Structural Equation Models 2019 / WEEK 3

Mikko Patronen

05 February, 2019

Exercise 3.1

The null hypothesis is that burnout is a multidimensional construct composed of three factors:

- Emotional Exhaustion (EE)
- Depersonalization (DP)
- Personal Accomplishment (PA)

Alternative hypothesis is that burnout is not a multidimensional construct composed of three factors.

Let us bring the data in R and prepare it for analysis:

```
library(psych)
```

```
## Warning: package 'psych' was built under R version 3.5.2
```

```
library(likert)
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'ggplot2'
```

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

```
## Loading required package: xtable
```

```
library(lavaan)
```

```
## This is lavaan 0.6-3
```

```
## lavaan is BETA software! Please report any bugs.
```

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

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

```
viikko3 <- read.fortran(file = "ELEMM1.DAT", c("22F1.0"))
colnames(viikko3) <- c("ITEM1", "ITEM2", "ITEM3", "ITEM4", "ITEM5",
                   "ITEM6", "ITEM7", "ITEM8", "ITEM9", "ITEM10",
                   "ITEM11", "ITEM12", "ITEM13", "ITEM14",
                   "ITEM15", "ITEM16", "ITEM17", "ITEM18",
                   "ITEM19", "ITEM20", "ITEM21", "ITEM22")
```

Then let us do some EDA to see whether the items are normally distributed:

```
describe(viikko3)
```

```
##        vars   n mean   sd median trimmed  mad min max range  skew kurtosis
## ITEM1     1 372 4.37 1.66    4.0    4.36 2.97   1   7     6 -0.11    -1.17
```

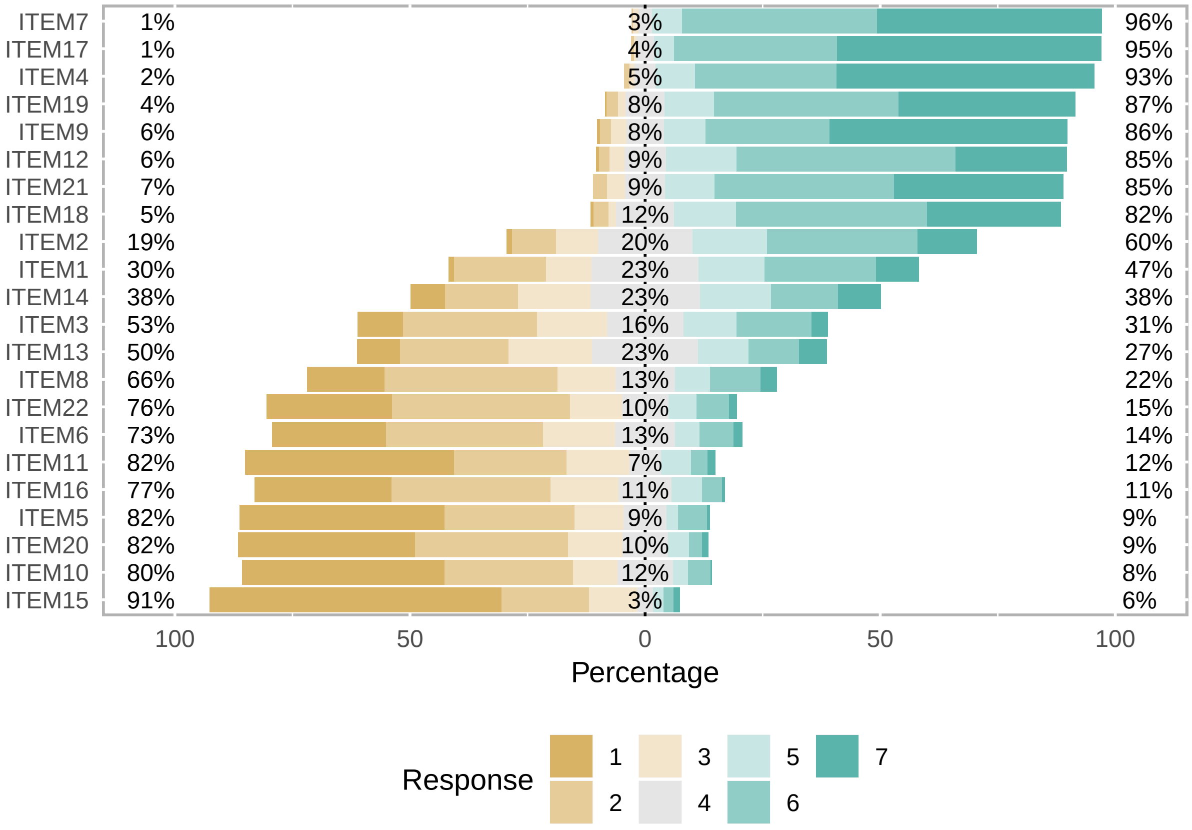
```
## ITEM2      2 372 4.87 1.55    5.0    4.97 1.48   1   7     6 -0.50    -0.71
## ITEM3      3 372 3.53 1.73    3.0    3.49 1.48   1   7     6  0.32    -1.11
## ITEM4      4 372 6.30 1.00    7.0    6.50 0.00   2   7     5 -1.80     3.63
## ITEM5      5 372 2.20 1.49    2.0    1.92 1.48   1   7     6  1.32     0.91
## ITEM6      6 372 2.71 1.58    2.0    2.50 1.48   1   7     6  0.92    -0.01
## ITEM7      7 372 6.31 0.84    6.0    6.46 1.48   2   7     5 -1.64     3.77
## ITEM8      8 372 3.04 1.73    2.0    2.89 1.48   1   7     6  0.74    -0.61
## ITEM9      9 372 6.03 1.32    7.0    6.29 0.00   1   7     6 -1.54     1.84
## ITEM10    10 372 2.20 1.45    2.0    1.96 1.48   1   7     6  1.20     0.56
## ITEM11    11 372 2.24 1.53    2.0    1.97 1.48   1   7     6  1.27     0.80
## ITEM12    12 372 5.70 1.19    6.0    5.86 1.48   1   7     6 -1.31     1.84
## ITEM13    13 372 3.59 1.68    3.5    3.52 2.22   1   7     6  0.35    -0.79
## ITEM14    14 372 4.03 1.73    4.0    4.01 1.48   1   7     6  0.03    -0.94
## ITEM15    15 372 1.77 1.30    1.0    1.47 0.00   1   7     6  2.09     4.24
## ITEM16    16 372 2.47 1.44    2.0    2.28 1.48   1   7     6  0.97     0.16
## ITEM17    17 372 6.41 0.85    7.0    6.58 0.00   2   7     5 -1.97     5.06
## ITEM18    18 372 5.70 1.27    6.0    5.87 1.48   1   7     6 -1.23     1.34
## ITEM19    19 372 5.95 1.19    6.0    6.15 1.48   1   7     6 -1.48     2.21
## ITEM20    20 372 2.24 1.41    2.0    2.01 1.48   1   7     6  1.29     1.17
## ITEM21    21 372 5.85 1.27    6.0    6.06 1.48   2   7     5 -1.29     1.16
## ITEM22    22 372 2.58 1.58    2.0    2.35 1.48   1   7     6  1.06     0.18
##          se
## ITEM1  0.09
## ITEM2  0.08
## ITEM3  0.09
## ITEM4  0.05
## ITEM5  0.08
## ITEM6  0.08
## ITEM7  0.04
## ITEM8  0.09
## ITEM9  0.07
## ITEM10 0.08
## ITEM11 0.08
## ITEM12 0.06
## ITEM13 0.09
## ITEM14 0.09
## ITEM15 0.07
## ITEM16 0.07
## ITEM17 0.04
## ITEM18 0.07
## ITEM19 0.06
## ITEM20 0.07
## ITEM21 0.07
## ITEM22 0.08
```

The skewness and kurtosis values for basicly all the items indicate that none of the items are normally distributed and therefore the maximum likelihood estimators are not suited for the task. Let us visualize the distributions of the items:

```
vk3 <- viikko3

vk3$ITEM1 = factor(vk3$ITEM1,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM2 = factor(vk3$ITEM2,
```

```r
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM3 = factor(vk3$ITEM3,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM4 = factor(vk3$ITEM4,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM5 = factor(vk3$ITEM5,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM6 = factor(vk3$ITEM6,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM7 = factor(vk3$ITEM7,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM8 = factor(vk3$ITEM8,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM9 = factor(vk3$ITEM9,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM10 = factor(vk3$ITEM10,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM11 = factor(vk3$ITEM11,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM12 = factor(vk3$ITEM12,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM13 = factor(vk3$ITEM13,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM14 = factor(vk3$ITEM14,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM15 = factor(vk3$ITEM15,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM16 = factor(vk3$ITEM16,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM17 = factor(vk3$ITEM17,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM18 = factor(vk3$ITEM18,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
vk3$ITEM19 = factor(vk3$ITEM19,
                   levels = c("1", "2", "3", "4", "5", "6", "7"),
                   ordered = TRUE)
```

```
vk3$ITEM20 = factor(vk3$ITEM20,
                    levels = c("1", "2", "3", "4", "5", "6", "7"),
                    ordered = TRUE)
vk3$ITEM21 = factor(vk3$ITEM21,
                    levels = c("1", "2", "3", "4", "5", "6", "7"),
                    ordered = TRUE)
vk3$ITEM22 = factor(vk3$ITEM22,
                    levels = c("1", "2", "3", "4", "5", "6", "7"),
                    ordered = TRUE)


Result = likert(vk3)

plot(Result,
     type="bar")
```



The visualization also shows that the items are not normally distributed.

ML estimator

Then let us specify the model according to instructions and visualize the model structure using ML estimator:

```
model_MBI <- "
EE =~ ITEM1 + ITEM2 + ITEM3 + ITEM6 + ITEM8 + ITEM13 + ITEM14 + ITEM16 + ITEM20
DP =~ ITEM5 + ITEM10 + ITEM11 + ITEM15 + ITEM22
PA =~ ITEM4 + ITEM7 + ITEM9 + ITEM12 + ITEM17 + ITEM18 + ITEM19 + ITEM21
"
```

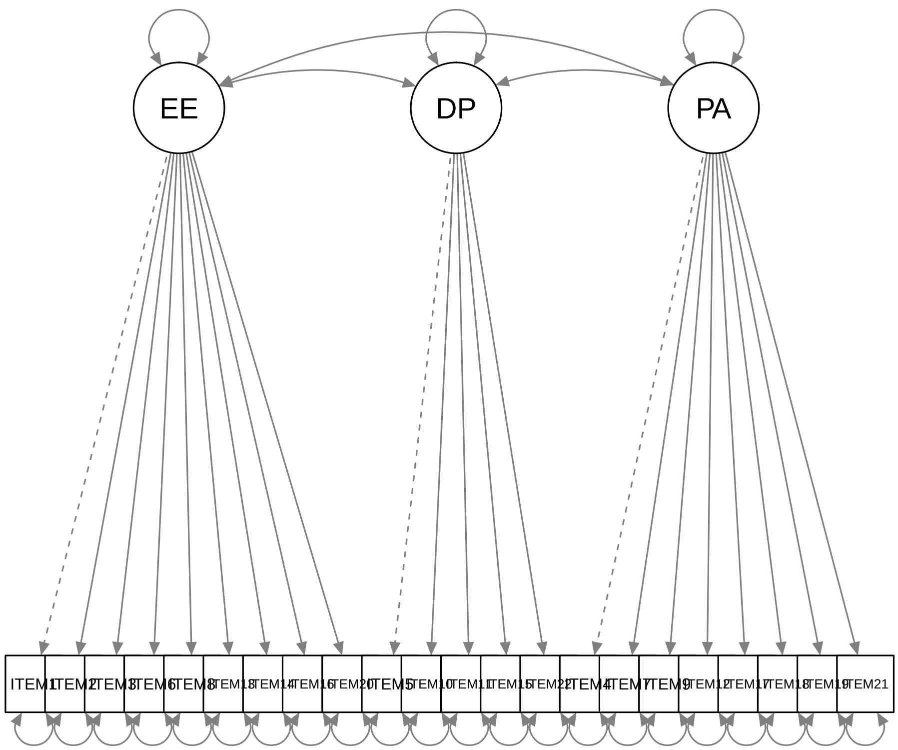
```r
fit_ML <- cfa(model_MBI, data = viikko3)
summary(fit_ML, fit.measures = T)
```

```
## lavaan 0.6-3 ended normally after 46 iterations
##
##   Optimization method                           NLMINB
##   Number of free parameters                         47
##
##   Number of observations                           372
##
##   Estimator                                         ML
##   Model Fit Test Statistic                     695.719
##   Degrees of freedom                               206
##   P-value (Chi-square)                           0.000
##
## Model test baseline model:
##
##   Minimum Function Test Statistic             3452.269
##   Degrees of freedom                               231
##   P-value                                        0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                    0.848
##   Tucker-Lewis Index (TLI)                       0.830
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -12811.043
##   Loglikelihood unrestricted model (H1)     -12463.184
##
##   Number of free parameters                         47
##   Akaike (AIC)                               25716.087
##   Bayesian (BIC)                             25900.275
##   Sample-size adjusted Bayesian (BIC)        25751.158
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.080
##   90 Percent Confidence Interval          0.073  0.087
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.073
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
```

```
##   EE =~
##     ITEM1             1.000
##     ITEM2             0.887    0.061   14.621    0.000
##     ITEM3             1.021    0.068   15.085    0.000
##     ITEM6             0.764    0.064   12.013    0.000
##     ITEM8             1.143    0.066   17.299    0.000
##     ITEM13            1.017    0.065   15.544    0.000
##     ITEM14            0.848    0.069   12.251    0.000
##     ITEM16            0.715    0.058   12.410    0.000
##     ITEM20            0.753    0.056   13.410    0.000
##   DP =~
##     ITEM5             1.000
##     ITEM10            1.142    0.127    8.986    0.000
##     ITEM11            1.353    0.142    9.511    0.000
##     ITEM15            0.905    0.109    8.318    0.000
##     ITEM22            0.768    0.121    6.361    0.000
##   PA =~
##     ITEM4             1.000
##     ITEM7             0.970    0.150    6.482    0.000
##     ITEM9             1.780    0.254    7.007    0.000
##     ITEM12            1.499    0.221    6.769    0.000
##     ITEM17            1.348    0.181    7.463    0.000
##     ITEM18            1.918    0.262    7.329    0.000
##     ITEM19            1.716    0.238    7.205    0.000
##     ITEM21            1.356    0.218    6.219    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   EE ~~
##     DP                0.701    0.099    7.061    0.000
##     PA               -0.192    0.042   -4.537    0.000
##   DP ~~
##     PA               -0.172    0.035   -4.850    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .ITEM1             1.128    0.095   11.861    0.000
##    .ITEM2             1.105    0.090   12.214    0.000
##    .ITEM3             1.301    0.108   12.031    0.000
##    .ITEM6             1.553    0.121   12.888    0.000
##    .ITEM8             0.852    0.081   10.553    0.000
##    .ITEM13            1.142    0.097   11.821    0.000
##    .ITEM14            1.804    0.140   12.844    0.000
##    .ITEM16            1.235    0.096   12.812    0.000
##    .ITEM20            1.075    0.085   12.585    0.000
##    .ITEM5             1.503    0.125   12.026    0.000
##    .ITEM10            1.169    0.107   10.901    0.000
##    .ITEM11            1.044    0.112    9.330    0.000
##    .ITEM15            1.106    0.093   11.838    0.000
##    .ITEM22            2.076    0.160   12.958    0.000
##    .ITEM4             0.802    0.062   12.901    0.000
##    .ITEM7             0.523    0.042   12.572    0.000
##    .ITEM9             1.117    0.093   11.952    0.000
##    .ITEM12            0.987    0.080   12.287    0.000
```

```
##    .ITEM17            0.375    0.035   10.739    0.000
##    .ITEM18            0.909    0.081   11.224    0.000
##    .ITEM19            0.844    0.073   11.557    0.000
##    .ITEM21            1.245    0.098   12.764    0.000
##     EE                1.625    0.190    8.551    0.000
##     DP                0.705    0.132    5.321    0.000
##     PA                0.193    0.048    4.047    0.000
```

```r
library(semPlot)
semPaths(fit_ML, layout='tree2')
```



The hypothesis that burnout has three factors is not supported by the results (chi square statistic = 695.719, p = 0.000), which suggest that the fit of the data to the model is not adequate and null hypothesis should be rejected. Also the indices CFI (0.848), TLI (0.83) support against hypothesis. However, RMSEA (0.08) modestly supports the null hypothesis.

MLM estimator

Then let us test the model fit with MLM estimator:

```r
fit_MLM <- cfa(model_MBI, data = viikko3, estimator="MLM")
summary(fit_MLM, fit.measures = T)
```

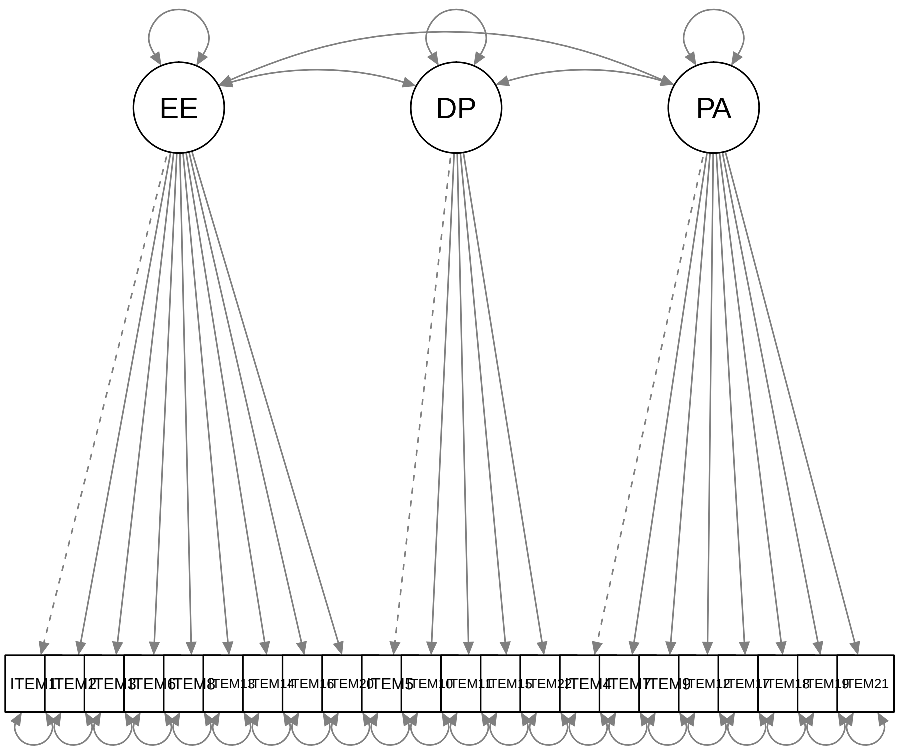
```
## lavaan 0.6-3 ended normally after 46 iterations
##
##   Optimization method                           NLMINB
##   Number of free parameters                         47
##
##   Number of observations                           372
##
##   Estimator                                         ML      Robust
```

```
##   Model Fit Test Statistic                     695.719     567.753
##   Degrees of freedom                               206         206
##   P-value (Chi-square)                           0.000       0.000
##   Scaling correction factor                                  1.225
##     for the Satorra-Bentler correction
##
## Model test baseline model:
##
##   Minimum Function Test Statistic             3452.269    2911.466
##   Degrees of freedom                               231         231
##   P-value                                        0.000       0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                    0.848       0.865
##   Tucker-Lewis Index (TLI)                       0.830       0.849
##
##   Robust Comparative Fit Index (CFI)                         0.861
##   Robust Tucker-Lewis Index (TLI)                            0.844
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -12811.043  -12811.043
##   Loglikelihood unrestricted model (H1)     -12463.184  -12463.184
##
##   Number of free parameters                         47          47
##   Akaike (AIC)                               25716.087   25716.087
##   Bayesian (BIC)                             25900.275   25900.275
##   Sample-size adjusted Bayesian (BIC)        25751.158   25751.158
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.080       0.069
##   90 Percent Confidence Interval          0.073  0.087       0.063  0.075
##   P-value RMSEA <= 0.05                          0.000       0.000
##
##   Robust RMSEA                                               0.076
##   90 Percent Confidence Interval                             0.069  0.084
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.073       0.073
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard Errors                           Robust.sem
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   EE =~
##     ITEM1             1.000
##     ITEM2             0.887    0.040   22.391    0.000
```

```
##     ITEM3            1.021    0.053   19.310    0.000
##     ITEM6            0.764    0.070   10.974    0.000
##     ITEM8            1.143    0.059   19.366    0.000
##     ITEM13           1.017    0.062   16.340    0.000
##     ITEM14           0.848    0.058   14.584    0.000
##     ITEM16           0.715    0.066   10.826    0.000
##     ITEM20           0.753    0.061   12.303    0.000
##   DP =~
##     ITEM5            1.000
##     ITEM10           1.142    0.152    7.509    0.000
##     ITEM11           1.353    0.162    8.368    0.000
##     ITEM15           0.905    0.123    7.366    0.000
##     ITEM22           0.768    0.122    6.284    0.000
##   PA =~
##     ITEM4            1.000
##     ITEM7            0.970    0.128    7.563    0.000
##     ITEM9            1.780    0.322    5.529    0.000
##     ITEM12           1.499    0.241    6.232    0.000
##     ITEM17           1.348    0.200    6.757    0.000
##     ITEM18           1.918    0.298    6.435    0.000
##     ITEM19           1.716    0.287    5.978    0.000
##     ITEM21           1.356    0.227    5.984    0.000
##
## Covariances:
##                   Estimate  Std.Err  z-value  P(>|z|)
##   EE ~~
##     DP               0.701    0.106    6.608    0.000
##     PA              -0.192    0.040   -4.796    0.000
##   DP ~~
##     PA              -0.172    0.036   -4.777    0.000
##
## Variances:
##                   Estimate  Std.Err  z-value  P(>|z|)
##    .ITEM1            1.128    0.093   12.177    0.000
##    .ITEM2            1.105    0.088   12.506    0.000
##    .ITEM3            1.301    0.106   12.317    0.000
##    .ITEM6            1.553    0.134   11.550    0.000
##    .ITEM8            0.852    0.082   10.450    0.000
##    .ITEM13           1.142    0.124    9.173    0.000
##    .ITEM14           1.804    0.142   12.730    0.000
##    .ITEM16           1.235    0.110   11.278    0.000
##    .ITEM20           1.075    0.137    7.860    0.000
##    .ITEM5            1.503    0.179    8.381    0.000
##    .ITEM10           1.169    0.147    7.959    0.000
##    .ITEM11           1.044    0.141    7.398    0.000
##    .ITEM15           1.106    0.153    7.220    0.000
##    .ITEM22           2.076    0.184   11.266    0.000
##    .ITEM4            0.802    0.113    7.124    0.000
##    .ITEM7            0.523    0.075    7.010    0.000
##    .ITEM9            1.117    0.149    7.487    0.000
##    .ITEM12           0.987    0.126    7.852    0.000
##    .ITEM17           0.375    0.056    6.635    0.000
##    .ITEM18           0.909    0.143    6.376    0.000
##    .ITEM19           0.844    0.111    7.622    0.000
```

```
##    .ITEM21            1.245    0.133    9.338    0.000
##     EE                1.625    0.148   11.004    0.000
##     DP                0.705    0.158    4.452    0.000
##     PA                0.193    0.050    3.839    0.000
```

```r
semPaths(fit_MLM, layout='tree2')
```



With this estimator the hypothesis that burnout has three factors is also not supported by the results (chi square statistic = 567.753, p = 0.000), which suggest that the fit of the data to the model is not adequate and null hypothesis should be rejected. Here the indices increase small amounts CFI (0.865), TLI (0.849) and RMSEA decreases small amount (0.076) but still these indices support against hypothesis.

Exercise 3.2

Unfortunately I wasted a huge amount of time with solving R error message that had to do with lavaan and sem -packages. The sem -package was causing problems that made R to not “find file” when using cfa. I managed to solve this problem just in time before deadline but unfortunately I did not have enough time to do this second excercise. But I learned a lot of R instead :D