COS-D419 Factor Analysis and Structural Equation Models 2023, Assignment 2

Ita Puusepp

2023-02-05

CFA & self-concept (SC)

Exercise 2.1

Specify and test the hypothesis given on the page 1 of the lecture material.

Draw conclusions based on the χ^2 statistic and the CFI, TLI, RMSEA, and SRMR indices.

What can you say about the parameter estimates?

Visualize the model.

Read in the data set:

Start by downloading the data file from Moodle to your Project folder!

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   1.0.1
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.3.0      v stringr 1.5.0
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(readr)

orig_data <- read_csv("ASC7INDM.CSV", show_col_types = FALSE)

# we will only use a subset of the data here:
SCdata <- orig_data %>% dplyr::select(starts_with("SDQ2N"))
glimpse(SCdata)
```

```
## Rows: 265
## Columns: 16
## $ SDQ2N01 <dbl> 6, 6, 4, 5, 6, 5, 1, 2, 5, 4, 2, 5, 6, 4, 4, 6, 6, 6, 5, 6, 6,~
```

```
## $ SDQ2N13 <dbl> 5, 6, 6, 5, 5, 5, 6, 1, 5, 6, 6, 5, 6, 3, 5, 6, 6, 6, 4, 5, 5,~
## $ SDQ2N25 <dbl> 4, 6, 6, 5, 5, 5, 1, 6, 6, 3, 6, 6, 6, 5, 5, 6, 6, 6, 6, 5, 4,~
## $ SDQ2N37 <dbl> 6, 6, 2, 6, 4, 3, 6, 4, 6, 6, 6, 5, 5, 5, 4, 5, 6, 4, 4, 6, 6,~
## $ SDQ2N04 <dbl> 3, 6, 6, 5, 3, 3, 4, 4, 6, 6, 5, 6, 5, 4, 4, 4, 4, 6, 5, 5, 3,~
## $ SDQ2N16 <dbl> 4, 6, 4, 6, 4, 2, 6, 4, 6, 5, 6, 6, 5, 5, 5, 5, 6, 5, 4, 6, 6,~
## $ SDQ2N28 <dbl> 4, 6, 6, 5, 4, 4, 6, 4, 6, 6, 6, 6, 5, 5, 5, 5, 6, 4, 2, 4, 4,~
## $ SDQ2N40 <dbl> 6, 6, 3, 6, 4, 4, 6, 6, 6, 6, 6, 6, 6, 5, 4, 4, 6, 6, 5, 5, 5,~
## $ SDQ2N10 <dbl> 2, 5, 6, 5, 4, 4, 1, 6, 5, 4, 2, 6, 5, 5, 5, 3, 4, 6, 5, 4, 6,~
## $ SDQ2N22 <dbl> 6, 6, 5, 6, 6, 4, 6, 6, 6, 6, 6, 6, 6, 5, 6, 6, 6, 6, 6, 3, 6,~
## $ SDQ2N34 <dbl> 1, 6, 4, 3, 5, 5, 1, 1, 5, 4, 5, 6, 5, 2, 5, 2, 3, 2, 1, 3, 3,~
## $ SDQ2N46 <dbl> 5, 6, 5, 5, 6, 6, 6, 5, 6, 6, 6, 6, 6, 6, 2, 5, 6, 6, 6, 6, 6,~
## $ SDQ2N07 <dbl> 6, 6, 6, 6, 3, 4, 5, 3, 6, 5, 6, 6, 6, 6, 4, 4, 6, 6, 6, 6, 3,~
## $ SDQ2N19 <dbl> 6, 6, 6, 6, 4, 5, 6, 4, 6, 6, 5, 6, 6, 6, 5, 5, 6, 6, 5, 5, 5,~
## $ SDQ2N31 <dbl> 6, 6, 3, 6, 4, 4, 6, 4, 6, 6, 6, 6, 6, 6, 5, 5, 6, 6, 5, 5, 5,~
## $ SDQ2N43 <dbl> 6, 6, 1, 5, 5, 4, 5, 6, 6, 6, 6, 6, 6, 6, 5, 6, 6, 6, 5, 6, 5,~
```

Explore the data (always do that!):

```r
library(psych)
```

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

## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```r
library(dplyr)
library(knitr)
library(tidyr)
library(corrplot)
```

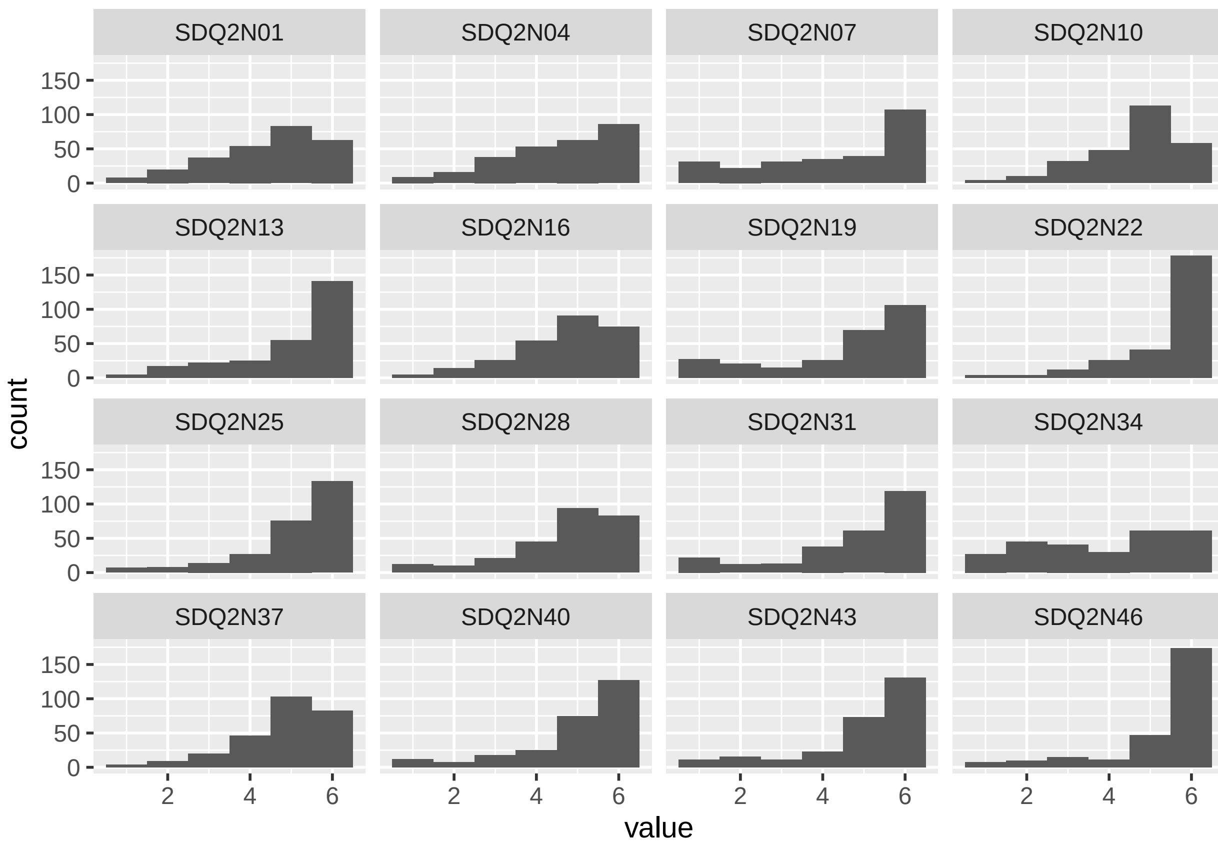
```
## corrplot 0.92 loaded
```

```r
# these are just examples - you can well do other things, too!

# basic statistics:
SCdata %>% describe() %>%
  as.data.frame() %>%
  select(mean, sd, min, max) %>%
  kable(digits = 2)
```

	mean	sd	min	max
SDQ2N01	4.41	1.35	1	6
SDQ2N13	5.00	1.36	1	6
SDQ2N25	5.10	1.23	1	6
SDQ2N37	4.83	1.14	1	6
SDQ2N04	4.52	1.40	1	6
SDQ2N16	4.65	1.24	1	6

	mean	sd	min	max
SDQ2N28	4.69	1.33	1	6
SDQ2N40	4.98	1.36	1	6
SDQ2N10	4.62	1.15	1	6
SDQ2N22	5.38	1.09	1	6
SDQ2N34	3.89	1.70	1	6
SDQ2N46	5.27	1.30	1	6
SDQ2N07	4.32	1.78	1	6
SDQ2N19	4.54	1.69	1	6
SDQ2N31	4.74	1.57	1	6
SDQ2N43	4.98	1.40	1	6

```r
# histograms:
SCdata %>% pivot_longer(cols = everything()) %>%
  ggplot(aes(x = value)) +
  geom_histogram(binwidth = 1) +
  facet_wrap(~name, nrow = 4, ncol = 4)
```

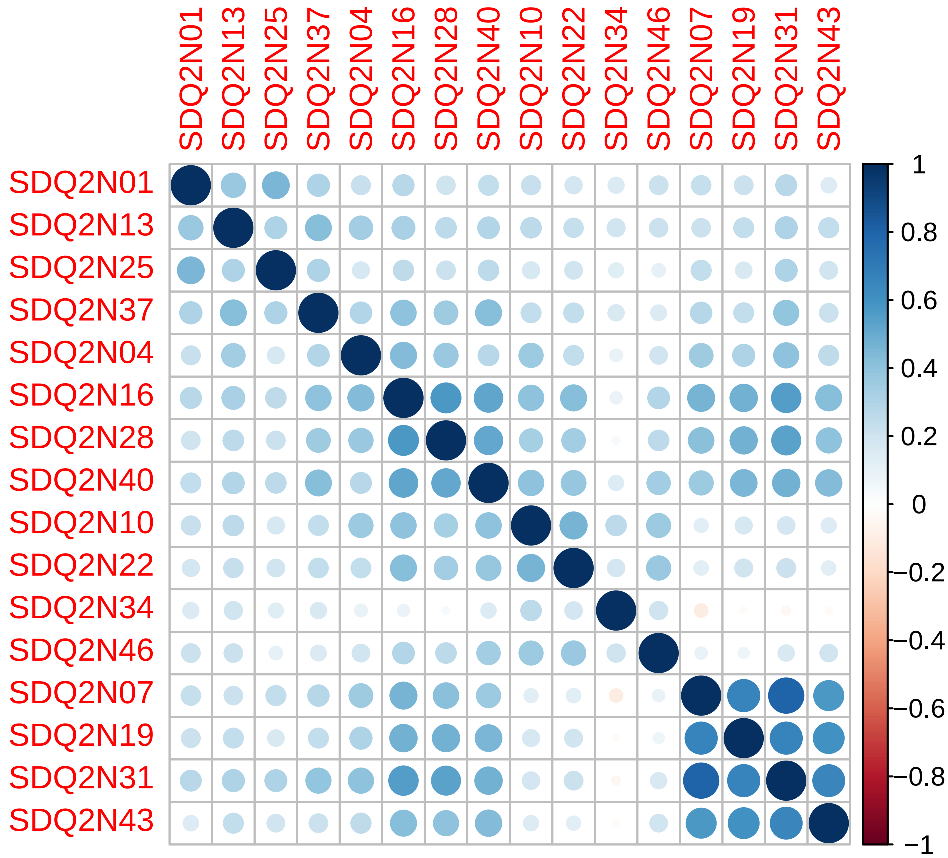



```r
# correlation matrix:
SCdata %>% cor() %>% round(digits = 2)
```

```
##          SDQ2N01 SDQ2N13 SDQ2N25 SDQ2N37 SDQ2N04 SDQ2N16 SDQ2N28 SDQ2N40 SDQ2N10
## SDQ2N01     1.00    0.37    0.45    0.31    0.22    0.28    0.21    0.25    0.23
## SDQ2N13     0.37    1.00    0.30    0.42    0.35    0.33    0.26    0.29    0.26
```

```
## SDQ2N25    0.45    0.30    1.00    0.31    0.18    0.26    0.21    0.26    0.18
## SDQ2N37    0.31    0.42    0.31    1.00    0.29    0.41    0.35    0.43    0.25
## SDQ2N04    0.22    0.35    0.18    0.29    1.00    0.43    0.37    0.27    0.37
## SDQ2N16    0.28    0.33    0.26    0.41    0.43    1.00    0.57    0.52    0.41
## SDQ2N28    0.21    0.26    0.21    0.35    0.37    0.57    1.00    0.51    0.34
## SDQ2N40    0.25    0.29    0.26    0.43    0.27    0.52    0.51    1.00    0.41
## SDQ2N10    0.23    0.26    0.18    0.25    0.37    0.41    0.34    0.41    1.00
## SDQ2N22    0.18    0.23    0.19    0.24    0.24    0.43    0.35    0.39    0.47
## SDQ2N34    0.16    0.20    0.14    0.16    0.10    0.09    0.04    0.14    0.26
## SDQ2N46    0.22    0.21    0.11    0.15    0.19    0.29    0.26    0.34    0.36
## SDQ2N07    0.23    0.21    0.25    0.29    0.35    0.46    0.41    0.37    0.12
## SDQ2N19    0.22    0.25    0.17    0.25    0.31    0.48    0.47    0.45    0.18
## SDQ2N31    0.27    0.31    0.30    0.40    0.40    0.55    0.54    0.48    0.19
## SDQ2N43    0.15    0.25    0.19    0.21    0.25    0.43    0.40    0.44    0.14
##         SDQ2N22 SDQ2N34 SDQ2N46 SDQ2N07 SDQ2N19 SDQ2N31 SDQ2N43
## SDQ2N01    0.18    0.16    0.22    0.23    0.22    0.27    0.15
## SDQ2N13    0.23    0.20    0.21    0.21    0.25    0.31    0.25
## SDQ2N25    0.19    0.14    0.11    0.25    0.17    0.30    0.19
## SDQ2N37    0.24    0.16    0.15    0.29    0.25    0.40    0.21
## SDQ2N04    0.24    0.10    0.19    0.35    0.31    0.40    0.25
## SDQ2N16    0.43    0.09    0.29    0.46    0.48    0.55    0.43
## SDQ2N28    0.35    0.04    0.26    0.41    0.47    0.54    0.40
## SDQ2N40    0.39    0.14    0.34    0.37    0.45    0.48    0.44
## SDQ2N10    0.47    0.26    0.36    0.12    0.18    0.19    0.14
## SDQ2N22    1.00    0.19    0.37    0.13    0.20    0.22    0.12
## SDQ2N34    0.19    1.00    0.21   -0.11   -0.03   -0.05   -0.03
## SDQ2N46    0.37    0.21    1.00    0.09    0.07    0.16    0.19
## SDQ2N07    0.13   -0.11    0.09    1.00    0.66    0.80    0.58
## SDQ2N19    0.20   -0.03    0.07    0.66    1.00    0.66    0.61
## SDQ2N31    0.22   -0.05    0.16    0.80    0.66    1.00    0.65
## SDQ2N43    0.12   -0.03    0.19    0.58    0.61    0.65    1.00
```

```
# correlation plot:
SCdata %>% cor() %>% corrplot()
```



The descriptives indicate that for each item the whole scale of the response options has been used (from min 1 to max 6). As visible from the means and histograms, some items seem to have almost a ceiling effect (e.g., SDQ2N22, SDQ2N46). The correlation plot indicates that the MSC items are especially highly correlated with each other, but also items of other scales can be seen to correlate more strongly with each other.

Define and estimate a CFA model:

```
library(lavaan) # install.packages("lavaan")
```

```
## This is lavaan 0.6-13
## lavaan is FREE software! Please report any bugs.
```

```
##
## Attaching package: 'lavaan'
```

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

```
# Define a CFA model using the lavaan package:
# NOTE: with the model definitions in lavaan syntax you have to
# SELECT all the code up to ' and then press Ctrl+Enter / Cmd+Enter
# when activating operations individually.
```

```r
model1 <- '# CFA model of self-concept (SC):
           GSC =~ SDQ2N01 + SDQ2N13 + SDQ2N25 + SDQ2N37
           ASC =~ SDQ2N04 + SDQ2N16 + SDQ2N28 + SDQ2N40
           ESC =~ SDQ2N10 + SDQ2N22 + SDQ2N34 + SDQ2N46
           MSC =~ SDQ2N07 + SDQ2N19 + SDQ2N31 + SDQ2N43
          '

# Estimate the model using the data defined earlier:
cfa1 <- cfa(model1, data = SCdata)

# Numerical summary of the model:
summary(cfa1, fit.measures = TRUE, standardized = TRUE)
```

```
## lavaan 0.6.13 ended normally after 49 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        38
##
##   Number of observations                           265
##
## Model Test User Model:
##
##   Test statistic                               159.112
##   Degrees of freedom                                98
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              1703.155
##   Degrees of freedom                               120
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.961
##   Tucker-Lewis Index (TLI)                       0.953
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -6562.678
##   Loglikelihood unrestricted model (H1)      -6483.122
##
##   Akaike (AIC)                               13201.356
##   Bayesian (BIC)                             13337.386
##   Sample-size adjusted Bayesian (SABIC)      13216.905
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.049
##   90 Percent confidence interval - lower         0.034
##   90 Percent confidence interval - upper         0.062
##   P-value H_0: RMSEA <= 0.050                    0.556
```

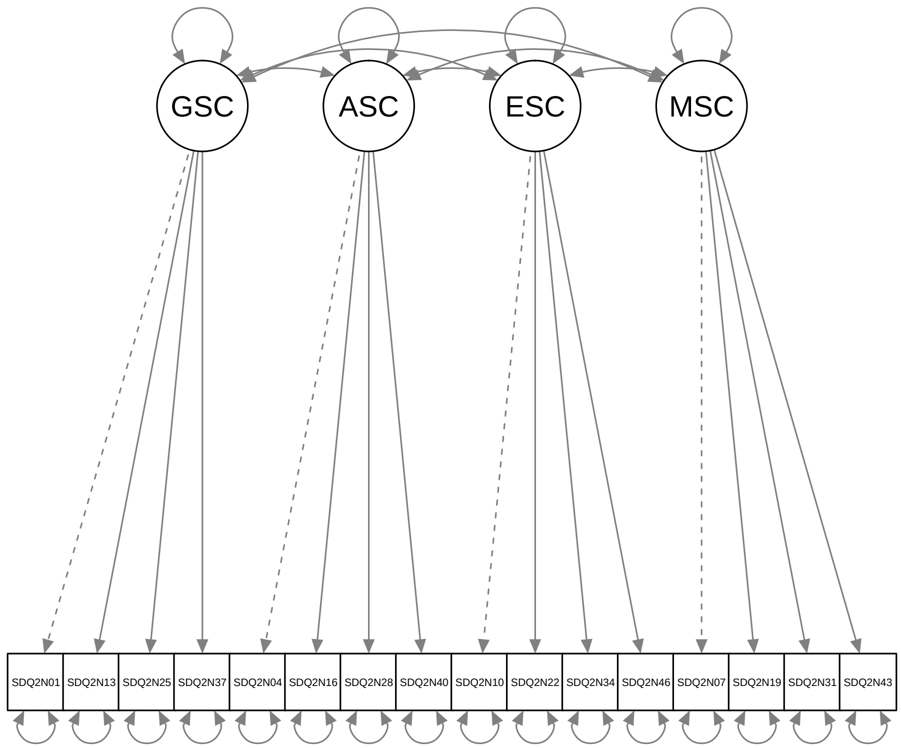
```
##   P-value H_0: RMSEA >= 0.080                    0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.048
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   GSC =~
##     SDQ2N01           1.000                               0.783    0.582
##     SDQ2N13           1.083    0.154    7.044    0.000    0.848    0.626
##     SDQ2N25           0.851    0.132    6.455    0.000    0.666    0.544
##     SDQ2N37           0.934    0.131    7.131    0.000    0.731    0.640
##   ASC =~
##     SDQ2N04           1.000                               0.749    0.536
##     SDQ2N16           1.279    0.150    8.520    0.000    0.958    0.774
##     SDQ2N28           1.247    0.154    8.097    0.000    0.934    0.703
##     SDQ2N40           1.259    0.156    8.048    0.000    0.943    0.695
##   ESC =~
##     SDQ2N10           1.000                               0.817    0.711
##     SDQ2N22           0.889    0.103    8.658    0.000    0.727    0.668
##     SDQ2N34           0.670    0.148    4.539    0.000    0.548    0.322
##     SDQ2N46           0.843    0.117    7.225    0.000    0.689    0.532
##   MSC =~
##     SDQ2N07           1.000                               1.519    0.854
##     SDQ2N19           0.841    0.058   14.495    0.000    1.277    0.755
##     SDQ2N31           0.952    0.049   19.516    0.000    1.446    0.923
##     SDQ2N43           0.655    0.049   13.298    0.000    0.995    0.712
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   GSC ~~
##     ASC               0.415    0.078    5.292    0.000    0.707    0.707
##     ESC               0.355    0.072    4.947    0.000    0.555    0.555
##     MSC               0.635    0.118    5.387    0.000    0.534    0.534
##   ASC ~~
##     ESC               0.464    0.078    5.921    0.000    0.758    0.758
##     MSC               0.873    0.134    6.519    0.000    0.767    0.767
##   ESC ~~
##     MSC               0.331    0.100    3.309    0.001    0.266    0.266
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .SDQ2N01           1.198    0.126    9.537    0.000    1.198    0.661
##    .SDQ2N13           1.119    0.124    9.019    0.000    1.119    0.609
##    .SDQ2N25           1.056    0.107    9.897    0.000    1.056    0.704
##    .SDQ2N37           0.771    0.087    8.821    0.000    0.771    0.591
##    .SDQ2N04           1.394    0.128   10.900    0.000    1.394    0.713
```

```
##    .SDQ2N16           0.616    0.068    9.020    0.000    0.616    0.402
##    .SDQ2N28           0.896    0.090    9.959    0.000    0.896    0.506
##    .SDQ2N40           0.952    0.095   10.029    0.000    0.952    0.517
##    .SDQ2N10           0.653    0.082    7.941    0.000    0.653    0.494
##    .SDQ2N22           0.657    0.075    8.735    0.000    0.657    0.554
##    .SDQ2N34           2.590    0.233   11.128    0.000    2.590    0.896
##    .SDQ2N46           1.201    0.118   10.183    0.000    1.201    0.717
##    .SDQ2N07           0.854    0.100    8.551    0.000    0.854    0.270
##    .SDQ2N19           1.228    0.121   10.153    0.000    1.228    0.429
##    .SDQ2N31           0.365    0.065    5.649    0.000    0.365    0.148
##    .SDQ2N43           0.964    0.092   10.473    0.000    0.964    0.493
##     GSC               0.613    0.137    4.464    0.000    1.000    1.000
##     ASC               0.561    0.126    4.453    0.000    1.000    1.000
##     ESC               0.668    0.116    5.749    0.000    1.000    1.000
##     MSC               2.307    0.273    8.460    0.000    1.000    1.000
```

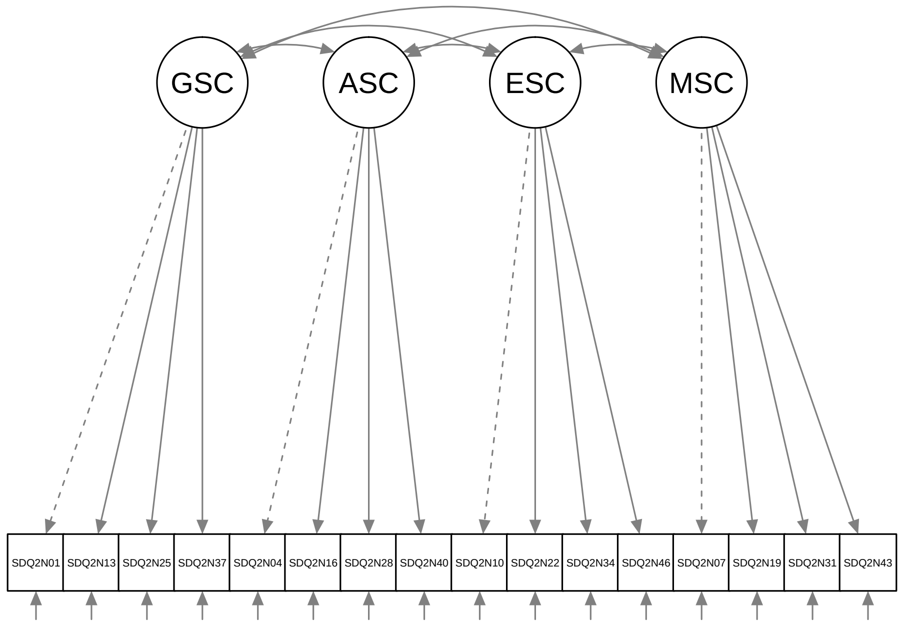
Model fit indices indicate a a very good fit, $\chi^2(98) = 159.11$, p $<$ .001, CFI = .96, TLI = .95, RMSEA = .05, and SRMR = .05. Although the p-value χ^2 test is $<$ .05, this might result from a large sample size, and all the other indices show a very good fit. Additionally, all of the factor loadings are significant at level p $<$ .001 and the standardized loadings range from .53 to .92, except for item SDQ2N34 that has a quite low loading of .32 (significant though). The same item also has quite a high residual variance, while otherwise the residual variances seem fine. Now looking back at the correlation plot, it is visible from there already that SDQ2N34 does not correlate as strongly with other items. So basically, the correlation plot can already give some insight regarding the results of the FA?

Visualize the CFA model:

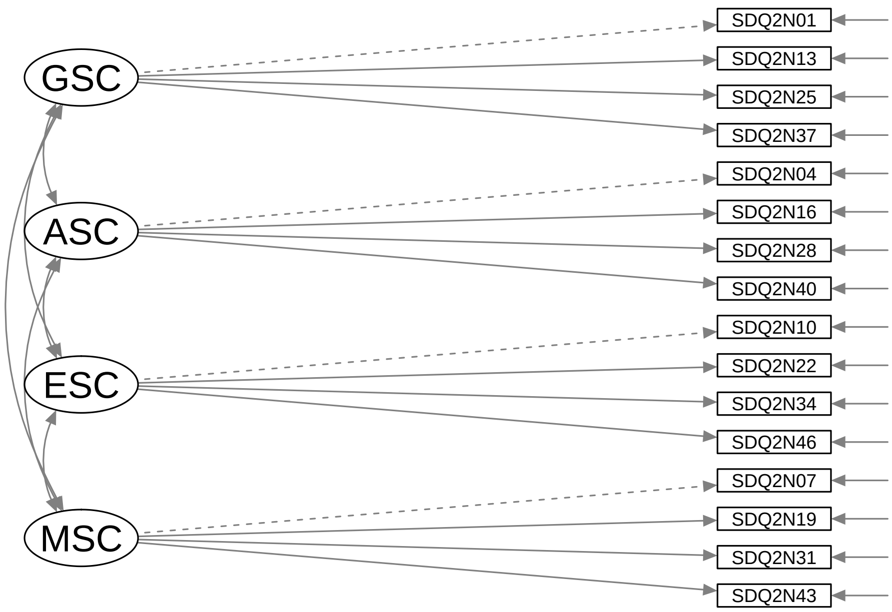
```
library(semPlot) # install.packages("semPlot")

# Path model based on the model object (TRY AND MODIFY!)
semPaths(cfa1) # default options
```




```
# LISREL style (introduced by K. Jöreskog in the 1970s - still the standard):
semPaths(cfa1, style = "lisrel")
```




```r
# "some" more options added:
semPaths(cfa1, style = "lisrel", layout = "tree2", what = "path", whatLabels = "name",
    intercepts = FALSE, residuals = TRUE, thresholds = FALSE, reorder = FALSE,
    rotation = 2,
    latents = c("MSC","ESC", "ASC" ,"GSC"),
    sizeLat = 10, sizeLat2 = 5,
    manifests = rev(colnames(SCdata)),
    sizeMan = 10, sizeMan2 = 2
    )
```



Exercise 2.2

Specify and test these two additional hypotheses (again draw conclusions based on the χ^2 statistic and the CFI, TLI, RMSEA, and SRMR indices):

- *Hypothesis 2:* SC is a two-factor structure consisting of GSC and ASC (so that the four GSC measures load onto the GSC and all other onto the ASC).

```
model2 <- '# two-factor solution:
           GSC =~ SDQ2N01 + SDQ2N13 + SDQ2N25 + SDQ2N37
           ASC =~ SDQ2N04 + SDQ2N16 + SDQ2N28 + SDQ2N40 + SDQ2N10 + SDQ2N22 + SDQ2N34 + SDQ2N46 + SDQ2N0
          '

# Estimate the model using the data defined earlier:
cfa2 <- cfa(model2, data = SCdata)

# Numerical summary of the model:
summary(cfa2, fit.measures = TRUE, standardized = TRUE)
```

```
## lavaan 0.6.13 ended normally after 38 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        33
```

```
##
##   Number of observations                           265
##
## Model Test User Model:
##
##   Test statistic                               457.653
##   Degrees of freedom                               103
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              1703.155
##   Degrees of freedom                               120
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.776
##   Tucker-Lewis Index (TLI)                       0.739
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -6711.949
##   Loglikelihood unrestricted model (H1)      -6483.122
##
##   Akaike (AIC)                               13489.897
##   Bayesian (BIC)                             13608.028
##   Sample-size adjusted Bayesian (SABIC)      13503.401
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.114
##   90 Percent confidence interval - lower         0.103
##   90 Percent confidence interval - upper         0.125
##   P-value H_0: RMSEA <= 0.050                    0.000
##   P-value H_0: RMSEA >= 0.080                    1.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.101
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   GSC =~
##     SDQ2N01           1.000                               0.801    0.595
##     SDQ2N13           1.048    0.151    6.930    0.000    0.839    0.619
##     SDQ2N25           0.860    0.131    6.542    0.000    0.688    0.562
##     SDQ2N37           0.890    0.128    6.957    0.000    0.712    0.623
```

```
##   ASC =~
##     SDQ2N04          1.000                               0.679    0.485
##     SDQ2N16          1.263    0.170    7.440    0.000    0.857    0.692
##     SDQ2N28          1.276    0.177    7.221    0.000    0.866    0.651
##     SDQ2N40          1.235    0.176    7.026    0.000    0.838    0.618
##     SDQ2N10          0.581    0.123    4.736    0.000    0.394    0.343
##     SDQ2N22          0.558    0.117    4.786    0.000    0.378    0.348
##     SDQ2N34          0.065    0.161    0.406    0.685    0.044    0.026
##     SDQ2N46          0.514    0.132    3.885    0.000    0.349    0.270
##     SDQ2N07          2.069    0.262    7.885    0.000    1.404    0.790
##     SDQ2N19          1.871    0.242    7.721    0.000    1.270    0.751
##     SDQ2N31          2.021    0.247    8.192    0.000    1.372    0.875
##     SDQ2N43          1.442    0.193    7.481    0.000    0.979    0.700
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   GSC ~~
##     ASC              0.340    0.068    4.975    0.000    0.626    0.626
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .SDQ2N01          1.170    0.127    9.216    0.000    1.170    0.646
##    .SDQ2N13          1.134    0.127    8.906    0.000    1.134    0.617
##    .SDQ2N25          1.026    0.107    9.582    0.000    1.026    0.684
##    .SDQ2N37          0.799    0.090    8.842    0.000    0.799    0.612
##    .SDQ2N04          1.495    0.134   11.171    0.000    1.495    0.764
##    .SDQ2N16          0.799    0.076   10.490    0.000    0.799    0.521
##    .SDQ2N28          1.018    0.095   10.695    0.000    1.018    0.576
##    .SDQ2N40          1.138    0.105   10.828    0.000    1.138    0.618
##    .SDQ2N10          1.166    0.103   11.364    0.000    1.166    0.882
##    .SDQ2N22          1.043    0.092   11.360    0.000    1.043    0.879
##    .SDQ2N34          2.888    0.251   11.510    0.000    2.888    0.999
##    .SDQ2N46          1.554    0.136   11.425    0.000    1.554    0.927
##    .SDQ2N07          1.191    0.123    9.654    0.000    1.191    0.377
##    .SDQ2N19          1.247    0.124   10.067    0.000    1.247    0.436
##    .SDQ2N31          0.575    0.073    7.852    0.000    0.575    0.234
##    .SDQ2N43          0.996    0.095   10.442    0.000    0.996    0.510
##     GSC              0.641    0.142    4.508    0.000    1.000    1.000
##     ASC              0.461    0.114    4.034    0.000    1.000    1.000
```

First, model fit indices show a poor fit, $\chi^2(103) = 457.65$, p $< .001$, CFI and TLI are unacceptably low ($<$ .80), RMSEA and SRMR are greater than .10. Also, the factor loadings of multiple items in ASC show very low and even insignificant loadings. And these numerous items with lower loadings also demonstrate very high (nearly 1.00!!) residual variances (especially the items SDQ2N10, N22, N34, N46 that are intended to assess ESC). Also, from the previous 4-factor model we saw that ASC and MSC latent factors correlate rather highly with each other, while ESC and MSC correlate rather weakly with each other).

- *Hypothesis 3:* SC is a one-factor structure.

```
model3 <- '# single-factor solution:
            SC =~ SDQ2N01 + SDQ2N13 + SDQ2N25 + SDQ2N37 + SDQ2N04 + SDQ2N16 + SDQ2N28 + SDQ2N40 + SDQ2N1
           '
```

```r
# Estimate the model using the data defined earlier:
cfa3 <- cfa(model3, data = SCdata)

# Numerical summary of the model:
summary(cfa3, fit.measures = TRUE, standardized = TRUE)
```

```
## lavaan 0.6.13 ended normally after 43 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        32
##
##   Number of observations                           265
##
## Model Test User Model:
##
##   Test statistic                               531.918
##   Degrees of freedom                               104
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              1703.155
##   Degrees of freedom                               120
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.730
##   Tucker-Lewis Index (TLI)                       0.688
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -6749.081
##   Loglikelihood unrestricted model (H1)      -6483.122
##
##   Akaike (AIC)                               13562.162
##   Bayesian (BIC)                             13676.713
##   Sample-size adjusted Bayesian (SABIC)      13575.256
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.125
##   90 Percent confidence interval - lower         0.114
##   90 Percent confidence interval - upper         0.135
##   P-value H_0: RMSEA <= 0.050                    0.000
##   P-value H_0: RMSEA >= 0.080                    1.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.104
##
## Parameter Estimates:
```

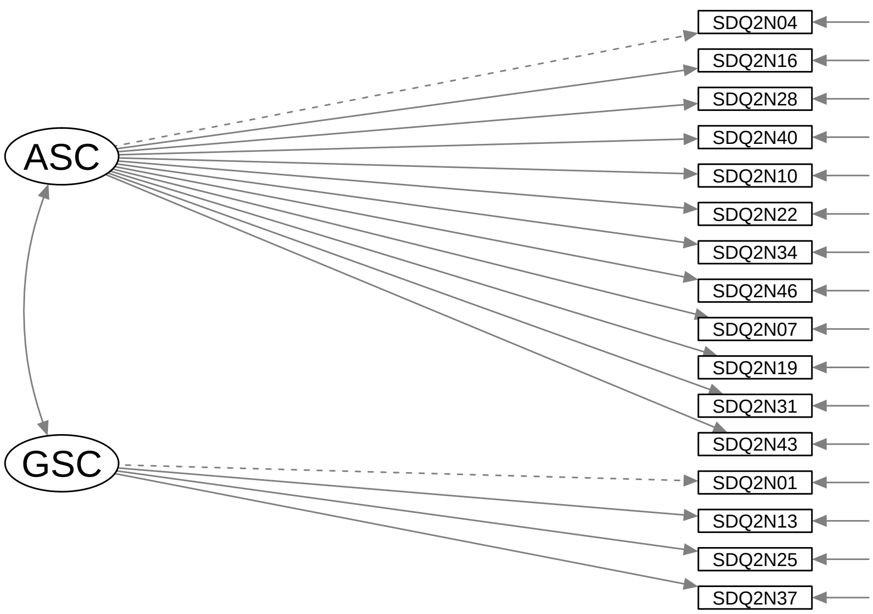
```
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   SC =~
##     SDQ2N01           1.000                               0.496    0.368
##     SDQ2N13           1.158    0.247    4.690    0.000    0.574    0.423
##     SDQ2N25           0.903    0.209    4.330    0.000    0.448    0.366
##     SDQ2N37           1.126    0.224    5.018    0.000    0.558    0.489
##     SDQ2N04           1.407    0.278    5.063    0.000    0.698    0.499
##     SDQ2N16           1.772    0.310    5.716    0.000    0.878    0.709
##     SDQ2N28           1.775    0.317    5.605    0.000    0.880    0.662
##     SDQ2N40           1.744    0.315    5.541    0.000    0.865    0.637
##     SDQ2N10           0.859    0.197    4.362    0.000    0.426    0.370
##     SDQ2N22           0.816    0.187    4.371    0.000    0.405    0.372
##     SDQ2N34           0.181    0.222    0.815    0.415    0.090    0.053
##     SDQ2N46           0.756    0.202    3.732    0.000    0.375    0.289
##     SDQ2N07           2.743    0.471    5.826    0.000    1.360    0.765
##     SDQ2N19           2.505    0.434    5.768    0.000    1.242    0.735
##     SDQ2N31           2.711    0.454    5.970    0.000    1.344    0.857
##     SDQ2N43           1.929    0.341    5.659    0.000    0.956    0.684
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .SDQ2N01           1.565    0.138   11.335    0.000    1.565    0.864
##    .SDQ2N13           1.508    0.134   11.266    0.000    1.508    0.821
##    .SDQ2N25           1.299    0.115   11.338    0.000    1.299    0.866
##    .SDQ2N37           0.994    0.089   11.160    0.000    0.994    0.761
##    .SDQ2N04           1.469    0.132   11.140    0.000    1.469    0.751
##    .SDQ2N16           0.762    0.073   10.368    0.000    0.762    0.497
##    .SDQ2N28           0.994    0.093   10.633    0.000    0.994    0.562
##    .SDQ2N40           1.093    0.102   10.742    0.000    1.093    0.594
##    .SDQ2N10           1.140    0.101   11.333    0.000    1.140    0.863
##    .SDQ2N22           1.022    0.090   11.332    0.000    1.022    0.862
##    .SDQ2N34           2.882    0.250   11.508    0.000    2.882    0.997
##    .SDQ2N46           1.535    0.135   11.409    0.000    1.535    0.916
##    .SDQ2N07           1.311    0.132    9.913    0.000    1.311    0.415
##    .SDQ2N19           1.316    0.129   10.186    0.000    1.316    0.460
##    .SDQ2N31           0.650    0.078    8.367    0.000    0.650    0.265
##    .SDQ2N43           1.040    0.099   10.520    0.000    1.040    0.532
##     SC                0.246    0.083    2.972    0.003    1.000    1.000
```

And the fit of this model is even poorer than that of the previous model. CFI and TLI are even lower and RMSEA and SRMR even higher. Although the χ^2 test was significant even in the first well-fitting 4-factor model (probably resultsing from a large sample size), we can see that the worse the model fit gets, the greater also the ratio between the χ^2 test statistic and the degrees of freedom (4-factor model: 159/98, 2-factor model: 458/103, 1-factor model: 532/104).

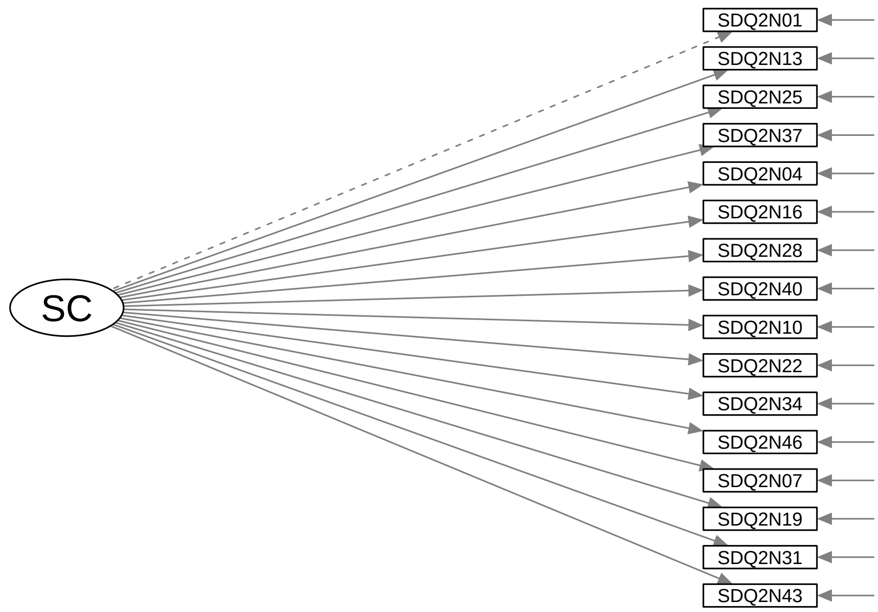
As to parameter estimates, the items intended to measure GSC and ESC have especially high residual variance and low or even insignificant factor loadings. There seems to be something more mutual in the items intended to assess ASC and MSC - it is plausible that participants consider math skills as an important indicator of academic skills and therefore also ASC and MSC items have greater variance in common when compared to GSC and ESC items.

Visualize the models and compare them with the four-factor model analyzed in Exercise 2.1.

```
semPaths(cfa2, style = "lisrel", layout = "tree2", what = "path", whatLabels = "name",
    intercepts = FALSE, residuals = TRUE, thresholds = FALSE, reorder = FALSE,
    rotation = 2,
    latents = c("GSC","ASC"),
    sizeLat = 10, sizeLat2 = 5,
    manifests = rev(colnames(SCdata)),
    sizeMan = 10, sizeMan2 = 2
    )
```




```
semPaths(cfa3, style = "lisrel", layout = "tree2", what = "path", whatLabels = "name",
    intercepts = FALSE, residuals = TRUE, thresholds = FALSE, reorder = FALSE,
    rotation = 2,
    latents = c("SC"),
    sizeLat = 10, sizeLat2 = 5,
    manifests = rev(colnames(SCdata)),
    sizeMan = 10, sizeMan2 = 2
    )
```


SC
SDQ2N01
SDQ2N13
SDQ2N25
SDQ2N37
SDQ2N04
SDQ2N16
SDQ2N28
SDQ2N40
SDQ2N10
SDQ2N22
SDQ2N34
SDQ2N46
SDQ2N07
SDQ2N19
SDQ2N31
SDQ2N43