Assignment 2

SEM 2: Structural Equation Modeling

Please hand in a `.pdf` file containing your report and a `.R` containing your codes or screenshots of every Jasp analysis. The deadline of this assignment is Tuesday May 21 13:00.

Assignment

Part 1

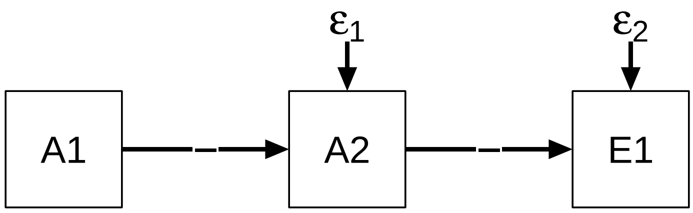
You can obtain the a dataset (named `bfi`) in R as follows:

```
library("psych")
data("bfi")
```

This data contains observations of a questionaire designed to measure the big 5 of personality:

Item label	Item description
A1	Am indifferent to the feelings of others
A2	Inquire about others' well-being
A3	Know how to comfort others
A4	Love children
A5	Make people feel at ease
C1	Am exacting in my work
C2	Continue until everything is perfect
C3	Do things according to a plan
C4	Do things in a half-way manner
C5	Waste my time
E1	Don't talk a lot
E2	Find it difficult to approach others
E3	Know how to captivate people
E4	Make friends easily
E5	Take charge
N1	Get angry easily
N2	Get irritated easily
N3	Have frequent mood swings
N4	Often feel blue
N5	Panic easily
O1	Am full of ideas
O2	Avoid difficult reading material
O3	Carry the conversation to a higher level
O4	Spend time reflecting on things
O5	Will not probe deeply into a subject
gender	Males = 1 Females = 2
education	1 = HS, 2 = finished HS, 3 = some college, 4 = college graduate 5 = graduate degree
age	Age in years

Of note, these items are *not* from the Big Five Inventory. See `?bfi` for details. For simplicity, treat all variables (including, e.g., gender and education) as continuous (do not use polychoric correlations or robust estimators). We can think of theories on how these variables interact with one-another. For example, I can reason that being "indifferent to the feelings of others" (A1) causes someone not to "inquire about others' well-being" (A2), and that in turn not inquiring about others well-being causes to "don't talk a lot" (E1).



The code for this path diagram are available here. I can fit this model in lavaan:

```
library("lavaan")
Model <- '
E1 ~ A2
A2 ~ A1
'
fit <- sem(Model, bfi)
fit
```

```
## lavaan 0.6-3 ended normally after 15 iterations
##
##   Optimization method                           NLMINB
##   Number of free parameters                          4
##
##                                                   Used       Total
##   Number of observations                          2737        2800
##
##   Estimator                                         ML
##   Model Fit Test Statistic                       3.316
##   Degrees of freedom                                 1
##   P-value (Chi-square)                           0.069
```

```
round(fitMeasures(fit)[c("rmsea","cfi","tli","srmr")],2)
```

```
## rmsea   cfi   tli  srmr
##  0.03  0.99  0.98  0.01
```

```
parameterEstimates(fit)
```

```
##   lhs op rhs    est    se       z pvalue ci.lower ci.upper
## 1  E1  ~  A2 -0.291 0.026 -11.193      0   -0.342   -0.240
## 2  A2  ~  A1 -0.285 0.015 -18.983      0   -0.314   -0.255
## 3  E1 ~~  E1  2.541 0.069  36.993      0    2.406    2.676
## 4  A2 ~~  A2  1.217 0.033  36.993      0    1.152    1.281
## 5  A1 ~~  A1  1.977 0.000      NA     NA    1.977    1.977
```

This model fits well!

Question 1 (1 point) Suppose we intervene on this system and make someone very indifferent to the feelings of others (such that that person responds higher on item A1). If the above causal model is true, do we then expect the person to talk *more*, *less*, or *the same* as compared to people we did not intervene on? Tip: you do not have to calculate anything, conceptual reasoning is enough. ▪

Question 2 (1 point) Think of causal model using variables from the `bfi` dataset. Use 4 or 5 variables, and make sure your theorized model is *acyclic*. Draw the path diagram of the model (1 point), including the expected sign (- or +, or use colors) of each regression (you may simplify the path diagram by not drawing exogenous variables as I did above). You do not have to explain your reasoning behind the model.

Question 3 (2 points) Fit your hypothesized model using Lavaan (1 point) and judge the fit (50 words max; 1 point). You do not have to modify the model to improve fit.
Note: when fitting models in lavaan using `sem()`, the function automatically adds covariances between dependent variables when only observed variables are used. The reason for this is to be similar to multivariate regression in which you aim to explain the variance, not the covariance, of outcome variables. We do not want this though. The option that should be usable to disable this behavior appears bugged, so the only way to disable this behavior (unless you use the `lavaan()` function and specify all parameters) is to manually force some residual covariances to be zero:

```
Model <- '
E1 ~ A2
A1 ~ A2
E1 ~~ 0*A1
'
```

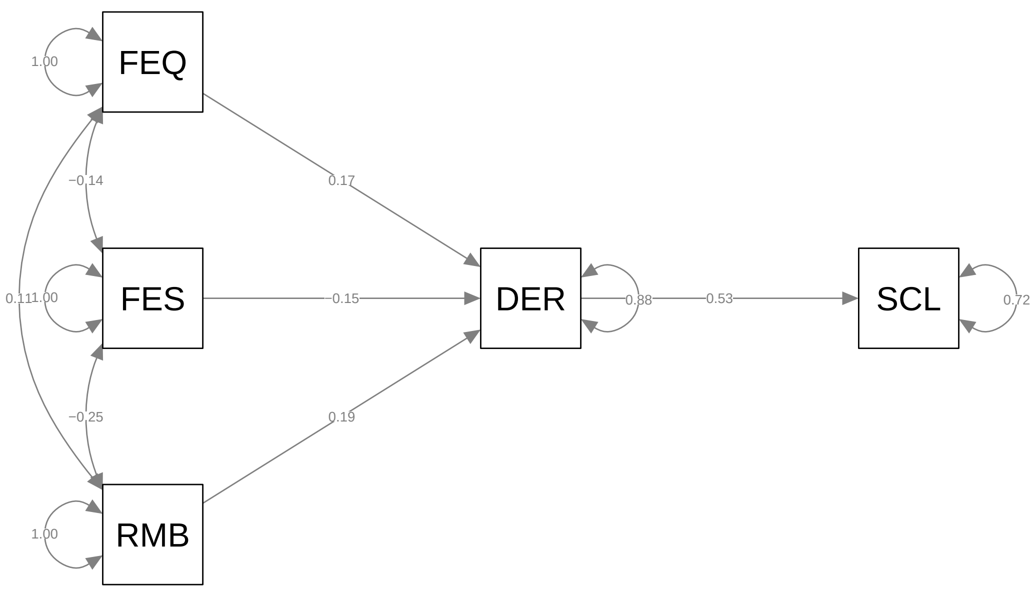
Use *semPlot* to check if you specified the correct model!

Question 4 (1 point) Is an equivalent model possible using the same variables? Why (not)?

Part 2

Given the following SEM analysis, which was originally reported by Suveg, Morelen, Brewer, and Thomassin (2010):

```
library("semPlot")
library("lavaan")
suveg.r <- c(
'1.00
-0.25 1.00
 0.11 -0.14 1.00
 0.25 -0.22 0.21 1.00
0.18 -0.15 0.19 0.53 1.00')

suveg.r <- getCov(suveg.r, names = c("RMBI", "FES", "FEQN", "DERS", "SCL90ANX"))
sd.suveg <- c(0.33, 0.62, 1.00, 0.54, 0.47)
suveg <- cor2cov(R = suveg.r, sds = sd.suveg)

mod1 <- '

ENTER MODEL HERE

'

fit1 <- sem(mod1, sample.cov = suveg, sample.nobs = 676, fixed.x = FALSE)

library("semPlot")
semPaths(fit1,"mod","std", layout = "tree2", rotation = 2,
         sizeMan = 10, curve = 2)
```



Question 5 (1 point) Fill in the model to reproduce the above analysis

We can obtain the implied variances and covariances using:

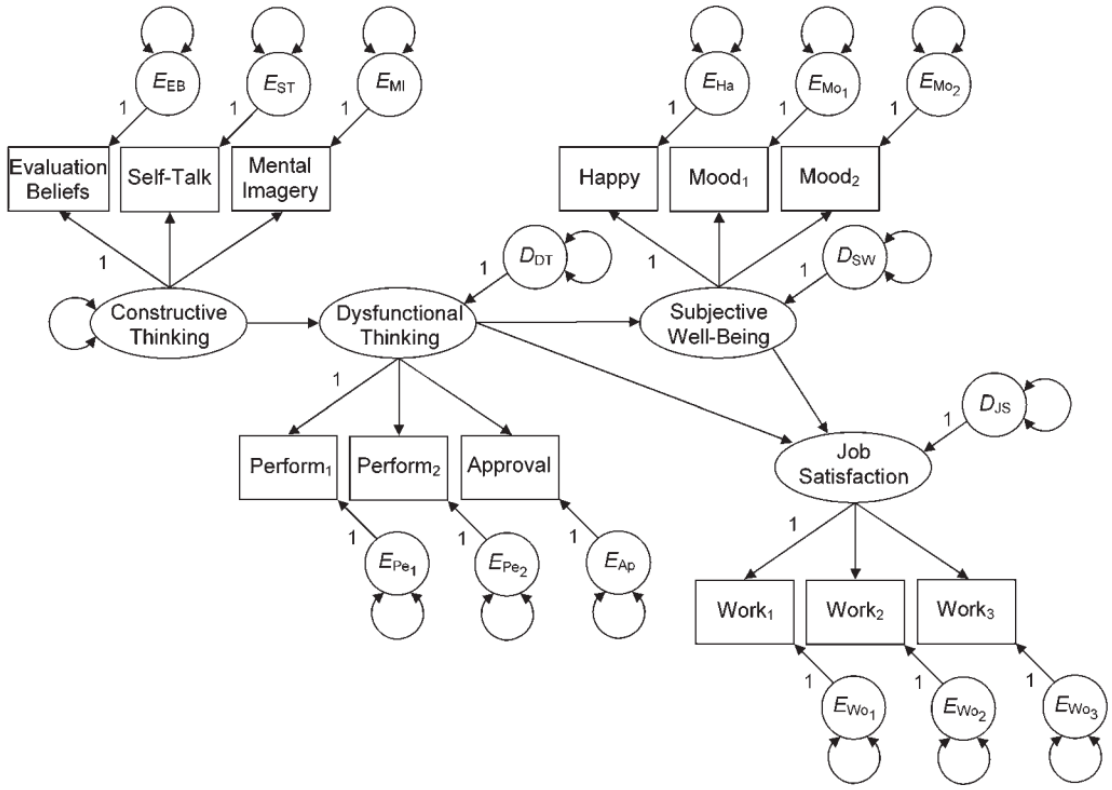
```
lavInspect(fit1, "sigma")

##           DERS   SCL90A RMBI   FES    FEQN
## DERS       0.291
## SCL90ANX   0.134  0.221
## RMBI       0.044  0.021  0.109
## FES       -0.074 -0.034 -0.051  0.384
## FEQN       0.113  0.052  0.036 -0.087  0.999
```

Question 6 (2 points) Using the Schur complement and the *model implied* variances and covariances, compute the conditional covariance between FES and SCL given DER. Round your result to 2 digits. Are these variables conditionally independent? ■

Question 7 (1 point) Can the arrow DER $\rightarrow$ SCL be changed into DER $\leftarrow$ SCL? Why (not)? ■

Part 3

Last week, we looked at a SEM analysis by Houghton and Jinkerson (2007), also reported by Kline (2015):



Replicating the analysis should have lead to the following model parameters: The original article contains a correlation matrix, standard deviations, and the sample size ($N = 263$). With this information, we can also construct the variance–covariance matrix, which I prepared for you in the file `houghton.csv` on Canvas, which can be loaded in R as follows:

```
covMat <- as.matrix(read.csv("houghton.csv"))
rownames(covMat) <- colnames(covMat)
mod <- '
consThinking =~ evaluation_beliefs + self_talk + mental_imagery
dysThining =~ perform1 + perform2 + approval
subWell =~ happy + mood1 + mood2
JobSat =~ work1 + work2 + work3
JobSat ~ subWell + dysThining
subWell ~ dysThining
dysThining ~ consThinking
'
library("lavaan")
fit <- lavaan::sem(mod, sample.cov = covMat, sample.nobs = 263)
fit

## lavaan 0.6-3 ended normally after 43 iterations
##
##   Optimization method                           NLMINB
##   Number of free parameters                         28
##
##   Number of observations                           263
##
```

```
##   Estimator                                         ML
##   Model Fit Test Statistic                      66.313
##   Degrees of freedom                                50
##   P-value (Chi-square)                           0.061
```

A not well known functionality of the semPlot package is that it can be used to perform matrix algebra using the terms `Lambda`, `Psi`, `Theta` and `Beta`. For example, this command obtains the factor loadings structure:

```
semMatrixAlgebra(fit, Lambda)

## model set to 'mplus'

##                     consThinking dysThining   subWell    JobSat
## evaluation_beliefs      1.000000  0.0000000 0.0000000 0.0000000
## self_talk               1.060066  0.0000000 0.0000000 0.0000000
## mental_imagery          1.860517  0.0000000 0.0000000 0.0000000
## perform1                0.000000  1.0000000 0.0000000 0.0000000
## perform2                0.000000  1.1262473 0.0000000 0.0000000
## approval                0.000000  0.9912901 0.0000000 0.0000000
## happy                   0.000000  0.0000000 1.0000000 0.0000000
## mood1                   0.000000  0.0000000 1.7678091 0.0000000
## mood2                   0.000000  0.0000000 0.8120267 0.0000000
## work1                   0.000000  0.0000000 0.0000000 1.0000000
## work2                   0.000000  0.0000000 0.0000000 1.0308362
## work3                   0.000000  0.0000000 0.0000000 0.8918587
```

and this command will compute $\mathbf{\Sigma}$:

```
Sigma <- semMatrixAlgebra(fit, Lambda %*% Imin(Beta, inverse = TRUE) %*%
                   Psi %*% t(Imin(Beta, inverse = TRUE)) %*%
                   t(Lambda) + Theta)

## model set to 'mplus'

round(Sigma, 2)

##                     evaluation_beliefs self_talk mental_imagery perform1
## evaluation_beliefs                0.50      0.23           0.40    -0.03
## self_talk                         0.23      1.26           0.42    -0.03
## mental_imagery                    0.40      0.42           1.00    -0.06
## perform1                         -0.03     -0.03          -0.06     0.34
## perform2                         -0.03     -0.04          -0.06     0.27
## approval                         -0.03     -0.03          -0.06     0.23
## happy                             0.01      0.01           0.02    -0.08
## mood1                             0.02      0.02           0.03    -0.14
## mood2                             0.01      0.01           0.02    -0.06
## work1                             0.02      0.02           0.03    -0.13
## work2                             0.02      0.02           0.03    -0.14
## work3                             0.02      0.02           0.03    -0.12
##                     perform2 approval happy mood1 mood2 work1 work2 work3
## evaluation_beliefs     -0.03    -0.03  0.01  0.02  0.01  0.02  0.02  0.02
## self_talk              -0.04    -0.03  0.01  0.02  0.01  0.02  0.02  0.02
## mental_imagery         -0.06    -0.06  0.02  0.03  0.02  0.03  0.03  0.03
## perform1                0.27     0.23 -0.08 -0.14 -0.06 -0.13 -0.14 -0.12
## perform2                0.37     0.26 -0.09 -0.16 -0.07 -0.15 -0.15 -0.13
## approval                0.26     0.53 -0.08 -0.14 -0.06 -0.13 -0.14 -0.12
## happy                  -0.09    -0.08  0.31  0.21  0.09  0.13  0.13  0.11
```

```
## mood1                  -0.16    -0.14  0.21  0.58  0.17  0.22  0.23  0.20
## mood2                  -0.07    -0.06  0.09  0.17  0.27  0.10  0.11  0.09
## work1                  -0.15    -0.13  0.13  0.22  0.10  0.88  0.64  0.55
## work2                  -0.15    -0.14  0.13  0.23  0.11  0.64  1.03  0.57
## work3                  -0.13    -0.12  0.11  0.20  0.09  0.55  0.57  0.87
```

(the function `Imin(..., inverse = TRUE)` is a helper function to compute $(\boldsymbol{I}-\boldsymbol{B})^{-1}$).

Question 8 (1 point) Use `semMatrixAlgebra` to compute the variance–covariance matrix of the latent variables (tip: see the slides). ■

Question 9 (1 point) Compute $\text{Cov}(\text{ConstructiveThinking}, \text{JobSatisfaction} \mid \text{DysfunctionalThinking})$ ■

Question 10 (1 point) Compute $\mathscr{E}(\text{JobSatisfaction} \mid \text{See}(\text{ConstructiveThinking} = 10))$, assuming all variables are centered. ■

Question 11 (optional; 1 bonus point) Compute $\mathscr{E}(\text{JobSatisfaction} \mid \text{Do}(\text{DysfunctionalThinking} = 10))$, assuming all variables are centered. ■

Question 12 (1 point) Are the following statements true or false given the model? Briefly (one short sentence) explain your answer.

- Intervening on the variable “Happy” will impact the variables “Mood1” and “Mood2”.
- A person that has a high job satisfaction is more likely to score higher on the variable “happy”.
- A model in which Subjective well-being $\rightarrow$ Job satisfaction is changed into Job satisfaction $\rightarrow$ Subjective well-being is equivalent to the model above.

■

References

Houghton, J. D., & Jinkerson, D. L. (2007). Constructive thought strategies and job satisfaction: A preliminary examination. *Journal of Business and Psychology*, *22*(1), 45–53.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Suveg, C., Morelen, D., Brewer, G. A., & Thomassin, K. (2010). The emotion dysregulation model of anxiety: A preliminary path analytic examination. *Journal of Anxiety Disorders*, *24*(8), 924–930.