Answers to exercises week 3 Latent Variable Models

```
library("lavaan")
```

Exercise 4.1

We read in the data:

```
WISC.cor <- lav_matrix_lower2full(c(
  1.00,
  0.31, 1.00,
  0.36, 0.40, 1.00,
  0.51, 0.46, 0.45, 1.00,
  0.29, 0.40, 0.33, 0.43, 1.00,
  0.39, 0.29, 0.27, 0.36, 0.33, 1.00,
  0.32, 0.27, 0.29, 0.33, 0.24, 0.28, 1.00,
  0.22, 0.32, 0.15, 0.22, 0.27, 0.12, 0.26, 1.00
))
WISC.means <- c(7.83, 5.50, 5.67, 21.50, 7.67, 8.00, 6.50, 34.83)
WISC.sds <- c(2.69, 1.50, 2.36, 6.06, 1.85, 2.18, 5.97, 9.94)
WISC.cov <- cor2cov(WISC.cor, sds=WISC.sds)

WISCIV.cor <- lav_matrix_lower2full(c(
  1.00,
  0.46, 1.00,
  0.58, 0.55, 1.00,
  0.63, 0.43, 0.73, 1.00,
  0.27, 0.51, 0.37, 0.33, 1.00,
  0.45, 0.38, 0.37, 0.43, 0.13, 1.00,
  0.33, 0.52, 0.49, 0.41, 0.29, 0.43, 1.00,
  0.15, 0.27, 0.16, 0.09, 0.12, 0.25, 0.23, 1.00
))
WISCIV.means <- c(15.17, 15.00, 11.83, 21.67, 12.17, 17.83, 18.67, 45.83)
WISCIV.sds <- c(4.93, 4.10, 5.20, 6.54, 2.72, 5.35, 9.36, 10.44)
WISCIV.cov <- cor2cov(WISCIV.cor, sds=WISCIV.sds)

WISC.names <- c("Compr", "Arith", "Simil", "Vocab", "DigSpan", "PictCompl",
                "BlockDes", "Cod")

names(WISC.means) <- names(WISCIV.sds) <- names(WISC.sds) <-
  names(WISCIV.sds) <- rownames(WISC.cov) <- colnames(WISC.cov) <-
  rownames(WISCIV.cov) <- colnames(WISCIV.cov) <- WISC.names

WISC.cov.list <- list(WISC.cov, WISCIV.cov)
WISC.mean.list <- list(WISC.means, WISCIV.means)
WISC.n.list <- list(WISC.n = 200, WISCIV.n = 200)
```

We specify the model:

```
WISC.mod <- '
  # verbal comprehension
  VC =~ Simil + Vocab + Compr
  # perceptual reasoning
  PR =~ PictCompl + BlockDes
  # working memory
  WM =~ Arith + DigSpan + Cod
  # Fix intercepts of marker indicators to 0:
  Simil ~ 0*1
  PictCompl ~ 0*1
  Arith ~ 0*1
  # Set means of latent variables free:
  VC ~ NA*1
  PR ~ NA*1
  WM ~ NA*1
'
```

a) We first test for configural invariance:

```
WISC.conf.fit <- cfa(WISC.mod, sample.cov = WISC.cov.list,
                     sample.mean = WISC.mean.list,
                     sample.nobs = WISC.n.list)
summary(WISC.conf.fit, standardized = TRUE)
```

```
## lavaan 0.6-12 ended normally after 255 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        54
##
##   Number of observations per group:
##     Group 1                                        200
##     Group 2                                        200
##
## Model Test User Model:
##
##   Test statistic                                62.187
##   Degrees of freedom                                34
##   P-value (Chi-square)                           0.002
##   Test statistic for each group:
##     Group 1                                     18.998
##     Group 2                                     43.189
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
##
## Group 1 [Group 1]:
```

```
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   VC =~
##     Simil             1.000                               1.393    0.592
##     Vocab             3.346    0.457    7.328    0.000    4.659    0.771
##     Compr             1.238    0.185    6.693    0.000    1.725    0.643
##   PR =~
##     PictCompl         1.000                               1.198    0.551
##     BlockDes          2.526    0.498    5.077    0.000    3.026    0.508
##   WM =~
##     Arith             1.000                               1.018    0.681
##     DigSpan           1.126    0.171    6.568    0.000    1.147    0.621
##     Cod               4.049    0.847    4.780    0.000    4.123    0.416
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   VC ~~
##     PR                1.513    0.310    4.876    0.000    0.907    0.907
##     WM                1.187    0.224    5.287    0.000    0.837    0.837
##   PR ~~
##     WM                1.000    0.211    4.751    0.000    0.820    0.820
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .Simil             0.000                               0.000    0.000
##    .PictCompl         0.000                               0.000    0.000
##    .Arith             0.000                               0.000    0.000
##     VC                5.670    0.166   34.062    0.000    4.071    4.071
##     PR                8.000    0.154   52.028    0.000    6.677    6.677
##     WM                5.500    0.106   51.985    0.000    5.401    5.401
##    .Vocab             2.531    2.641    0.958    0.338    2.531    0.419
##    .Compr             0.808    1.072    0.754    0.451    0.808    0.301
##    .BlockDes        -13.707    4.010   -3.418    0.001  -13.707   -2.302
##    .DigSpan           1.476    0.953    1.550    0.121    1.476    0.800
##    .Cod              12.562    4.713    2.665    0.008   12.562    1.267
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .Simil             3.602    0.419    8.603    0.000    3.602    0.650
##    .Vocab            14.833    2.454    6.044    0.000   14.833    0.406
##    .Compr             4.226    0.518    8.151    0.000    4.226    0.587
##    .PictCompl         3.293    0.451    7.308    0.000    3.293    0.696
##    .BlockDes         26.304    3.282    8.015    0.000   26.304    0.742
##    .Arith             1.202    0.179    6.722    0.000    1.202    0.537
##    .DigSpan           2.090    0.272    7.691    0.000    2.090    0.614
##    .Cod              81.310    8.768    9.273    0.000   81.310    0.827
##     VC                1.939    0.472    4.108    0.000    1.000    1.000
##     PR                1.436    0.458    3.134    0.002    1.000    1.000
##     WM                1.037    0.231    4.496    0.000    1.000    1.000
##
##
## Group 2 [Group 2]:
##
```

```
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   VC =~
##     Simil             1.000                               4.460    0.860
##     Vocab             1.233    0.094   13.137    0.000    5.499    0.843
##     Compr             0.788    0.072   10.915    0.000    3.516    0.715
##   PR =~
##     PictCompl         1.000                               3.194    0.599
##     BlockDes          2.100    0.309    6.794    0.000    6.707    0.718
##   WM =~
##     Arith             1.000                               3.761    0.920
##     DigSpan           0.397    0.062    6.447    0.000    1.495    0.551
##     Cod               0.820    0.216    3.790    0.000    3.085    0.296
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   VC ~~
##     PR               10.806    1.934    5.586    0.000    0.758    0.758
##     WM               10.853    1.637    6.630    0.000    0.647    0.647
##   PR ~~
##     WM                9.022    1.594    5.661    0.000    0.751    0.751
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .Simil             0.000                               0.000    0.000
##    .PictCompl         0.000                               0.000    0.000
##    .Arith             0.000                               0.000    0.000
##     VC               11.830    0.367   32.254    0.000    2.653    2.653
##     PR               17.830    0.377   47.250    0.000    5.582    5.582
##     WM               15.000    0.289   51.869    0.000    3.988    3.988
##    .Vocab             7.085    1.161    6.103    0.000    7.085    1.086
##    .Compr             5.844    0.901    6.490    0.000    5.844    1.188
##    .BlockDes        -18.769    5.566   -3.372    0.001  -18.769   -2.010
##    .DigSpan           6.209    0.940    6.608    0.000    6.209    2.288
##    .Cod              33.526    3.323   10.088    0.000   33.526    3.219
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .Simil             7.014    1.237    5.670    0.000    7.014    0.261
##    .Vocab            12.323    1.991    6.190    0.000   12.323    0.290
##    .Compr            11.823    1.394    8.481    0.000   11.823    0.489
##    .PictCompl        18.276    2.247    8.133    0.000   18.276    0.642
##    .BlockDes         42.184    7.143    5.906    0.000   42.184    0.484
##    .Arith             2.580    1.642    1.571    0.116    2.580    0.154
##    .DigSpan           5.127    0.578    8.869    0.000    5.127    0.696
##    .Cod              98.931   10.059    9.835    0.000   98.931    0.912
##     VC               19.891    2.790    7.129    0.000    1.000    1.000
##     PR               10.204    2.546    4.008    0.000    1.000    1.000
##     WM               14.146    2.316    6.109    0.000    1.000    1.000
```

```
fit.indices <- c("chisq", "df", "pvalue", "cfi", "rmsea", "rmsea.ci.lower",
                 "rmsea.ci.upper", "srmr", "aic", "bic")
fitMeasures(WISC.conf.fit, fit.indices)
```

```
##          chisq              df          pvalue             cfi           rmsea
##         62.187          34.000           0.002           0.968           0.064
## rmsea.ci.lower rmsea.ci.upper            srmr             aic             bic
##          0.038           0.089           0.037       17617.734       17833.273
```

The model fits well according to the CFI and SRMR, but not according to the χ^2 and RMSEA. However, the CI around the RMSEA contains .05. Looking at the standardized loadings in both groups, we see that all values are significant (all p values $< .001$) and substantial (all `std.all` values $> .40$, with only one exception: the coding subtest on the working memory factor in the WISC-IV, but this subtest also shows the lowest standardized loading in the WISC). So we will assume the model fits well enough. Also, we should note that the second group (WISC-IV) contributes more to the overall χ^2 value.

We inspect the residuals:

```
residuals(WISC.conf.fit, type = "cor.bentler")
```

In the WISC-IV, we see substantial residuals between Picture Completion and Digit Span and Coding. This can perhaps be solved by allowing Picture Completion to load on the Working Memory factor, but the residuals in the WISC group do not suggest the same problem. So in this case, I prefer to keep the models for the two intelligence tests the same.

Next, we test for metric (loading) invariance

```
WISC.metr.fit <- cfa(WISC.mod, sample.cov = WISC.cov.list,
                     sample.mean = WISC.mean.list,
                     sample.nobs = WISC.n.list,
                     group.equal = c("loadings"))
fitMeasures(WISC.metr.fit, fit.indices)
```

```
##          chisq              df          pvalue             cfi           rmsea
##        132.004          39.000           0.000           0.896           0.109
## rmsea.ci.lower rmsea.ci.upper            srmr             aic             bic
##          0.089           0.130           0.093       17677.551       17873.133
```

```
lavTestLRT(WISC.conf.fit, WISC.metr.fit)
```

```
## Chi-Squared Difference Test
##
##               Df   AIC   BIC   Chisq Chisq diff Df diff Pr(>Chisq)
## WISC.conf.fit 34 17618 17833  62.187
## WISC.metr.fit 39 17678 17873 132.004     69.817       5  1.119e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lavTestScore(WISC.metr.fit)
```

```
## $test
##
## total score test:
```

```
##
##    test     X2 df p.value
## 1 score 65.047  5       0
##
## $uni
##
## univariate score tests:
##
##    lhs op   rhs     X2 df p.value
## 1 .p2. == .p35. 27.669  1   0.000
## 2 .p3. == .p36.  0.052  1   0.819
## 3 .p5. == .p38.  0.680  1   0.410
## 4 .p7. == .p40. 12.023  1   0.001
## 5 .p8. == .p41. 11.753  1   0.001
```

We did not obtain full metric invariance, according to $\Delta\chi^2$ and ΔCFI. Also, RMSEA and SRMR indicate inadequate model fit for the metric invariance model. The modification indices for three of the five equality restrictions are very high, so metric invariance does not appear tenable at all. We will lift all equality restrictions on the loadings and continue testing equality of intercepts (scalar invariance). (Note that many researchers would argue you should not test for equivalence of intercepts, if metric invariance was not obtained, but for the sake of example, we will continue here):

```
WISC.scal.fit <- cfa(WISC.mod, sample.cov = WISC.cov.list,
                     sample.mean = WISC.mean.list,
                     sample.nobs = WISC.n.list,
                     group.equal = c("intercepts"))
fitMeasures(WISC.scal.fit, fit.indices)
```

```
##          chisq             df         pvalue            cfi          rmsea
##         96.400         39.000          0.000          0.936          0.086
## rmsea.ci.lower rmsea.ci.upper           srmr            aic            bic
##          0.064          0.108          0.071      17641.947      17837.529
```

```
lavTestLRT(WISC.scal.fit, WISC.conf.fit)
```

```
## Chi-Squared Difference Test
##
##               Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## WISC.conf.fit 34 17618 17833 62.187
## WISC.scal.fit 39 17642 17838 96.400     34.213       5   2.16e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lavTestScore(WISC.scal.fit)
```

```
## $test
##
## total score test:
##
##    test     X2 df p.value
## 1 score 32.087  5       0
```

```
##
## $uni
##
## univariate score tests:
##
##     lhs op   rhs    X2 df p.value
## 1 .p29. == .p62. 0.008  1   0.927
## 2 .p30. == .p63. 9.670  1   0.002
## 3 .p31. == .p64. 0.350  1   0.554
## 4 .p32. == .p65. 6.645  1   0.010
## 5 .p33. == .p66. 8.080  1   0.004
```

Again, the equality restrictions do not seem tenable according to $\Delta\chi^2$ and ΔCFI. Also, RMSEA and SRMR indicate inadequate model fit for the model with equal intercepts.

The univariate score tests indicate that three out of five equality restrictions seem untenable, which indicates that equality of intercepts is also not tenable. For the sake of example, we continue with testing the equality of measurement error variances (uniqueness invariance):

```
WISC.uni.fit <- cfa(WISC.mod, sample.cov = WISC.cov.list,
                    sample.mean = WISC.mean.list,
                    sample.nobs = WISC.n.list, meanstructure = TRUE,
                    group.equal = c("residuals"))
```

```
## Warning in lav_object_post_check(object): lavaan WARNING: some estimated lv
## variances are negative
```

We get a warning about negative LV variances, this obviously indicates a problem with our model. Restricting all measurement error variances to be equal between the two tests seems not a good idea. For the sake of example, let us still inspect the model fit:

```
fitMeasures(WISC.uni.fit, fit.indices)
```

```
##          chisq             df         pvalue            cfi          rmsea
##        189.581         42.000          0.000          0.835          0.133
## rmsea.ci.lower rmsea.ci.upper           srmr            aic            bic
##          0.114          0.152          0.105      17729.128      17912.735
```

```
lavTestLRT(WISC.uni.fit, WISC.conf.fit)
```

```
## Chi-Squared Difference Test
##
##               Df   AIC   BIC   Chisq Chisq diff Df diff Pr(>Chisq)
## WISC.conf.fit 34 17618 17833  62.187
## WISC.uni.fit  42 17729 17913 189.581     127.39       8  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lavTestScore(WISC.uni.fit)
```

```
## $test
##
## total score test:
##
##    test     X2 df p.value
## 1 score 85.914  8       0
##
## $uni
##
## univariate score tests:
##
##     lhs op   rhs     X2 df p.value
## 1 .p15. == .p48. 17.838  1   0.000
## 2 .p16. == .p49.  0.042  1   0.839
## 3 .p17. == .p50. 40.746  1   0.000
## 4 .p18. == .p51.  0.017  1   0.897
## 5 .p19. == .p52.  0.581  1   0.446
## 6 .p20. == .p53.  6.855  1   0.009
## 7 .p21. == .p54. 30.483  1   0.000
## 8 .p22. == .p55.  1.489  1   0.222
```

According to all fit indices, restricting the measurement error variances to be equal across the two WISC tests is not tenable. The univariate score tests indicate that especially the residual variances of the Similarities, Comprehension, Arithmetic and Digit Span subtests should not be restricted to equality between the two tests.

Our analyses thus indicate that more than half of the measurement parameters are different between groups. It therefore seems more appropriate to estimate all measurement parameters freely in each of the two groups, and to interpret the differences in the parameters in the least restrictive, configural invariance model.

b) The parameter estimates of the configural invariant model indicate that the standardized loadings of all subtests are higher for the WISC-IV than for the original WISC, except for the Digit Span and Coding subtests. The intercepts of all subtests increased, with exception of the intercept of BlockDesign. So given the same value of the latent VC and WM factors, we expect a higher subtest score on the WISC-IV than on the original WISC. The standardized residuals indicate that the reliabilities of all subtests were higher for the WISC-IV than for the original WISC. Furthermore, we see that the means for the latent factors are higher on the WISC-IV than on the WISC. This is an interesting finding, as the tests were administered to the same group of respondents, so we would expect the same average levels of intelligence on both WISC tests.

Additional exercise 1

As both WISC tests were administered to the same persons, it could make more sense to assume the scale of the latent variables are the same for each of the groups. Therefore, the standardized latent variable approach to identification may make more sense for this problem.

```
WISC.mod <- '
  # verbal comprehension
  VC =~ Simil + Vocab + Compr
  # perceptual reasoning
```