using an EFA and then a CFA in new data that a model with two correlated factors was best*
  - *the two factors were labelled 'verbal aggression' and 'physical aggression'*



Alpha and omega for our aggression subscales

- We can use the omega() function from the psych package to compute internal consistency for each set of 5 items
- We let omega() know which items we wish to compute internal consistency for
  - *Here the first five items of our aggression measure*
- We let omega() know that these items consistute one factor by setting nfactors=1

```
library(psych)
```

```
##
## Attaching package: 'psych'
```

```
## The following object is masked from 'package:lavaan':
##
##     cor2cov
```

```
omega_verbal<-omega(agg.items[ ,c(1:5)], nfactors=1) ##omega for the verbal aggression factor (items 1-5)
```

```
## Loading required namespace: GPArotation
```

```
## Omega_h for 1 factor is not meaningful, just omega_t
```

omega() output

```
omega_verbal
```

```
## Omega
## Call: omegah(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip,
##     digits = digits, title = title, sl = sl, labels = labels,
##     plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option,
##     covar = covar)
## Alpha:                 0.87
## G.6:                   0.84
## Omega Hierarchical:    0.87
## Omega H asymptotic:    1
## Omega Total            0.87
##
## Schmid Leiman Factor loadings greater than  0.2
##          g  F1*   h2   u2 p2
## item1 0.72      0.51 0.49  1
## item2 0.80      0.64 0.36  1
## item3 0.73      0.53 0.47  1
## item4 0.69      0.47 0.53  1
## item5 0.85      0.72 0.28  1
##
## With eigenvalues of:
##   g F1*
## 2.9 0.0
##
## general/max  Inf   max/min =   NaN
## mean percent general =  1    with sd =  0 and cv of  0
## Explained Common Variance of the general factor =  1
##
## The degrees of freedom are 5  and the fit is  0
## The number of observations was  1000  with Chi Square =  1.17  with prob <  0.95
## The root mean square of the residuals is  0
## The df corrected root mean square of the residuals is  0.01
## RMSEA index =  0  and the 10 % confidence intervals are  0 0.005
## BIC =  -33.37
##
## Compare this with the adequacy of just a general factor and no group factors
## The degrees of freedom for just the general factor are 5  and the fit is  0
## The number of observations was  1000  with Chi Square =  1.17  with prob <  0.95
## The root mean square of the residuals is  0
## The df corrected root mean square of the residuals is  0.01
##
## RMSEA index =  0  and the 10 % confidence intervals are  0 0.005
## BIC =  -33.37
##
## Measures of factor score adequacy
##                                                  g F1*
## Correlation of scores with factors            0.94   0
## Multiple R square of scores with factors      0.88   0
## Minimum correlation of factor score estimates 0.76  -1
##
##  Total, General and Subset omega for each subset
##                                                  g  F1*
## Omega total for total scores and subscales    0.87 0.87
## Omega general for total scores and subscales  0.87 0.87
## Omega group for total scores and subscales    0.00 0.00
```

- ‘Alpha’ gives us our Cronbach’s alpha value
- ‘omega Total’ gives us our omega value

```r
library(psych)
omega(agg.items[ ,c(1:5)], nfactors=1) ## calculate alpha and omega for the verbal aggression factor
```

```
## Omega_h for 1 factor is not meaningful, just omega_t
```

```
## Omega 
## Call: omegah(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip, 
##     digits = digits, title = title, sl = sl, labels = labels, 
##     plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option, 
##     covar = covar)
## Alpha:                 0.87 
## G.6:                   0.84 
## Omega Hierarchical:    0.87 
## Omega H asymptotic:    1 
## Omega Total            0.87 
## 
## Schmid Leiman Factor loadings greater than  0.2 
##          g  F1*   h2   u2 p2
## item1 0.72      0.51 0.49  1
## item2 0.80      0.64 0.36  1
## item3 0.73      0.53 0.47  1
## item4 0.69      0.47 0.53  1
## item5 0.85      0.72 0.28  1
## 
## With eigenvalues of:
##   g F1* 
## 2.9 0.0 
## 
## general/max  Inf   max/min =   NaN
## mean percent general =  1    with sd =  0 and cv of  0 
## Explained Common Variance of the general factor =  1 
## 
## The degrees of freedom are 5  and the fit is  0 
## The number of observations was  1000  with Chi Square =  1.17  with prob <  0.95
## The root mean square of the residuals is  0 
## The df corrected root mean square of the residuals is  0.01
## RMSEA index =  0  and the 10 % confidence intervals are  0 0.005
## BIC =  -33.37
## 
## Compare this with the adequacy of just a general factor and no group factors
## The degrees of freedom for just the general factor are 5  and the fit is  0 
## The number of observations was  1000  with Chi Square =  1.17  with prob <  0.95
## The root mean square of the residuals is  0 
## The df corrected root mean square of the residuals is  0.01 
## 
## RMSEA index =  0  and the 10 % confidence intervals are  0 0.005
## BIC =  -33.37 
## 
## Measures of factor score adequacy             
##                                                  g F1*
## Correlation of scores with factors            0.94   0
## Multiple R square of scores with factors      0.88   0
## Minimum correlation of factor score estimates 0.76  -1
## 
##  Total, General and Subset omega for each subset
##                                                  g  F1*
## Omega total for total scores and subscales    0.87 0.87
## Omega general for total scores and subscales  0.87 0.87
## Omega group for total scores and subscales    0.00 0.00
```

Alpha and omega

- We can do the same for the physical aggression items

```
omega_physical<-omega(agg.items[ ,c(6:10)], nfactors=1) ## calculate alpha and omega for the physical aggression
  factor
```

```
## Omega_h for 1 factor is not meaningful, just omega_t
```

omega() output for physical aggression

```
omega_physical
```

```
## Omega
## Call: omegah(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip,
##     digits = digits, title = title, sl = sl, labels = labels,
##     plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option,
##     covar = covar)
## Alpha:                 0.89
## G.6:                   0.87
## Omega Hierarchical:    0.89
## Omega H asymptotic:    1
## Omega Total            0.89
##
## Schmid Leiman Factor loadings greater than  0.2
##           g  F1*   h2   u2 p2
## item6  0.66      0.44 0.56  1
## item7  0.90      0.81 0.19  1
## item8  0.92      0.85 0.15  1
## item9  0.70      0.49 0.51  1
## item10 0.74      0.55 0.45  1
##
## With eigenvalues of:
##   g F1*
## 3.1 0.0
##
## general/max  Inf   max/min =   NaN
## mean percent general =  1    with sd =  0 and cv of  0
## Explained Common Variance of the general factor =  1
##
## The degrees of freedom are 5  and the fit is  0.01
## The number of observations was  1000  with Chi Square =  6.81  with prob <  0.24
## The root mean square of the residuals is  0.01
## The df corrected root mean square of the residuals is  0.01
## RMSEA index =  0.019  and the 10 % confidence intervals are  0 0.051
## BIC =  -27.73
##
## Compare this with the adequacy of just a general factor and no group factors
## The degrees of freedom for just the general factor are 5  and the fit is  0.01
## The number of observations was  1000  with Chi Square =  6.81  with prob <  0.24
## The root mean square of the residuals is  0.01
## The df corrected root mean square of the residuals is  0.01
##
## RMSEA index =  0.019  and the 10 % confidence intervals are  0 0.051
## BIC =  -27.73
##
## Measures of factor score adequacy
##                                                  g F1*
## Correlation of scores with factors            0.96   0
## Multiple R square of scores with factors      0.93   0
## Minimum correlation of factor score estimates 0.86  -1
##
##  Total, General and Subset omega for each subset
##                                                  g  F1*
## Omega total for total scores and subscales    0.89 0.89
## Omega general for total scores and subscales  0.89 0.89
## Omega group for total scores and subscales    0.00 0.00
```

How to solve the problem of attenuation due to unreliability?

- Traditional method was to apply a formula to correct correlations for unreliability:

$$\frac{r_{xy}}{r_{xx} * r_{yy}}$$

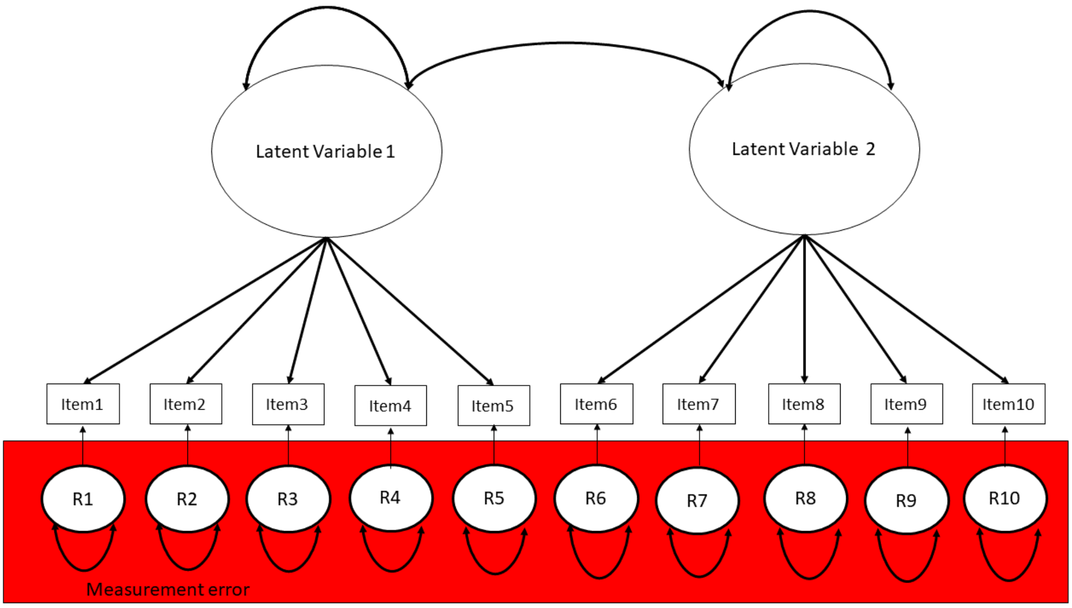
- Where: - r_{xy} is the uncorrected correlation between variables x and y - r_{xx} is an estimate of the reliability of variable x - r_{yy} is an estimate of the reliability of variable y

- However, this requires multiple steps (compute reliability, correct correlations)

- Further complicated when it’s a whole correlation matrix that requires correction

- SEM can solve the problem in a single step

Addressing attenuation due to unreliability with SEM



- SEM can solve the problem of attenuation due to unreliability
- It uses latent variable measurement models from CFA
- These models separate out systematic variance (latent common factors) and measurement error variance (residual factors)
- The relations between constructs are tested using the latent common factors i.e., the error-free parts

Fitting structural equation models

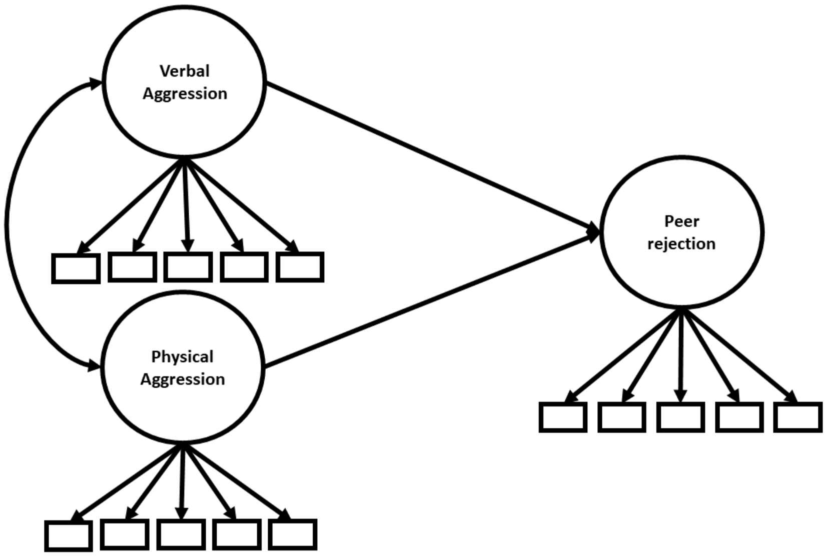
- Fitting SEMs follows the same process as CFA and path analyis:
  - *Model specification*
  - *Model estimation*
  - *Model evaluation*
  - *(Model modification)*
  - *Model interpretation*
- However, we usually want to test our measurement models first using CFAs for each construct

An example SEM model

- Imagine we wanted to know whether verbal and physical aggression predicted peer rejection in children, accounting for imperfect reliability
- We have a sample of n=570
- We have a 10-item aggression measure to measure verbal and physical aggression (5 items each)
- We have a 5-item peer rejection measure
- We can fit a SEM to assess whether latent verbal and physical aggression factors predict a latent peer rejection factor

Our model

- The model we want to test looks like:



Step 1: check the measurement models

- First we would conduct a CFA for aggression and a CFA for peer rejection
  - *i.e., we first test our proposed measurement models*
- We do this as a first step because mis-fit is most often due to measurement rather than structural part of the model

CFA for aggression

- We fit a two-factor CFA with correlated factors for aggression
- By default, the first loading for each factor will be fixed to 1 for scaling/identification

```
##CFA for aggression

agg.CFA<-'Vagg=~agg1+agg2+agg3+agg4+agg5

   Pagg=~agg6+agg7+agg8+agg9+agg10

   Vagg~~Pagg'

agg.CFA.est<-cfa(agg.CFA, data=agg.PR.data)
summary(agg.CFA.est, fit.measures=T, standardized=T)
```

```
## lavaan 0.6-5 ended normally after 30 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         21
##
##   Number of observations                           570
##
## Model Test User Model:
##
##   Test statistic                                46.611
##   Degrees of freedom                                34
##   P-value (Chi-square)                           0.073
##
## Model Test Baseline Model:
##
##   Test statistic                              3343.183
##   Degrees of freedom                                45
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.996
##   Tucker-Lewis Index (TLI)                       0.995
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -6599.772
##   Loglikelihood unrestricted model (H1)      -6576.466
##
##   Akaike (AIC)                               13241.543
##   Bayesian (BIC)                             13332.802
##   Sample-size adjusted Bayesian (BIC)        13266.136
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.026
##   90 Percent confidence interval - lower         0.000
##   90 Percent confidence interval - upper         0.042
##   P-value RMSEA <= 0.05                          0.994
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.021
```

```
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   Vagg =~
##     agg1              1.000                               0.756    0.737
##     agg2              1.145    0.059   19.279    0.000    0.866    0.830
##     agg3              0.910    0.058   15.721    0.000    0.688    0.680
##     agg4              0.913    0.058   15.620    0.000    0.691    0.676
##     agg5              1.139    0.058   19.593    0.000    0.862    0.844
##   Pagg =~
##     agg6              1.000                               0.691    0.678
##     agg7              1.395    0.073   19.020    0.000    0.964    0.898
##     agg8              1.331    0.070   18.995    0.000    0.919    0.897
##     agg9              1.155    0.071   16.263    0.000    0.798    0.747
##     agg10             1.022    0.065   15.766    0.000    0.706    0.722
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   Vagg ~~
##     Pagg              0.386    0.037   10.393    0.000    0.739    0.739
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .agg1              0.480    0.033   14.488    0.000    0.480    0.456
##    .agg2              0.340    0.027   12.380    0.000    0.340    0.312
##    .agg3              0.550    0.036   15.164    0.000    0.550    0.538
##    .agg4              0.567    0.037   15.204    0.000    0.567    0.543
##    .agg5              0.300    0.025   11.843    0.000    0.300    0.288
##    .agg6              0.560    0.036   15.707    0.000    0.560    0.540
##    .agg7              0.222    0.021   10.816    0.000    0.222    0.193
##    .agg8              0.206    0.019   10.917    0.000    0.206    0.196
##    .agg9              0.504    0.033   15.123    0.000    0.504    0.442
##    .agg10             0.458    0.030   15.374    0.000    0.458    0.479
##     Vagg              0.572    0.058    9.862    0.000    1.000    1.000
##     Pagg              0.477    0.053    8.943    0.000    1.000    1.000
```

CFA for peer rejection

- We fit a CFA for peer rejection
- By default, the loading for the first item will be fixed to 1 for scaling/identification

```r
##CFA for aggression

PR.CFA<-'PR=~PR1+PR2+PR3+PR4+PR5'


PR.CFA.est<-cfa(PR.CFA, data=agg.PR.data)
summary(PR.CFA.est, fit.measures=T, standardized=T)
```

```
## lavaan 0.6-5 ended normally after 24 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         10
##
##   Number of observations                           570
##
## Model Test User Model:
##
##   Test statistic                                 7.228
##   Degrees of freedom                                 5
##   P-value (Chi-square)                           0.204
##
## Model Test Baseline Model:
##
##   Test statistic                              1978.376
##   Degrees of freedom                                10
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.999
##   Tucker-Lewis Index (TLI)                       0.998
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -3097.913
##   Loglikelihood unrestricted model (H1)      -3094.298
##
##   Akaike (AIC)                                6215.825
##   Bayesian (BIC)                              6259.281
##   Sample-size adjusted Bayesian (BIC)         6227.536
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.028
##   90 Percent confidence interval - lower         0.000
##   90 Percent confidence interval - upper         0.069
##   P-value RMSEA <= 0.05                          0.771
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.009
##
## Parameter Estimates:
##
```

```
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   PR =~
##     PR1               1.000                               0.769    0.776
##     PR2               1.047    0.051   20.476    0.000    0.805    0.798
##     PR3               1.211    0.050   24.145    0.000    0.931    0.914
##     PR4               1.176    0.051   22.847    0.000    0.904    0.871
##     PR5               1.031    0.052   19.959    0.000    0.792    0.781
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .PR1               0.391    0.026   14.868    0.000    0.391    0.398
##    .PR2               0.370    0.025   14.536    0.000    0.370    0.364
##    .PR3               0.171    0.017   10.020    0.000    0.171    0.165
##    .PR4               0.259    0.021   12.521    0.000    0.259    0.241
##    .PR5               0.401    0.027   14.790    0.000    0.401    0.390
##     PR                0.591    0.055   10.783    0.000    1.000    1.000
```

Step 2: specify the SEM model

- Assuming the measurement models show good fit, we proceed to specifying the full SEM
- The SEM specification combines the measurement models with the hypothesised structural relations between the latent variables
- Just as with CFA and path analysis, model must be identified
  - *The number of 'knowns' are at least as many as the 'unknowns'*

```
agg.PR.model<-'
# aggression measurement model
   Vagg=~agg1+agg2+agg3+agg4+agg5

   Pagg=~agg6+agg7+agg8+agg9+agg10

   Vagg~~Pagg

# peer rejection measurement model
   PR=~PR1+PR2+PR3+PR4+PR5

#structural part of the model

  PR~Vagg + Pagg        # Peer rejection is regressed on verbal and physical aggression'
```

Step 3: estimate the SEM model

- As for CFA and path analysis, we can use maximum likelihood estimation to estimate the parameters
- As for path analysis we can do this using the sem() function from lavaan
- We provide the name of the model and the dataset

```
agg.PR.est<-sem(agg.PR.model, data= agg.PR.data)
```

Step 4: evaluate the model

- We look at the fit statistics and check they are satisfactory
- We are looking for TLI and CFI>.95; RMSEA and SRMR<.05
- We can inspect the fit statistics using the summary() function, setting fit.measures=T

```
summary(agg.PR.est, fit.measures=T)
```

```
## lavaan 0.6-5 ended normally after 35 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         33
##
##   Number of observations                           570
##
## Model Test User Model:
##
##   Test statistic                               107.429
##   Degrees of freedom                                87
##   P-value (Chi-square)                           0.068
##
## Model Test Baseline Model:
##
##   Test statistic                              5508.400
##   Degrees of freedom                               105
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.996
##   Tucker-Lewis Index (TLI)                       0.995
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -9631.059
##   Loglikelihood unrestricted model (H1)      -9577.344
##
##   Akaike (AIC)                               19328.118
##   Bayesian (BIC)                             19471.524
##   Sample-size adjusted Bayesian (BIC)        19366.763
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.020
##   90 Percent confidence interval - lower         0.000
##   90 Percent confidence interval - upper         0.032
##   P-value RMSEA <= 0.05                          1.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.023
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard errors                             Standard
##
```

```
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Vagg =~
##     agg1              1.000
##     agg2              1.146    0.059   19.313    0.000
##     agg3              0.909    0.058   15.731    0.000
##     agg4              0.914    0.058   15.659    0.000
##     agg5              1.138    0.058   19.605    0.000
##   Pagg =~
##     agg6              1.000
##     agg7              1.397    0.074   18.982    0.000
##     agg8              1.332    0.070   18.948    0.000
##     agg9              1.161    0.071   16.281    0.000
##     agg10             1.025    0.065   15.762    0.000
##   PR =~
##     PR1               1.000
##     PR2               1.048    0.051   20.549    0.000
##     PR3               1.212    0.050   24.258    0.000
##     PR4               1.172    0.051   22.823    0.000
##     PR5               1.031    0.052   20.012    0.000
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   PR ~
##     Vagg              0.241    0.070    3.442    0.001
##     Pagg              0.319    0.077    4.156    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Vagg ~~
##     Pagg              0.386    0.037   10.390    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .agg1              0.480    0.033   14.509    0.000
##    .agg2              0.339    0.027   12.418    0.000
##    .agg3              0.551    0.036   15.183    0.000
##    .agg4              0.566    0.037   15.211    0.000
##    .agg5              0.301    0.025   11.929    0.000
##    .agg6              0.562    0.036   15.723    0.000
##    .agg7              0.223    0.020   10.898    0.000
##    .agg8              0.208    0.019   11.035    0.000
##    .agg9              0.500    0.033   15.115    0.000
##    .agg10             0.457    0.030   15.379    0.000
##    .PR1               0.391    0.026   14.913    0.000
##    .PR2               0.368    0.025   14.578    0.000
##    .PR3               0.169    0.017   10.115    0.000
##    .PR4               0.264    0.021   12.745    0.000
##    .PR5               0.400    0.027   14.835    0.000
##     Vagg              0.572    0.058    9.870    0.000
##     Pagg              0.476    0.053    8.927    0.000
##    .PR                0.450    0.043   10.596    0.000
```

Step 5: interpret the model

- We can see whether the regression paths are significant using the summary() function
- We can also look at the standardised coefficients by setting standardized=T

```
summary(agg.PR.est, fit.measures=T, standardized=T)
```

```
## lavaan 0.6-5 ended normally after 35 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         33
##
##   Number of observations                           570
##
## Model Test User Model:
##
##   Test statistic                               107.429
##   Degrees of freedom                                87
##   P-value (Chi-square)                           0.068
##
## Model Test Baseline Model:
##
##   Test statistic                              5508.400
##   Degrees of freedom                               105
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.996
##   Tucker-Lewis Index (TLI)                       0.995
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -9631.059
##   Loglikelihood unrestricted model (H1)      -9577.344
##
##   Akaike (AIC)                               19328.118
##   Bayesian (BIC)                             19471.524
##   Sample-size adjusted Bayesian (BIC)        19366.763
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.020
##   90 Percent confidence interval - lower         0.000
##   90 Percent confidence interval - upper         0.032
##   P-value RMSEA <= 0.05                          1.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.023
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard errors                             Standard
##
## Latent Variables:
```

```
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   Vagg =~
##     agg1             1.000                               0.757    0.738
##     agg2             1.146    0.059   19.313    0.000    0.867    0.830
##     agg3             0.909    0.058   15.731    0.000    0.688    0.680
##     agg4             0.914    0.058   15.659    0.000    0.692    0.677
##     agg5             1.138    0.058   19.605    0.000    0.861    0.843
##   Pagg =~
##     agg6             1.000                               0.690    0.677
##     agg7             1.397    0.074   18.982    0.000    0.963    0.898
##     agg8             1.332    0.070   18.948    0.000    0.918    0.896
##     agg9             1.161    0.071   16.281    0.000    0.800    0.749
##     agg10            1.025    0.065   15.762    0.000    0.707    0.723
##   PR =~
##     PR1              1.000                               0.769    0.776
##     PR2              1.048    0.051   20.549    0.000    0.806    0.799
##     PR3              1.212    0.050   24.258    0.000    0.932    0.915
##     PR4              1.172    0.051   22.823    0.000    0.902    0.869
##     PR5              1.031    0.052   20.012    0.000    0.793    0.782
##
## Regressions:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   PR ~
##     Vagg             0.241    0.070    3.442    0.001    0.237    0.237
##     Pagg             0.319    0.077    4.156    0.000    0.286    0.286
##
## Covariances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   Vagg ~~
##     Pagg             0.386    0.037   10.390    0.000    0.739    0.739
##
## Variances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .agg1             0.480    0.033   14.509    0.000    0.480    0.456
##    .agg2             0.339    0.027   12.418    0.000    0.339    0.311
##    .agg3             0.551    0.036   15.183    0.000    0.551    0.538
##    .agg4             0.566    0.037   15.211    0.000    0.566    0.542
##    .agg5             0.301    0.025   11.929    0.000    0.301    0.289
##    .agg6             0.562    0.036   15.723    0.000    0.562    0.542
##    .agg7             0.223    0.020   10.898    0.000    0.223    0.194
##    .agg8             0.208    0.019   11.035    0.000    0.208    0.197
##    .agg9             0.500    0.033   15.115    0.000    0.500    0.439
##    .agg10            0.457    0.030   15.379    0.000    0.457    0.478
##    .PR1              0.391    0.026   14.913    0.000    0.391    0.398
##    .PR2              0.368    0.025   14.578    0.000    0.368    0.362
##    .PR3              0.169    0.017   10.115    0.000    0.169    0.163
##    .PR4              0.264    0.021   12.745    0.000    0.264    0.245
##    .PR5              0.400    0.027   14.835    0.000    0.400    0.389
##     Vagg             0.572    0.058    9.870    0.000    1.000    1.000
##     Pagg             0.476    0.053    8.927    0.000    1.000    1.000
##    .PR               0.450    0.043   10.596    0.000    0.762    0.762
```

Making model modifications in SEM

- Our initially hypothesised model may not be optimal
  - *We didn’t include paths that we should have (check expected parameter changes and modification indices)*
  - *Some included paths are non-significant and could be trimmed*
- These issues may affect the measurement or structural part of the model
  - *But more often mis-fit relates to the measurement part*
  - *Aim to make any modifications in the measurement part in initial CFAs before fitting the full SEM*
- Carefully consider before making modifications
  - *Can they be theoretically justified?*
  - *Am I likely to be just capitalising on chance?*
- Aim to replicate the modified model in new data

Check modification indices and expected parameter changes

```
modindices(agg.PR.est, sort=T)
```

```
##        lhs op   rhs     mi    epc sepc.lv sepc.all sepc.nox
## 68    agg1 ~~  agg3 10.595 -0.082  -0.082   -0.160   -0.160
## 95    agg3 ~~  agg5  8.947  0.069   0.069    0.170    0.170
## 137   agg7 ~~  agg9  7.264 -0.056  -0.056   -0.167   -0.167
## 45    Vagg =~   PR4  6.615 -0.102  -0.077   -0.075   -0.075
## 102   agg3 ~~   PR2  5.852  0.051   0.051    0.113    0.113
## 65      PR =~  agg9  5.242  0.112   0.086    0.081    0.081
## 131   agg6 ~~   PR1  4.855  0.047   0.047    0.101    0.101
## 123   agg5 ~~   PR2  4.643 -0.037  -0.037   -0.112   -0.112
## 119   agg5 ~~  agg8  4.337 -0.031  -0.031   -0.122   -0.122
## 169    PR3 ~~   PR4  4.283  0.038   0.038    0.180    0.180
## 106   agg4 ~~  agg5  4.227 -0.048  -0.048   -0.116   -0.116
## 167    PR2 ~~   PR4  4.102 -0.036  -0.036   -0.117   -0.117
## 104   agg3 ~~   PR4  4.019 -0.038  -0.038   -0.099   -0.099
## 51    Pagg =~  agg5  3.541 -0.142  -0.098   -0.096   -0.096
## 50    Pagg =~  agg4  3.537  0.161   0.111    0.108    0.108
## 160 agg10 ~~   PR4  3.301  0.031   0.031    0.090    0.090
## 168    PR2 ~~   PR5  3.158  0.034   0.034    0.089    0.089
## 117   agg5 ~~  agg6  3.095 -0.036  -0.036   -0.088   -0.088
## 170    PR3 ~~   PR5  3.041 -0.031  -0.031   -0.117   -0.117
## 72    agg1 ~~  agg7  2.948 -0.030  -0.030   -0.093   -0.093
## 74    agg1 ~~  agg9  2.910  0.039   0.039    0.080    0.080
## 140   agg7 ~~   PR2  2.806  0.026   0.026    0.090    0.090
## 118   agg5 ~~  agg7  2.711  0.025   0.025    0.097    0.097
## 154   agg9 ~~   PR3  2.622  0.026   0.026    0.089    0.089
## 141   agg7 ~~   PR3  2.592 -0.020  -0.020   -0.101   -0.101
## 39    Vagg =~  agg8  2.564 -0.095  -0.072   -0.070   -0.070
## 92    agg2 ~~   PR4  2.508 -0.026  -0.026   -0.085   -0.085
## 88    agg2 ~~ agg10  2.378 -0.030  -0.030   -0.077   -0.077
## 124   agg5 ~~   PR3  2.238  0.021   0.021    0.091    0.091
## 44    Vagg =~   PR3  2.209  0.053   0.040    0.040    0.040
## 47    Pagg =~  agg1  2.199  0.121   0.084    0.081    0.081
## 144   agg8 ~~  agg9  2.193  0.029   0.029    0.091    0.091
## 93    agg2 ~~   PR5  1.966  0.026   0.026    0.071    0.071
## 62      PR =~  agg6  1.832 -0.069  -0.053   -0.052   -0.052
## 146   agg8 ~~   PR1  1.649 -0.019  -0.019   -0.068   -0.068
## 40    Vagg =~  agg9  1.613  0.094   0.071    0.067    0.067
## 38    Vagg =~  agg7  1.415  0.074   0.056    0.052    0.052
## 153   agg9 ~~   PR2  1.410 -0.024  -0.024   -0.056   -0.056
## 64      PR =~  agg8  1.307 -0.043  -0.033   -0.032   -0.032
## 156   agg9 ~~   PR5  1.255  0.023   0.023    0.052    0.052
## 37    Vagg =~  agg6  1.218 -0.084  -0.064   -0.063   -0.063
## 99    agg3 ~~  agg9  1.215 -0.027  -0.027   -0.051   -0.051
## 138   agg7 ~~ agg10  1.164  0.021   0.021    0.065    0.065
## 60      PR =~  agg4  1.132  0.055   0.042    0.042    0.042
## 107   agg4 ~~  agg6  1.119  0.027   0.027    0.048    0.048
## 125   agg5 ~~   PR4  1.091  0.016   0.016    0.057    0.057
## 163    PR1 ~~   PR3  1.054 -0.018  -0.018   -0.068   -0.068
## 75    agg1 ~~ agg10  1.050  0.022   0.022    0.048    0.048
## 122   agg5 ~~   PR1  1.017 -0.018  -0.018   -0.052   -0.052
## 97    agg3 ~~  agg7  0.976  0.018   0.018    0.052    0.052
## 132   agg6 ~~   PR2  0.925 -0.020  -0.020   -0.044   -0.044
## 48    Pagg =~  agg2  0.906 -0.074  -0.051   -0.049   -0.049
## 61      PR =~  agg5  0.898 -0.042  -0.032   -0.031   -0.031
## 126   agg5 ~~   PR5  0.865 -0.017  -0.017   -0.048   -0.048
## 101   agg3 ~~   PR1  0.855  0.020   0.020    0.043    0.043
## 55    Pagg =~   PR4  0.844 -0.040  -0.027   -0.026   -0.026
## 91    agg2 ~~   PR3  0.770  0.013   0.013    0.052    0.052
## 53    Pagg =~   PR2  0.687  0.039   0.027    0.027    0.027
```

```
## 66      PR =~ agg10  0.651  0.037   0.029    0.029    0.029
## 113   agg4 ~~   PR2  0.650  0.017   0.017    0.038    0.038
## 70    agg1 ~~  agg5  0.615  0.018   0.018    0.047    0.047
## 133   agg6 ~~   PR3  0.593 -0.013  -0.013   -0.042   -0.042
## 71    agg1 ~~  agg6  0.588  0.018   0.018    0.035    0.035
## 105   agg3 ~~   PR5  0.569 -0.016  -0.016   -0.035   -0.035
## 115   agg4 ~~   PR4  0.536 -0.014  -0.014   -0.036   -0.036
## 134   agg6 ~~   PR4  0.515 -0.013  -0.013   -0.035   -0.035
## 130   agg6 ~~ agg10  0.505 -0.017  -0.017   -0.033   -0.033
## 159 agg10 ~~   PR3  0.476 -0.010  -0.010   -0.038   -0.038
## 73    agg1 ~~  agg8  0.457  0.011   0.011    0.036    0.036
## 82    agg2 ~~  agg4  0.417  0.016   0.016    0.036    0.036
## 145   agg8 ~~ agg10  0.405 -0.012  -0.012   -0.038   -0.038
## 158 agg10 ~~   PR2  0.399 -0.012  -0.012   -0.029   -0.029
## 76    agg1 ~~   PR1  0.394  0.013   0.013    0.030    0.030
## 164    PR1 ~~   PR4  0.384  0.011   0.011    0.035    0.035
## 152   agg9 ~~   PR1  0.375 -0.013  -0.013   -0.029   -0.029
## 147   agg8 ~~   PR2  0.362  0.009   0.009    0.032    0.032
## 128   agg6 ~~  agg8  0.344  0.011   0.011    0.034    0.034
## 81    agg2 ~~  agg3  0.315 -0.013  -0.013   -0.031   -0.031
## 150   agg8 ~~   PR5  0.297 -0.008  -0.008   -0.029   -0.029
## 121   agg5 ~~ agg10  0.297  0.010   0.010    0.028    0.028
## 112   agg4 ~~   PR1  0.273  0.011   0.011    0.024    0.024
## 129   agg6 ~~  agg9  0.267  0.013   0.013    0.024    0.024
## 139   agg7 ~~   PR1  0.258  0.008   0.008    0.027    0.027
## 79    agg1 ~~   PR4  0.244 -0.009  -0.009   -0.025   -0.025
## 151   agg9 ~~ agg10  0.238 -0.011  -0.011   -0.023   -0.023
## 109   agg4 ~~  agg8  0.235  0.009   0.009    0.026    0.026
## 67    agg1 ~~  agg2  0.220  0.011   0.011    0.027    0.027
## 110   agg4 ~~  agg9  0.175  0.010   0.010    0.019    0.019
## 127   agg6 ~~  agg7  0.172  0.008   0.008    0.024    0.024
## 46    Vagg =~   PR5  0.165  0.018   0.014    0.014    0.014
## 43    Vagg =~   PR2  0.156  0.017   0.013    0.013    0.013
## 49    Pagg =~  agg3  0.141  0.032   0.022    0.022    0.022
## 57      PR =~  agg1  0.126  0.017   0.013    0.013    0.013
## 116   agg4 ~~   PR5  0.118  0.008   0.008    0.016    0.016
## 165    PR1 ~~   PR5  0.117  0.007   0.007    0.017    0.017
## 56    Pagg =~   PR5  0.116  0.017   0.012    0.011    0.011
## 52    Pagg =~   PR1  0.111  0.016   0.011    0.011    0.011
## 42    Vagg =~   PR1  0.101  0.014   0.011    0.011    0.011
## 136   agg7 ~~  agg8  0.090  0.006   0.006    0.029    0.029
## 89    agg2 ~~   PR1  0.088 -0.005  -0.005   -0.015   -0.015
## 83    agg2 ~~  agg5  0.078  0.007   0.007    0.021    0.021
## 41    Vagg =~ agg10  0.063  0.018   0.013    0.014    0.014
## 157 agg10 ~~   PR1  0.056  0.005   0.005    0.011    0.011
## 54    Pagg =~   PR3  0.055 -0.009  -0.006   -0.006   -0.006
## 90    agg2 ~~   PR2  0.054 -0.004  -0.004   -0.012   -0.012
## 96    agg3 ~~  agg6  0.051 -0.006  -0.006   -0.010   -0.010
## 166    PR2 ~~   PR3  0.045  0.004   0.004    0.015    0.015
## 149   agg8 ~~   PR4  0.045  0.003   0.003    0.012    0.012
## 100   agg3 ~~ agg10  0.044  0.005   0.005    0.010    0.010
## 120   agg5 ~~  agg9  0.040 -0.004  -0.004   -0.010   -0.010
## 161 agg10 ~~   PR5  0.037 -0.004  -0.004   -0.009   -0.009
## 85    agg2 ~~  agg7  0.032  0.003   0.003    0.010    0.010
## 111   agg4 ~~ agg10  0.031  0.004   0.004    0.008    0.008
## 108   agg4 ~~  agg7  0.027 -0.003  -0.003   -0.009   -0.009
## 84    agg2 ~~  agg6  0.023 -0.003  -0.003   -0.007   -0.007
## 63      PR =~  agg7  0.018 -0.005  -0.004   -0.004   -0.004
## 94    agg3 ~~  agg4  0.016 -0.003  -0.003   -0.006   -0.006
## 142   agg7 ~~   PR4  0.015 -0.002  -0.002   -0.007   -0.007
## 155   agg9 ~~   PR4  0.015  0.002   0.002    0.006    0.006
## 87    agg2 ~~  agg9  0.014  0.002   0.002    0.006    0.006
## 135   agg6 ~~   PR5  0.014 -0.003  -0.003   -0.005   -0.005
## 80    agg1 ~~   PR5  0.013  0.002   0.002    0.005    0.005
## 103   agg3 ~~   PR3  0.011 -0.002  -0.002   -0.006   -0.006
## 162    PR1 ~~   PR2  0.009  0.002   0.002    0.005    0.005
## 143   agg7 ~~   PR5  0.008  0.001   0.001    0.005    0.005
## 86    agg2 ~~  agg8  0.005  0.001   0.001    0.004    0.004
## 148   agg8 ~~   PR3  0.005  0.001   0.001    0.004    0.004
```

```
## 114   agg4 ~~   PR3  0.004 -0.001  -0.001   -0.004   -0.004
## 69    agg1 ~~  agg4  0.003 -0.001  -0.001   -0.003   -0.003
## 171    PR4 ~~   PR5  0.003 -0.001  -0.001   -0.003   -0.003
## 77    agg1 ~~   PR2  0.002  0.001   0.001    0.002    0.002
## 78    agg1 ~~   PR3  0.002 -0.001  -0.001   -0.003   -0.003
## 59      PR =~  agg3  0.001 -0.002  -0.001   -0.001   -0.001
## 58      PR =~  agg2  0.001 -0.001  -0.001   -0.001   -0.001
## 98    agg3 ~~  agg8  0.000  0.000   0.000    0.000    0.000
```

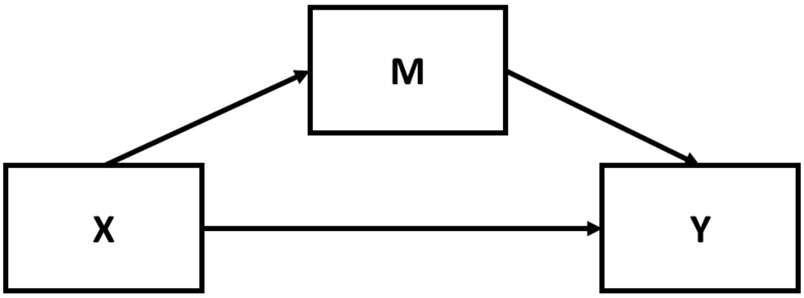
Reporting SEMs

- Main principles: transparency and reproducibility
- Method
  - *Describe the measurement model specification and criteria used to evaluate it (model fit etc.)*
  - *Describe the SEM model specification and criteria used to evaluate it*
  - *Explain how the model specification operationalises your hypothesis/hypotheses*
- Results
  - *Fit for the initial CFAs*
  - *Fit for the SEM (SRMR, RMSEA, TLI, CFI)*
  - *Any modifications made and why*
  - *All parameter estimates from the SEM*
    - diagram can again be helpful for visualising model
    - may need to show the structural and measurement parts of the model separately for visual clarity

Cautions regarding the use of SEM

- We assume the paths represent causal relations but this is an assumption
  - *Especially when using cross-sectional data*
- Well-fitting models do not guarantee that we have found the ‘correct’ model
- Our parameter estimates are correct only if the model is correctly specified

Mediation with SEM



- Last week we saw how we could test mediation (indirect effects) using path analysis
- These analyses can be affected by attenuation due to unreliability
- We can use latent measurement models for our predictor(s), mediator(s) and outcome(s) to overcome this

A SEM mediation example

- ADHD symptoms are known to be associated with:
  - *emotional dysregulation*
  - *depression*
- A researcher wants to test the hypothesis that emotional dysregulation mediates the relation between ADHD and depression
- We have:
  - *A 5-item measure of ADHD symptoms*
  - *A 5-item measure of emotional dysregulation*
  - *A 5-item measure of depression*
  - *n=720 participants*
- We will use SEM to test the researcher’s hypothesis

SEM mediation example

- The dataset

```
library(psych)
describe(ADHD_ED_dep)
```

```
##        vars   n  mean   sd median trimmed  mad   min  max range  skew kurtosis
## ADHD1     1 720 -0.03 1.02  -0.02   -0.02 0.98 -4.10 3.05  7.15 -0.12     0.16
## ADHD2     2 720 -0.02 1.03  -0.08   -0.04 0.95 -3.03 3.61  6.64  0.20     0.33
## ADHD3     3 720 -0.04 1.01  -0.08   -0.06 0.97 -2.64 2.75  5.39  0.15    -0.19
## ADHD4     4 720  0.02 1.02   0.01    0.03 1.09 -2.68 3.18  5.85 -0.01    -0.31
## ADHD5     5 720 -0.03 1.02  -0.03   -0.02 1.05 -2.99 3.03  6.03 -0.02    -0.19
## ED1       6 720  0.01 1.03   0.04    0.01 1.04 -3.18 3.22  6.40 -0.08    -0.07
## ED2       7 720  0.01 0.97  -0.03    0.01 0.98 -2.71 3.22  5.94  0.04    -0.21
## ED3       8 720  0.03 1.02   0.01    0.02 1.06 -3.08 3.12  6.20  0.11    -0.12
## ED4       9 720  0.00 1.02   0.03    0.00 1.05 -3.40 3.09  6.49 -0.01    -0.24
## ED5      10 720 -0.04 0.99  -0.07   -0.04 0.96 -3.69 2.59  6.28 -0.03     0.08
## Dep1     11 720  0.01 1.01   0.00    0.01 1.02 -2.58 3.20  5.79  0.01    -0.25
## Dep2     12 720  0.01 1.07   0.04    0.01 1.15 -3.06 3.55  6.61  0.04    -0.06
## Dep3     13 720  0.00 1.03  -0.03   -0.02 0.98 -3.08 2.74  5.82  0.09    -0.03
## Dep4     14 720  0.00 1.05  -0.02   -0.02 1.02 -3.15 3.23  6.39  0.15    -0.06
## Dep5     15 720  0.01 1.04  -0.02    0.01 1.04 -3.85 3.08  6.93 -0.01    -0.16
##          se
## ADHD1 0.04
## ADHD2 0.04
## ADHD3 0.04
## ADHD4 0.04
## ADHD5 0.04
## ED1   0.04
## ED2   0.04
## ED3   0.04
## ED4   0.04
## ED5   0.04
## Dep1  0.04
## Dep2  0.04
## Dep3  0.04
## Dep4  0.04
## Dep5  0.04
```

SEM mediation example - compare with path analysis

- For comparison let's first test mediation using path analysis
- First we have to create sum or average scores for each construct:

```
attach(ADHD_ED_dep)
ADHD_ED_dep$ADHD_score<-(ADHD1+ADHD2+ADHD3+ADHD4+ADHD5)/5 #ADHD mean score
ADHD_ED_dep$ED_score<-(ED1+ED2+ED3+ED4+ED5)/5 #emotional dysregulation mean score
ADHD_ED_dep$Dep_score<-(Dep1+Dep2+Dep3+Dep4+Dep5)/5 #depression mean score
detach(ADHD_ED_dep)

describe(ADHD_ED_dep)
```

```
##            vars   n  mean   sd median trimmed  mad   min  max range  skew
## ADHD1         1 720 -0.03 1.02  -0.02   -0.02 0.98 -4.10 3.05  7.15 -0.12
## ADHD2         2 720 -0.02 1.03  -0.08   -0.04 0.95 -3.03 3.61  6.64  0.20
## ADHD3         3 720 -0.04 1.01  -0.08   -0.06 0.97 -2.64 2.75  5.39  0.15
## ADHD4         4 720  0.02 1.02   0.01    0.03 1.09 -2.68 3.18  5.85 -0.01
## ADHD5         5 720 -0.03 1.02  -0.03   -0.02 1.05 -2.99 3.03  6.03 -0.02
## ED1           6 720  0.01 1.03   0.04    0.01 1.04 -3.18 3.22  6.40 -0.08
## ED2           7 720  0.01 0.97  -0.03    0.01 0.98 -2.71 3.22  5.94  0.04
## ED3           8 720  0.03 1.02   0.01    0.02 1.06 -3.08 3.12  6.20  0.11
## ED4           9 720  0.00 1.02   0.03    0.00 1.05 -3.40 3.09  6.49 -0.01
## ED5          10 720 -0.04 0.99  -0.07   -0.04 0.96 -3.69 2.59  6.28 -0.03
## Dep1         11 720  0.01 1.01   0.00    0.01 1.02 -2.58 3.20  5.79  0.01
## Dep2         12 720  0.01 1.07   0.04    0.01 1.15 -3.06 3.55  6.61  0.04
## Dep3         13 720  0.00 1.03  -0.03   -0.02 0.98 -3.08 2.74  5.82  0.09
## Dep4         14 720  0.00 1.05  -0.02   -0.02 1.02 -3.15 3.23  6.39  0.15
## Dep5         15 720  0.01 1.04  -0.02    0.01 1.04 -3.85 3.08  6.93 -0.01
## ADHD_score   16 720 -0.02 0.77   0.00   -0.02 0.75 -2.20 2.93  5.13  0.03
## ED_score     17 720  0.00 0.79  -0.01   -0.01 0.79 -2.51 2.52  5.04  0.12
## Dep_score    18 720  0.00 0.82  -0.04   -0.01 0.84 -2.48 2.38  4.86  0.12
##            kurtosis   se
## ADHD1          0.16 0.04
## ADHD2          0.33 0.04
## ADHD3         -0.19 0.04
## ADHD4         -0.31 0.04
## ADHD5         -0.19 0.04
## ED1           -0.07 0.04
## ED2           -0.21 0.04
## ED3           -0.12 0.04
## ED4           -0.24 0.04
## ED5            0.08 0.04
## Dep1          -0.25 0.04
## Dep2          -0.06 0.04
## Dep3          -0.03 0.04
## Dep4          -0.06 0.04
## Dep5          -0.16 0.04
## ADHD_score     0.12 0.03
## ED_score       0.08 0.03
## Dep_score     -0.14 0.03
```

SEM mediation example - estimate the reliability of the ADHD scores

```
omega(ADHD_ED_dep[ ,c('ADHD1','ADHD2','ADHD3','ADHD4','ADHD5')], nfactors=1)
```

```
## Omega_h for 1 factor is not meaningful, just omega_t
```

```
## Omega
## Call: omegah(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip,
##     digits = digits, title = title, sl = sl, labels = labels,
##     plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option,
##     covar = covar)
## Alpha:                 0.81
## G.6:                   0.78
## Omega Hierarchical:    0.81
## Omega H asymptotic:    1
## Omega Total            0.81
##
## Schmid Leiman Factor loadings greater than  0.2
##          g  F1*   h2   u2 p2
## ADHD1 0.52      0.27 0.73  1
## ADHD2 0.72      0.52 0.48  1
## ADHD3 0.61      0.37 0.63  1
## ADHD4 0.76      0.58 0.42  1
## ADHD5 0.79      0.62 0.38  1
##
## With eigenvalues of:
##   g F1*
## 2.4 0.0
##
## general/max  Inf   max/min =   NaN
## mean percent general =  1    with sd =  0 and cv of  0
## Explained Common Variance of the general factor =  1
##
## The degrees of freedom are 5  and the fit is  0.01
## The number of observations was  720  with Chi Square =  5.25  with prob <  0.39
## The root mean square of the residuals is  0.01
## The df corrected root mean square of the residuals is  0.02
## RMSEA index =  0.008  and the 10 % confidence intervals are  0 0.053
## BIC =  -27.65
##
## Compare this with the adequacy of just a general factor and no group factors
## The degrees of freedom for just the general factor are 5  and the fit is  0.01
## The number of observations was  720  with Chi Square =  5.25  with prob <  0.39
## The root mean square of the residuals is  0.01
## The df corrected root mean square of the residuals is  0.02
##
## RMSEA index =  0.008  and the 10 % confidence intervals are  0 0.053
## BIC =  -27.65
##
## Measures of factor score adequacy
##                                                  g F1*
## Correlation of scores with factors            0.91   0
## Multiple R square of scores with factors      0.83   0
## Minimum correlation of factor score estimates 0.67  -1
##
##  Total, General and Subset omega for each subset
##                                                  g  F1*
## Omega total for total scores and subscales    0.81 0.81
## Omega general for total scores and subscales  0.81 0.81
## Omega group for total scores and subscales    0.00 0.00
```

SEM mediation example - estimate the reliability of the ED scores

```
omega(ADHD_ED_dep[ ,c('ED1','ED2','ED3','ED4','ED5')], nfactors=1)
```

```
## Omega_h for 1 factor is not meaningful, just omega_t
```

```
## Omega
## Call: omegah(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip,
##     digits = digits, title = title, sl = sl, labels = labels,
##     plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option,
##     covar = covar)
## Alpha:                 0.84
## G.6:                   0.81
## Omega Hierarchical:    0.84
## Omega H asymptotic:    1
## Omega Total            0.84
##
## Schmid Leiman Factor loadings greater than  0.2
##        g  F1*   h2   u2 p2
## ED1 0.69      0.48 0.52  1
## ED2 0.67      0.46 0.54  1
## ED3 0.76      0.58 0.42  1
## ED4 0.83      0.69 0.31  1
## ED5 0.62      0.38 0.62  1
##
## With eigenvalues of:
##   g F1*
## 2.6 0.0
##
## general/max  4.662212e+16   max/min =   1
## mean percent general =  1    with sd =  0 and cv of  0
## Explained Common Variance of the general factor =  1
##
## The degrees of freedom are 5  and the fit is  0
## The number of observations was  720  with Chi Square =  2.48  with prob <  0.78
## The root mean square of the residuals is  0.01
## The df corrected root mean square of the residuals is  0.01
## RMSEA index =  0  and the 10 % confidence intervals are  0 0.034
## BIC =  -30.42
##
## Compare this with the adequacy of just a general factor and no group factors
## The degrees of freedom for just the general factor are 5  and the fit is  0
## The number of observations was  720  with Chi Square =  2.48  with prob <  0.78
## The root mean square of the residuals is  0.01
## The df corrected root mean square of the residuals is  0.01
##
## RMSEA index =  0  and the 10 % confidence intervals are  0 0.034
## BIC =  -30.42
##
## Measures of factor score adequacy
##                                                  g F1*
## Correlation of scores with factors            0.93   0
## Multiple R square of scores with factors      0.86   0
## Minimum correlation of factor score estimates 0.71  -1
##
##  Total, General and Subset omega for each subset
##                                                  g  F1*
## Omega total for total scores and subscales    0.84 0.84
## Omega general for total scores and subscales  0.84 0.84
## Omega group for total scores and subscales    0.00 0.00
```

SEM mediation example - estimate the reliability of the depression scores

```
omega(ADHD_ED_dep[ ,c('Dep1','Dep2','Dep3','Dep4','Dep5')], nfactors=1)
```

```
## Omega_h for 1 factor is not meaningful, just omega_t
```

```
## Omega
## Call: omegah(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip,
##     digits = digits, title = title, sl = sl, labels = labels,
##     plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option,
##     covar = covar)
## Alpha:                 0.85
## G.6:                   0.82
## Omega Hierarchical:    0.85
## Omega H asymptotic:    1
## Omega Total            0.85
##
## Schmid Leiman Factor loadings greater than  0.2
##         g  F1*   h2   u2 p2
## Dep1 0.61      0.37 0.63  1
## Dep2 0.73      0.54 0.46  1
## Dep3 0.81      0.66 0.34  1
## Dep4 0.77      0.59 0.41  1
## Dep5 0.71      0.50 0.50  1
##
## With eigenvalues of:
##   g F1*
## 2.7 0.0
##
## general/max  Inf   max/min =   NaN
## mean percent general =  1    with sd =  0 and cv of  0
## Explained Common Variance of the general factor =  1
##
## The degrees of freedom are 5  and the fit is  0
## The number of observations was  720  with Chi Square =  2.78  with prob <  0.73
## The root mean square of the residuals is  0.01
## The df corrected root mean square of the residuals is  0.01
## RMSEA index =  0  and the 10 % confidence intervals are  0 0.037
## BIC =  -30.11
##
## Compare this with the adequacy of just a general factor and no group factors
## The degrees of freedom for just the general factor are 5  and the fit is  0
## The number of observations was  720  with Chi Square =  2.78  with prob <  0.73
## The root mean square of the residuals is  0.01
## The df corrected root mean square of the residuals is  0.01
##
## RMSEA index =  0  and the 10 % confidence intervals are  0 0.037
## BIC =  -30.11
##
## Measures of factor score adequacy
##                                                  g F1*
## Correlation of scores with factors            0.93   0
## Multiple R square of scores with factors      0.86   0
## Minimum correlation of factor score estimates 0.72  -1
##
##  Total, General and Subset omega for each subset
##                                                  g  F1*
## Omega total for total scores and subscales    0.85 0.85
## Omega general for total scores and subscales  0.85 0.85
## Omega group for total scores and subscales    0.00 0.00
```

SEM mediation example - conduct the path analysis

```
#specify the model
path_analysis<-
'Dep_score~ADHD_score+a*ED_score
 ED_score~b*ADHD_score
ind:=a*b' #the indirect effect
#estimate the model
path_analysis.est<-sem(path_analysis, data=ADHD_ED_dep, se='bootstrap')
```

SEM mediation example - path analysis output

```
summary(path_analysis.est, ci=T, standardized=T)
```

```
## lavaan 0.6-5 ended normally after 16 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                          5
##
##   Number of observations                           720
##
## Model Test User Model:
##
##   Test statistic                                 0.000
##   Degrees of freedom                                 0
##
## Parameter Estimates:
##
##   Standard errors                            Bootstrap
##   Number of requested bootstrap draws             1000
##   Number of successful bootstrap draws            1000
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
##   Dep_score ~
##     ADHD_score        0.235    0.045    5.263    0.000    0.146    0.321
##     ED_score   (a)    0.368    0.045    8.184    0.000    0.277    0.458
##   ED_score ~
##     ADHD_score (b)    0.610    0.030   20.646    0.000    0.550    0.665
##    Std.lv  Std.all
##
##     0.235    0.220
##     0.368    0.353
##
##     0.610    0.596
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
##    .Dep_score         0.491    0.025   19.710    0.000    0.439    0.538
##    .ED_score          0.397    0.022   18.109    0.000    0.355    0.439
##    Std.lv  Std.all
##     0.491    0.734
##     0.397    0.645
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
##     ind               0.225    0.029    7.826    0.000    0.169    0.285
##    Std.lv  Std.all
##     0.225    0.210
```

SEM mediation example - conduct SEM mediation

```
#specify the model
SEM<-'
ADHD=~ADHD1+ADHD2+ADHD3+ADHD4+ADHD5 # ADHD measurement model
ED=~ED1+ED2+ED3+ED4+ED5 # emotional dysregulation measurement model
Dep=~Dep1+Dep2+Dep3+Dep4+Dep5 #depression measurement model

#structural part of the model

Dep~ADHD+a*ED
ED~b*ADHD

ind:=a*b # the indirect effect'

#estimate the model
SEM.est<-sem(SEM, data=ADHD_ED_dep, se='bootstrap')
```

SEM mediation example - SEM output

```r
#view the model output
summary(SEM.est, ci=T, fit.measures=T, standardized=T)
```

```
## lavaan 0.6-5 ended normally after 35 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         33
##
##   Number of observations                           720
##
## Model Test User Model:
##
##   Test statistic                                61.954
##   Degrees of freedom                                87
##   P-value (Chi-square)                           0.981
##
## Model Test Baseline Model:
##
##   Test statistic                              4459.242
##   Degrees of freedom                               105
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    1.000
##   Tucker-Lewis Index (TLI)                       1.007
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -13343.757
##   Loglikelihood unrestricted model (H1)     -13312.780
##
##   Akaike (AIC)                               26753.514
##   Bayesian (BIC)                             26904.629
##   Sample-size adjusted Bayesian (BIC)        26799.845
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.000
##   90 Percent confidence interval - lower         0.000
##   90 Percent confidence interval - upper         0.000
##   P-value RMSEA <= 0.05                          1.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.018
##
## Parameter Estimates:
##
##   Standard errors                            Bootstrap
##   Number of requested bootstrap draws             1000
##   Number of successful bootstrap draws            1000
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
##   ADHD =~
##     ADHD1             1.000                               1.000    1.000
##     ADHD2             1.420    0.116   12.221    0.000    1.219    1.667
##     ADHD3             1.181    0.098   12.089    0.000    1.005    1.399
```

```
##     ADHD4               1.494    0.121   12.385    0.000    1.276    1.746
##     ADHD5               1.561    0.121   12.873    0.000    1.338    1.834
##   ED =~
##     ED1                 1.000                               1.000    1.000
##     ED2                 0.908    0.055   16.475    0.000    0.798    1.022
##     ED3                 1.103    0.054   20.500    0.000    1.007    1.217
##     ED4                 1.197    0.063   18.958    0.000    1.084    1.336
##     ED5                 0.876    0.056   15.720    0.000    0.775    0.999
##   Dep =~
##     Dep1                1.000                               1.000    1.000
##     Dep2                1.267    0.081   15.717    0.000    1.127    1.437
##     Dep3                1.353    0.083   16.224    0.000    1.210    1.538
##     Dep4                1.309    0.078   16.710    0.000    1.163    1.485
##     Dep5                1.184    0.080   14.782    0.000    1.047    1.352
##    Std.lv  Std.all
##
##     0.519    0.512
##     0.737    0.719
##     0.613    0.605
##     0.775    0.759
##     0.810    0.797
##
##     0.708    0.688
##     0.643    0.665
##     0.781    0.765
##     0.847    0.833
##     0.620    0.624
##
##     0.617    0.609
##     0.782    0.732
##     0.834    0.811
##     0.807    0.772
##     0.730    0.703
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
##   Dep ~
##     ADHD                0.265    0.087    3.066    0.002    0.104    0.444
##     ED          (a)     0.361    0.063    5.707    0.000    0.237    0.483
##   ED ~
##     ADHD        (b)     0.988    0.092   10.768    0.000    0.827    1.191
##    Std.lv  Std.all
##
##     0.223    0.223
##     0.414    0.414
##
##     0.725    0.725
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
##    .ADHD1               0.760    0.046   16.521    0.000    0.668    0.851
##    .ADHD2               0.507    0.031   16.191    0.000    0.446    0.567
##    .ADHD3               0.652    0.038   17.182    0.000    0.576    0.729
##    .ADHD4               0.442    0.029   15.087    0.000    0.385    0.502
##    .ADHD5               0.376    0.027   13.932    0.000    0.321    0.432
##    .ED1                 0.558    0.033   17.085    0.000    0.494    0.623
##    .ED2                 0.519    0.030   17.358    0.000    0.461    0.579
##    .ED3                 0.431    0.030   14.595    0.000    0.370    0.488
##    .ED4                 0.316    0.023   13.496    0.000    0.268    0.359
##    .ED5                 0.603    0.034   17.662    0.000    0.533    0.665
##    .Dep1                0.645    0.035   18.399    0.000    0.576    0.717
##    .Dep2                0.528    0.035   15.255    0.000    0.458    0.597
##    .Dep3                0.363    0.030   12.119    0.000    0.302    0.416
##    .Dep4                0.442    0.031   14.481    0.000    0.387    0.510
##    .Dep5                0.545    0.035   15.745    0.000    0.475    0.614
##     ADHD                0.269    0.039    6.868    0.000    0.197    0.354
##    .ED                  0.238    0.030    7.830    0.000    0.183    0.301
##    .Dep                 0.245    0.028    8.782    0.000    0.191    0.302
##    Std.lv  Std.all
```

```
##     0.760    0.738
##     0.507    0.483
##     0.652    0.635
##     0.442    0.424
##     0.376    0.364
##     0.558    0.527
##     0.519    0.557
##     0.431    0.414
##     0.316    0.306
##     0.603    0.610
##     0.645    0.629
##     0.528    0.464
##     0.363    0.343
##     0.442    0.404
##     0.545    0.505
##     1.000    1.000
##     0.475    0.475
##     0.645    0.645
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
##     ind               0.356    0.065    5.459    0.000    0.228    0.494
##    Std.lv  Std.all
##     0.300    0.300
```

Path vs SEM models

- SEM models can adjust for attenuation due to unreliability
- This means that structural associations tend to be larger (and arguably more accurate)
- It makes SEM preferable to path analysis

SEM summary

- Full SEM models combine CFA and path analysis
- The steps in a SEM are:
  - *Test the measurement models (specification, estimation, evaluation & modification)*
  - *Specify the full SEM*
  - *Estimate the full SEM*
  - *Evaluate the full SEM*
- Paths are usally assumed to represent causal effects but this is only an assumption