Assignment 3

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 19 13:00.

Assignment

Part 1: Eurovision

Every year, with the exception of 2020, European countries (and Australia, for some reason) come together to celebrate diversity in culture and musical quality through the Eurovision song contest:

The setup of the Eurovision is that about 50 countries participate in semi-finals. The best songs are then selected to participate in the finale. Four countries are exempt from this—France, Germany, Spain, and the United Kingdom—as these are the countries that contribute the most to Eurovision. Songs from these countries will automatically enter the finale.

> **Question 1 (1 point)**
> Assume that the *quality of song* is independent of the *amount a country contributes to Eurovision*. If we only look at songs in the Eurovision finale, will these two aspects still be independent? Or can we expect that the quality of songs from France, Germany, Spain, and the United Kingdom to be, on average, *better* or *worse* than songs from other countries? Explain your answer.

After entering the finale, countries vote on other songs and the song with the most votes is proclaimed the winner.

> **Question 2 (1 point)**   Given the following statement:
>
> > The better the quality of the song, the more likely that song is to win.
>
> Which of the following statements is true?
> - The winning song is the song with the highest quality
> - Bribing the jury to give more points will increase the quality of the song
> - We can predict that the winning song had a high quality
> - Improving the quality of the song will increase the chance of winning

Part 2: SEM models

Last week, we looked at a SEM analysis by Houghton and Jinkerson (2007), also reported by Kline (2015):
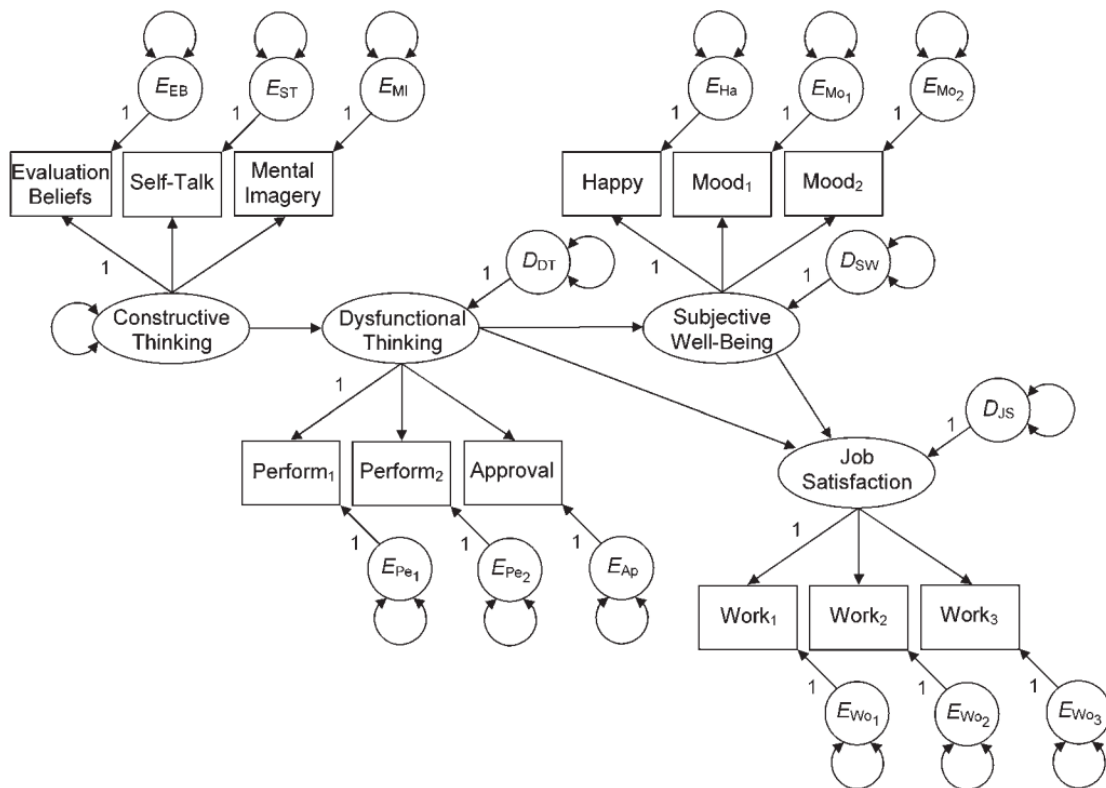
Replicating the analysis should have lead to the following model parameters: The original article contains a correlation matrix, standard deviations, and the sample size ($N = 263$). With this information, we can also construct the variance–covariance matrix, which I prepared for you in the file `houghton.csv` 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")

## This is lavaan 0.6-6.1525
## lavaan is BETA software!  Please report any bugs.

fit <- lavaan::sem(mod, sample.cov = covMat, sample.nobs = 263)
fit

## lavaan 0.6-6.1525 ended normally after 43 iterations
##
##   Estimator                                         ML
```

```
##    Optimization method                           NLMINB
##    Number of free parameters                         28
##
##    Number of observations                           263
##
## Model Test User Model:
##
##    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:

```
library("semPlot")
semMatrixAlgebra(fit, Lambda)

##                     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 Σ:

```
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 3 (1 point)** Use `semMatrixAlgebra` to compute the variance–covariance matrix of the latent variables (tip: see the slides). ∎

> **Question 4 (1 point)** Compute $\mathrm{Cov}(\mathrm{ConstructiveThinking}, \mathrm{JobSatisfaction} \mid \mathrm{DysfunctionalThinking})$ ∎

> **Question 5 (1 point)** Compute $\mathscr{E}(\mathrm{JobSatisfaction} \mid \mathrm{See}(\mathrm{ConstructiveThinking} = 10))$, assuming all variables are centered. ∎

> **Question 6 (1 point)** Compute $\mathscr{E}(\mathrm{JobSatisfaction} \mid \mathrm{Do}(\mathrm{ConstructiveThinking} = 10))$, assuming all variables are centered. ∎

> **Question 7 (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.
>
> ∎

Kuss et al. (2017) report the following SEM model:

Figure 2. *Structural equation model including standardised coefficients for direct (GSI) and indirect (GSI*Media & GSI*Substance) paths predicting excessive Internet use*

<div style="border-left: 4px solid orange; padding-left: 1em;">
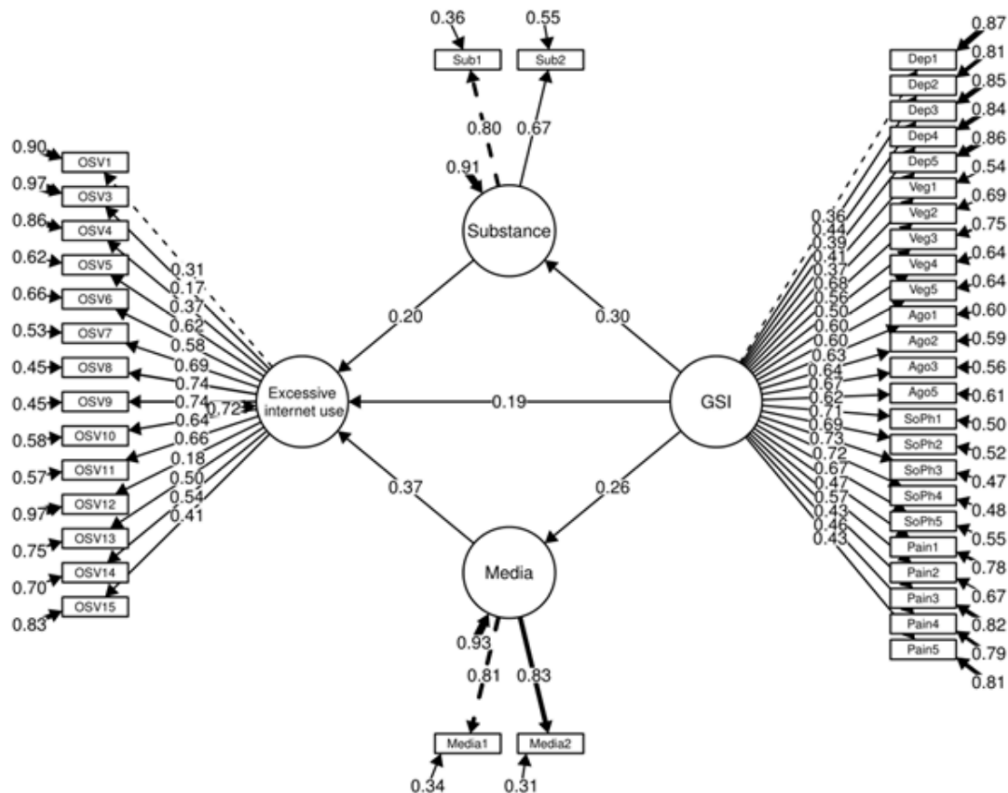
Question 8 (1 point)

The authors do not report the degrees of freedom of their model. Give the degrees of freedom (use only the figure above, do not read the paper itself).

</div>

<div style="border-left: 4px solid orange; padding-left: 1em;">

Question 9 (1 point) Are the following statements true or false?
- Excessive internet use ⊥⊥ GSI | Substance
- Substance ⊥⊥ Media | Excessive internet use
- Substance ⊥⊥ Media | GSI

</div>

<div style="border-left: 4px solid orange; padding-left: 1em;">

Question 10 (1 point) Do the following sugested replacements lead to an equivalent model or not?
- Replacing Substance → Excessive internet use with Substance ← Excessive internet use
- Replacing GSI → Substance with GSI ← Substance
- Replacing GSI → Excessive internet use with GSI ↔ Excessive internet use (a residual / exogenous covariance)

</div>

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.

Kuss, D. J., Dunn, T. J., Wölfling, K., Müller, K. W., Hędzelek, M., & Marcinkowski, J. (2017). Excessive internet use and psychopathology: The role of coping. *Clinical Neuropsychiatry: Journal of Treatment Evaluation*, *14*(1), 73–81.