Testing Invariance in a Higher-Order Model

Appendix to *Latent Variable Modeling using R: A Step-by-Step Guide*

A. Alexander Beaujean

July 2014

Contents

1 Background

In this appendix, I demonstrate examining invariance in a higher-order latent variable model. Such analyses combine the information I presented in chapter 4 (*Latent Variable Models with Multiple Groups*) and chapter 9 (*Hierarchical Latent Variable Models*).

2 Example

2.1 Data

The data originally come from (Jensen & Reynolds, 1982, pp. 425,430), but were re-analyzed by Dolan (2000) using a multi-group LVM perspective. The data are scores on 13 subscales from the Wechsler Intelligence Scale for Children-Revised (WISC-R), assessed on white ($n = 1868$) and black ($n = 305$) children from approximately ages 6 to 17 years old.

```
# load lavaan
library(lavaan)

white.cor <- c(1, 0.58, 0.51, 0.66, 0.51, 0.34, 0.25, 0.35, 0.37, 0.44, 0.34, 0.26, 0.22, 1,
    0.43, 0.63, 0.55, 0.33, 0.19, 0.4, 0.37, 0.45, 0.35, 0.25, 0.24, 1, 0.48, 0.4, 0.42, 0.32,
    0.3, 0.26, 0.41, 0.23, 0.29, 0.24, 1, 0.61, 0.36, 0.24, 0.38, 0.39, 0.43, 0.33, 0.29, 0.21,
    1, 0.23, 0.19, 0.35, 0.34, 0.38, 0.29, 0.23, 0.23, 1, 0.37, 0.16, 0.18, 0.29, 0.17, 0.28,
    0.18, 1, 0.16, 0.19, 0.27, 0.15, 0.25, 0.19, 1, 0.34, 0.47, 0.41, 0.15, 0.29, 1, 0.41,
```

```
    0.37, 0.22, 0.27, 1, 0.56, 0.3, 0.39, 1, 0.2, 0.31, 1, 0.18, 1)

white.mean <- c(10.41, 10.29, 10.37, 10.42, 10.44, 10.08, 10.09, 10.41, 10.37, 10.39, 10.73,
    10.22, 10.41)

white.sd <- c(2.91, 3.01, 2.84, 2.94, 2.81, 3, 2.87, 2.87, 2.91, 2.92, 3.01, 3.3, 3.06)

black.cor <- c(1, 0.55, 0.53, 0.63, 0.49, 0.43, 0.32, 0.42, 0.29, 0.37, 0.31, 0.21, 0.26, 1,
    0.46, 0.65, 0.48, 0.34, 0.21, 0.43, 0.36, 0.41, 0.36, 0.26, 0.24, 1, 0.52, 0.39, 0.5, 0.3,
    0.32, 0.23, 0.4, 0.28, 0.28, 0.22, 1, 0.63, 0.41, 0.25, 0.43, 0.36, 0.41, 0.34, 0.28, 0.25,
    1, 0.35, 0.24, 0.44, 0.38, 0.38, 0.35, 0.26, 0.3, 1, 0.43, 0.28, 0.3, 0.35, 0.25, 0.25,
    0.28, 1, 0.29, 0.26, 0.26, 0.17, 0.25, 0.26, 1, 0.37, 0.48, 0.49, 0.16, 0.36, 1, 0.37,
    0.41, 0.21, 0.32, 1, 0.57, 0.43, 0.29, 1, 0.39, 0.19, 1, 0.18, 1)

black.mean <- c(8.09, 7.91, 8.63, 7.86, 7.83, 9.18, 9.12, 8.12, 8.1, 7.7, 7.89, 8.86, 8.39)

black.sd <- c(2.65, 2.92, 2.75, 2.76, 2.53, 3.19, 2.95, 3.03, 3.03, 2.7, 2.96, 2.93, 3.12)

# put correlation data into matrix
white.cor <- vech.reverse(white.cor)
black.cor <- vech.reverse(black.cor)

# name variables in correlation matrix and mean/sd vectors
colnames(white.cor) <- rownames(white.cor) <- colnames(black.cor) <- rownames(black.cor) <-
names(white.mean) <- names(white.sd) <- names(black.mean) <- names(black.sd) <- c("info", "sim",
    "arith", "vocab", "comp", "ds", "ts", "pcomp", "parr", "bd", "oa", "cd", "mz")

# create covariance matrix
white.cov <- cor2cov(white.cor, white.sd)
black.cov <- cor2cov(black.cor, black.sd)
```

2.2 Model

Dolan (2000) used a three-factor model of the WISC-R data with a single higher-order factor, g (i.e., general intelligence). The higher-order model is shown in Figure 1.

```
# first order model
first.order.model <- '
verb =~ info + sim + arith + vocab + comp + pcomp + parr
perf =~ sim + comp + pcomp + parr + bd + oa + cd + mz
mem =~ info + arith + ds + ts + bd + cd + mz
'
```

In addition to the χ^2 statistic, Dolan (2000) used the following alternative fit measures to assess invariance: AIC, RMSEA, and NNFI.

```
# list fit measures of interest
fit.measures <- c("chisq", "df", "pvalue", "aic", "rmsea", "nnfi")
```

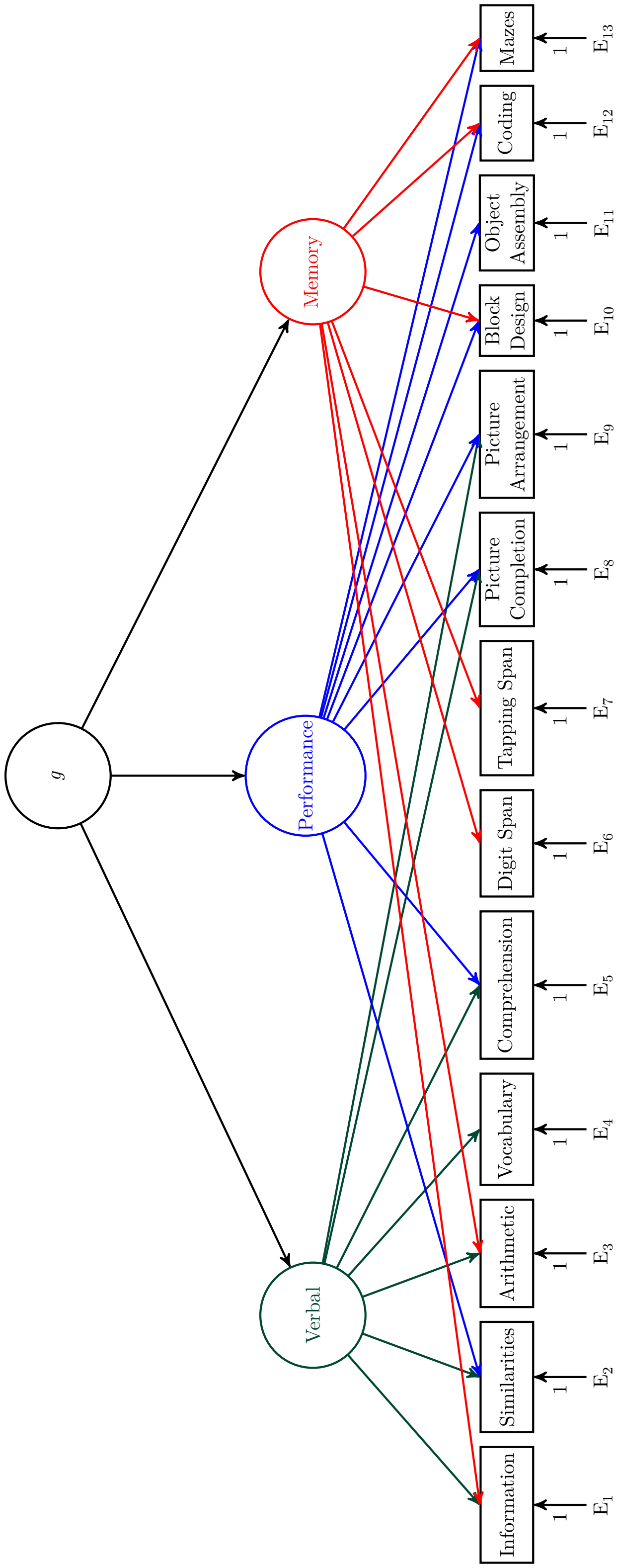

Figure 1 Path model of WISC-R subtests used by Dolan (2000).

Before assessing invariance, I need to assess model fit in each group separately. Dolan (2000) did not specify second-order factors when initially examining invariance of the first-order factors, which I mimic in the following analyses.

```
# white fit
white.fit <- cfa(first.order.model, sample.cov = white.cov, sample.nobs = 1868)
# summary(white.fit, standardized=TRUE)
fitMeasures(white.fit, fit.measures = fit.measures)
```

```
##      chisq         df     pvalue        aic      rmsea       nnfi
##    151.169     53.000      0.000 113589.828      0.031      0.982
```

```
# black fit
black.fit <- cfa(first.order.model, sample.cov = black.cov, sample.nobs = 305)
# summary(black.fit, standardized=TRUE)
fitMeasures(black.fit, fit.measures = fit.measures)
```

```
##     chisq        df    pvalue       aic     rmsea      nnfi
##    89.203    53.000     0.001 18316.070     0.047     0.962
```

2.3 Testing Invariance

Before examining invariance, I first need to combine the data into a single **R** objects.

```
# combine the two groups data
combined.cov <- list(white = white.cov, black = black.cov)
combined.mean <- list(white = white.mean, black = black.mean)
combined.n <- list(white = 1868, black = 305)
```

2.3.1 First-Order Factors

2.3.1.1 Configural Invariance

```
configural.first.fit <- cfa(first.order.model, sample.cov = combined.cov, sample.nobs = combined.n)
# summary(configural.first.fit)
fitMeasures(configural.first.fit, fit.measures = fit.measures)
```

```
##      chisq         df     pvalue        aic      rmsea       nnfi
##    240.372    106.000      0.000 131905.898      0.034      0.979
```

2.3.1.2 Weak Invariance

```
weak.first.fit <- cfa(first.order.model, sample.cov = combined.cov, , sample.nobs = combined.n,
    group.equal = c("loadings"))
# summary(weak.first.fit)
fitMeasures(weak.first.fit, fit.measures = fit.measures)
```

```
##      chisq         df     pvalue        aic      rmsea       nnfi
##    282.077    125.000      0.000 131909.604      0.034      0.979
```

2.3.1.3 Strong Invariance

```
strong.first.fit <- cfa(first.order.model, sample.cov = combined.cov, sample.mean = combined.mean,
    sample.nobs = combined.n, group.equal = c("loadings", "intercepts"))
# summary(strong.first.fit)
fitMeasures(strong.first.fit, fit.measures = fit.measures)


##      chisq         df     pvalue        aic      rmsea       nnfi
##    305.402    135.000      0.000 131964.929      0.034      0.979
```

2.3.1.4 Strict Invariance

```
strict.first.fit <- cfa(first.order.model, sample.cov = combined.cov, sample.mean = combined.mean,
    sample.nobs = combined.n, group.equal = c("loadings", "intercepts", "residuals"))
# summary(strict.first.fit)
fitMeasures(strict.first.fit, fit.measures = fit.measures)


##      chisq         df     pvalue        aic      rmsea       nnfi
##    327.173    148.000      0.000 131960.700      0.033      0.980
```

2.3.2 Second-Order Factor

To add the second-order factor, I need to re-specify the original LVM to include *g*.

```
second.order.model <- '
# first order model
verb =~ info + sim + arith + vocab + comp + pcomp + parr
perf =~ sim + comp + pcomp + parr + bd + oa + cd + mz
mem =~ info + arith + ds + ts + bd + cd + mz
# second order model
g=~ verb + perf + mem
'
```

2.3.2.1 Second-Order Loadings and First-Order (Residual) Variances

Following Dolan (2000), I first constrain the second-order loadings and residual variances of the first-order latent variables (but not the second-order factor) to be the same across groups. I do not model the means.

```
second.1.fit <- cfa(second.order.model, sample.cov=combined.cov, sample.mean=combined.mean,
sample.nobs=combined.n, group.equal=c("loadings", "residuals","lv.variances"),
group.partial=c("g~~g"))
# summary(second.1.fit.second.fit)
fitMeasures(second.1.fit, fit.measures=fit.measures)


##      chisq         df     pvalue        aic      rmsea       nnfi
##    310.034    143.000      0.000 131901.561      0.033      0.981
```

2.3.2.2 First Order (Residual) Means

Finally, I constrain the intercepts to be the same as well as the means of the first-order factors (but not the second-order factor) to be the same. This model specifies that the differences in subtests' means are solely due to difference in the second-order factor. To constraint the first-order factors' means, I set them to zero in both groups. To fit this model, I have to add constraints on the means for the first-order latent variables.

```
second.2.order.model <- '
# first order model
verb =~ info + sim + arith + vocab + comp + pcomp + parr
perf =~ sim + comp + pcomp + parr + bd + oa + cd + mz
mem =~ info + arith + ds + ts + bd + cd + mz
# second order model
g=~ verb + perf + mem
# constrain first order factor means to be zero
verb~0*1
perf~0*1
mem~0*1
'
```

```
second.2.fit <- cfa(second.2.order.model, sample.cov = combined.cov, sample.mean = combined.mean,
    sample.nobs = combined.n, meanstructure = TRUE, group.equal = c("loadings", "residuals",
        "lv.variances", "intercepts"), group.partial = c("g~~g"))
# summary(second.2.fit)
fitMeasures(second.2.fit, fit.measures = fit.measures)


##      chisq         df     pvalue        aic      rmsea       nnfi
##    392.373    155.000      0.000 132011.900      0.038      0.975
```

References

Dolan, C. V. (2000). Investigating Spearman's hypothesis by means of multi-group confirmatory factor analysis. *Multivariate Behavioral Research*, *35*, 21-50. doi: 10.1207/s15327906mbr3501_2

Jensen, A. R., & Reynolds, C. R. (1982). Race, social class and ability patterns on the WISC-R. *Personality and Individual Differences*, *3*, 423-438. doi: 10.1016/0191-8869(82)90007-1