2019R2 STAT6108 Assignment 4

Yiu Chung WONG 1155017920

```
# import data
teachers <- read.csv('hw4(2020).dat', header = FALSE, sep = '')
```

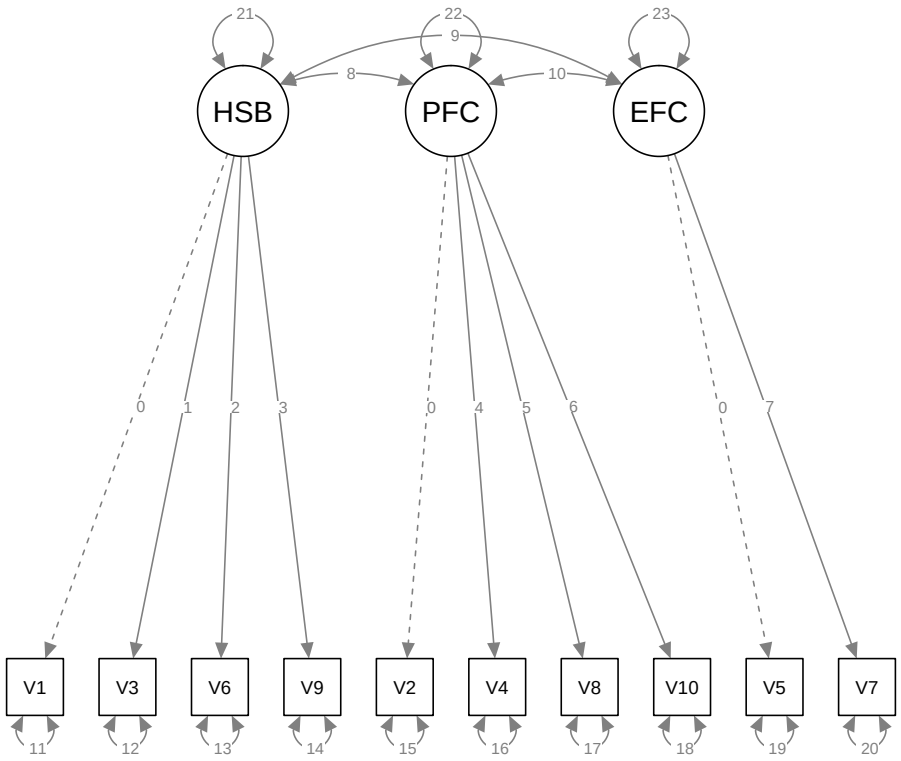
Model Setup

```
model <- 'HelpsSeekingBehavior =~ V1 + V3 + V6 + V9;
ProblemFocusedCoping =~ V2 + V4 + V8 + V10;
EmotionFocusedCoping =~ V5 + V7;
HelpsSeekingBehavior ~~ ProblemFocusedCoping;
HelpsSeekingBehavior ~~ EmotionFocusedCoping;
ProblemFocusedCoping ~~ EmotionFocusedCoping'
```

a)

Path diagram assuming unit loading identification

```
fit <- cfa(model, data = teachers)
semPlot::semPaths(object=fit, intercepts=FALSE, what="path", whatLabels="cons")
```



- Edge labeled 0 indicates the parameter is fixed, non-zero indicates free parameters.
- The factor loading of the first indicator is set to 1.0 for every latent variable to circumvent factor indeterminacy.
- The model satisfy the t-rule because the degrees of freedom is a positive integer.
- Latent factors are correlated, and degree of freedom i sgreater than zero. Hence the model is overidentified.

Two-indicator rule

- There are three factors
- Factor correlations are free
- At least 2 indicators per factor
- Each indicator loads on one factor
- Errors are uncorrelated

The model is indentifiable.

b)

```
n_variable <- 10
n_factor <- 3

p_star <- (n_variable * (n_variable+1))/2
q <- n_variable + (n_variable - n_factor) + (n_factor * (n_factor+1))/2

df <- p_star - q
```

- There are 23 free parameters in the the proposed model.
- The degrees of freedom is 32.

c)

ULI

```
uli <- lavaan(model, data=teachers, auto.var=TRUE, auto.fix.first=TRUE, std.lv=FALSE)
uliSummary <- summary(uli, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
```

UVI

```
uvi <- lavaan(model, data=teachers, auto.var=TRUE, auto.fix.first=FALSE, std.lv=TRUE)
uviSummary <- summary(uvi, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
```

Prepare for comparison

```
# row numbers of fixed parameters in parameter table
uliFixedParameters <- c(1,5,9)
uviFixedParameters <- c(24,25,26)

# round numbers to 3 decimal places
uliSummary$PE[,-c(1:3)] <- round(uliSummary$PE[,-c(1:3)],3)
uviSummary$PE[,-c(1:3)] <- round(uviSummary$PE[,-c(1:3)],3)
```

Parameter estimations that are common in both methods

```
# compare all paremeters except fixed parameters
combined <- rbind(uliSummary$PE, uviSummary$PE)
intersaction <- combined[duplicated(combined), , drop = FALSE]
intersaction[,-which(names(intersaction) == "std.nox")]
```

```
##    lhs op rhs exo   est    se     z pvalue std.lv std.all
## 50  V1 ~~  V1   0 0.211 0.026 8.190  0.000  0.211   0.426
## 51  V3 ~~  V3   0 0.528 0.060 8.759  0.000  0.528   0.485
## 52  V6 ~~  V6   0 0.404 0.043 9.433  0.000  0.404   0.592
## 53  V9 ~~  V9   0 0.172 0.039 4.454  0.000  0.172   0.205
## 54  V2 ~~  V2   0 0.391 0.052 7.479  0.000  0.391   0.469
## 55  V4 ~~  V4   0 0.114 0.036 3.150  0.002  0.114   0.209
## 56  V8 ~~  V8   0 0.321 0.033 9.837  0.000  0.321   0.747
## 57 V10 ~~ V10   0 0.508 0.052 9.853  0.000  0.508   0.751
## 58  V5 ~~  V5   0 0.225 0.073 3.072  0.002  0.225   0.314
## 59  V7 ~~  V7   0 0.534 0.086 6.246  0.000  0.534   0.525
## 63  V1 r2  V1   0 0.574    NA    NA     NA     NA      NA
## 64  V3 r2  V3   0 0.515    NA    NA     NA     NA      NA
## 65  V6 r2  V6   0 0.408    NA    NA     NA     NA      NA
## 66  V9 r2  V9   0 0.795    NA    NA     NA     NA      NA
## 67  V2 r2  V2   0 0.531    NA    NA     NA     NA      NA
## 68  V4 r2  V4   0 0.791    NA    NA     NA     NA      NA
## 69  V8 r2  V8   0 0.253    NA    NA     NA     NA      NA
## 70 V10 r2 V10   0 0.249    NA    NA     NA     NA      NA
## 71  V5 r2  V5   0 0.686    NA    NA     NA     NA      NA
## 72  V7 r2  V7   0 0.475    NA    NA     NA     NA      NA
```

- Error variance and R^2 are the same across two identification methods
- All error variance are sigificant

Parameter estimations that are different

```
nRowTable <- dim(uviSummary$PE)[1]
tail <- tail(duplicated(combined), nRowTable)
uliSummary$PE[!tail,-which(colnames(uliSummary$PE) == "std.nox")]
```

```
##                    lhs op                  rhs exo   est    se      z pvalue std.lv std.all
## 1  HelpsSeekingBehavior =~                   V1   0 1.000 0.000     NA     NA  0.534   0.758
## 2  HelpsSeekingBehavior =~                   V3   0 1.402 0.134 10.434  0.000  0.749   0.718
## 3  HelpsSeekingBehavior =~                   V6   0 0.989 0.107  9.221  0.000  0.528   0.639
## 4  HelpsSeekingBehavior =~                   V9   0 1.533 0.126 12.196  0.000  0.818   0.892
## 5  ProblemFocusedCoping =~                   V2   0 1.000 0.000     NA     NA  0.665   0.729
## 6  ProblemFocusedCoping =~                   V4   0 0.989 0.103  9.607  0.000  0.658   0.889
## 7  ProblemFocusedCoping =~                   V8   0 0.496 0.072  6.881  0.000  0.330   0.503
## 8  ProblemFocusedCoping =~                  V10   0 0.617 0.090  6.829  0.000  0.410   0.499
## 9  EmotionFocusedCoping =~                   V5   0 1.000 0.000     NA     NA  0.702   0.828
## 10 EmotionFocusedCoping =~                   V7   0 0.991 0.157  6.318  0.000  0.695   0.689
## 11 HelpsSeekingBehavior ~~ ProblemFocusedCoping   0 0.029 0.028  1.027  0.305  0.081   0.081
## 12 HelpsSeekingBehavior ~~ EmotionFocusedCoping   0 0.144 0.034  4.204  0.000  0.384   0.384
## 13 ProblemFocusedCoping ~~ EmotionFocusedCoping   0 0.210 0.046  4.586  0.000  0.449   0.449
## 24 HelpsSeekingBehavior ~~ HelpsSeekingBehavior   0 0.285 0.046  6.250  0.000  1.000   1.000
## 25 ProblemFocusedCoping ~~ ProblemFocusedCoping   0 0.442 0.079  5.601  0.000  1.000   1.000
## 26 EmotionFocusedCoping ~~ EmotionFocusedCoping   0 0.492 0.095  5.161  0.000  1.000   1.000
```

```
uviSummary$PE[!tail,-which(names(uviSummary$PE) == "std.nox")]
```

```
##                     lhs op                  rhs exo   est    se      z pvalue std.lv std.all
## 1  HelpsSeekingBehavior =~                   V1   0 0.534 0.043 12.500  0.000  0.534   0.758
## 2  HelpsSeekingBehavior =~                   V3   0 0.749 0.064 11.636  0.000  0.749   0.718
## 3  HelpsSeekingBehavior =~                   V6   0 0.528 0.053 10.026  0.000  0.528   0.639
## 4  HelpsSeekingBehavior =~                   V9   0 0.818 0.052 15.619  0.000  0.818   0.892
## 5  ProblemFocusedCoping =~                   V2   0 0.665 0.059 11.203  0.000  0.665   0.729
## 6  ProblemFocusedCoping =~                   V4   0 0.658 0.047 14.078  0.000  0.658   0.889
## 7  ProblemFocusedCoping =~                   V8   0 0.330 0.045  7.356  0.000  0.330   0.503
## 8  ProblemFocusedCoping =~                  V10   0 0.410 0.056  7.293  0.000  0.410   0.499
## 9  EmotionFocusedCoping =~                   V5   0 0.702 0.068 10.322  0.000  0.702   0.828
## 10 EmotionFocusedCoping =~                   V7   0 0.695 0.077  9.016  0.000  0.695   0.689
## 11 HelpsSeekingBehavior ~~ ProblemFocusedCoping   0 0.081 0.078  1.040  0.298  0.081   0.081
## 12 HelpsSeekingBehavior ~~ EmotionFocusedCoping   0 0.384 0.074  5.182  0.000  0.384   0.384
## 13 ProblemFocusedCoping ~~ EmotionFocusedCoping   0 0.449 0.073  6.178  0.000  0.449   0.449
## 24 HelpsSeekingBehavior ~~ HelpsSeekingBehavior   0 1.000 0.000     NA     NA  1.000   1.000
## 25 ProblemFocusedCoping ~~ ProblemFocusedCoping   0 1.000 0.000     NA     NA  1.000   1.000
## 26 EmotionFocusedCoping ~~ EmotionFocusedCoping   0 1.000 0.000     NA     NA  1.000   1.000
```

- Since the scales are different, the estimates (and their respected standard error) produced by the two identification methods do not match.
- Parameters that are sigificant in uli are also sigificant in uvi; the result of Wald test for sigificance are the same.
- Standardised latent variables and complete standardised solutions are equal.

Compare goodness-of-fit

```
identical(round(uliSummary$FIT, 3), round(uviSummary$FIT, 3))
```

```
## [1] TRUE
```

- Both identification methods yields identical Chi-square goodness-of-fit test, residuals, and other goodness-of-fit indicies.

d)

Goodness-of-fit evaluation

```
uliSummary$FIT
```

```
##             npar             fmin            chisq               df           pvalue
##           23.000            0.523          232.378           32.000            0.000
##    baseline.chisq      baseline.df  baseline.pvalue              cfi              tli
##          956.825           45.000            0.000            0.780            0.691
##             logl unrestricted.logl              aic              bic           ntotal
##        -2398.615        -2282.426         4843.230         4921.492          222.000
##             bic2             rmsea    rmsea.ci.lower    rmsea.ci.upper     rmsea.pvalue
##         4848.603            0.168            0.148            0.189            0.000
##             srmr
##            0.092
```

- H_0: $\Sigma = \Sigma(\theta)$
- Chi-square test statistics: 232.3784319; p-value: 0. H_0 is rejected at $\alpha = .95$.
- NNFI: $0.6909689 < 0.95$
- CFI: $0.7802445 < 0.95$
- RMSEA: $0.1679477 > 0.07$
- SRMR: $0.0921616 > 0.08$

Neither the goodness-of-fit test or the fit indicies pass the acceptable threshold levels. The proposed model is of poor fit.

e)

Modification indicies

```
mi <- modindices(uli, sort. = TRUE)
head(mi)[,-which(names(mi) == "sepc.nox")]
```

```
##    lhs op rhs     mi    epc sepc.lv sepc.all
## 48  V1 ~~  V6 56.721  0.195   0.195    0.668
## 57  V3 ~~  V9 45.750  0.360   0.360    1.193
## 89 V10 ~~  V5 25.264  0.165   0.165    0.488
## 90 V10 ~~  V7 23.206 -0.194  -0.194   -0.373
## 65  V6 ~~  V2 22.734  0.146   0.146    0.367
## 69  V6 ~~  V5 19.658 -0.132  -0.132   -0.438
```

- Modification indicies suggests there exist error covariance between V_1 and V_6, V_3 and V_9.
- This modification is justifiable as these indicators belong to the same latent factor according to the porposes model. Hence they are likely to subject to the same type of variance.

New model

```
newModel <- 'HelpsSeekingBehavior =~ V1 + V3 + V6 + V9;
ProblemFocusedCoping =~ V2 + V4 + V8 + V10;
EmotionFocusedCoping =~ V5 + V7;

HelpsSeekingBehavior ~~ ProblemFocusedCoping;
HelpsSeekingBehavior ~~ EmotionFocusedCoping;
ProblemFocusedCoping ~~ EmotionFocusedCoping;

V1~~V6;
V3~~V9'

uliNew <- lavaan(newModel, data=teachers, auto.var=TRUE, auto.fix.first=TRUE, std.lv=FALSE)
uliNewSummary <- summary(uliNew, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
```

```
## lavaan 0.6-5 ended normally after 36 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         25
##
##   Number of observations                           222
##
## Model Test User Model:
##
##   Test statistic                               175.853
##   Degrees of freedom                                30
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                               956.825
##   Degrees of freedom                                45
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.840
##   Tucker-Lewis Index (TLI)                       0.760
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -2370.352
##   Loglikelihood unrestricted model (H1)      -2282.426
##
##   Akaike (AIC)                                4790.705
##   Bayesian (BIC)                              4875.772
##   Sample-size adjusted Bayesian (BIC)         4796.545
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.148
##   90 Percent confidence interval - lower         0.127
##   90 Percent confidence interval - upper         0.170
```

```
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.086
##
## 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
##   HelpsSeekingBehavior =~
##     V1                       1.000                               0.523    0.743
##     V3                       1.219    0.262    4.647    0.000    0.637    0.611
##     V6                       0.906    0.089   10.218    0.000    0.474    0.574
##     V9                       1.551    0.293    5.299    0.000    0.811    0.884
##   ProblemFocusedCoping =~
##     V2                       1.000                               0.666    0.730
##     V4                       0.984    0.102    9.642    0.000    0.656    0.887
##     V8                       0.494    0.072    6.873    0.000    0.329    0.502
##     V10                      0.618    0.090    6.851    0.000    0.412    0.501
##   EmotionFocusedCoping =~
##     V5                       1.000                               0.723    0.854
##     V7                       0.933    0.147    6.352    0.000    0.675    0.669
##
## Covariances:
##                           Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   HelpsSeekingBehavior ~~
##     ProblmFcsdCpng           0.031    0.028    1.083    0.279    0.088    0.088
##     EmotinFcsdCpng           0.156    0.040    3.880    0.000    0.412    0.412
##   ProblemFocusedCoping ~~
##     EmotinFcsdCpng           0.218    0.046    4.700    0.000    0.452    0.452
##  .V1 ~~
##    .V6                       0.146    0.051    2.877    0.004    0.146    0.456
##  .V3 ~~
##    .V9                       0.148    0.101    1.455    0.146    0.148    0.417
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .V1                0.222    0.053    4.186    0.000    0.222    0.449
##    .V3                0.682    0.111    6.145    0.000    0.682    0.627
##    .V6                0.458    0.062    7.438    0.000    0.458    0.671
##    .V9                0.184    0.118    1.555    0.120    0.184    0.219
##    .V2                0.389    0.052    7.461    0.000    0.389    0.467
##    .V4                0.117    0.036    3.239    0.001    0.117    0.213
##    .V8                0.322    0.033    9.835    0.000    0.322    0.748
##    .V10               0.507    0.052    9.842    0.000    0.507    0.749
##    .V5                0.195    0.076    2.578    0.010    0.195    0.271
##    .V7                0.563    0.083    6.765    0.000    0.563    0.553
##     HelpsSekngBhvr    0.273    0.064    4.247    0.000    1.000    1.000
##     ProblmFcsdCpng    0.444    0.079    5.619    0.000    1.000    1.000
```

```
##     EmotinFcsdCpng    0.523    0.098    5.320    0.000    1.000    1.000
##
## R-Square:
##                     Estimate
##     V1                 0.551
##     V3                 0.373
##     V6                 0.329
##     V9                 0.781
##     V2                 0.533
##     V4                 0.787
##     V8                 0.252
##     V10                0.251
##     V5                 0.729
##     V7                 0.447
```

Likelihood ratio test

```
lmtest::lrtest(uli, uliNew)
```

```
## Likelihood ratio test
##
## Model 1: uli
## Model 2: uliNew
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1  23 -2398.6
## 2  25 -2370.3  2 56.525  5.317e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Likelihood ratio test reveals sigificant difference between models with and without error covariance between V_1 and V_6, V_3 and V_9.
- Hence the new model is a better moder.

Goodness-of-fit of the new model

- H_0: $\Sigma = \Sigma(\theta)$
- Chi-square test statistics: 175.8530083; p-value: 0. H_0 is rejected at $\alpha = .95$.
- NNFI: $0.760064 < 0.95$
- CFI: $0.8400427 < 0.95$
- RMSEA: $0.147986 > 0.07$
- SRMR: $0.0864558 > 0.08$

The new model has better commonly used goodness-of-fit indicies (closer to cutoff) across all measures. However, non of which passes the recommended cutoff. Chi-square goodness-of-fit test has a lower test statistics but still rejects H_0.

Finally, only one of the newly added parameter, the error covariance between V_1 and V_6 has a significant value. This suggests the error covariance between V_3 and V_9 might be spurious and can be omitted afterall.

One could endlessly add new parameter according to the modification indicies until a non-significant Chi-square test statistic is reached. This is highly data driven and defeats much the purpose of the analysis.

In summary, the poposed model may not be a good fit.