Intelligence 57 (2016) 60–63

Contents lists available at ScienceDirect

Intelligence

On the distortion of model fit in comparing the bifactor model and the higher-order factor model☆

Dylan Molenaar

Psychological Methods, Department of Psychology, University of Amsterdam, Weesperplein 4, 1018, XA, Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 8 February 2016
Received in revised form 28 February 2016
Accepted 1 March 2016
Available online 7 May 2016

Keywords:
Intelligence
Bifactor model
Higher-order factor model

ABSTRACT

Gignac (2016) showed that if the constraint by which the higher-order factor model is nested within the bifactor model is violated (the so-called 'proportionality constraint'), model misfit relates strongly to the magnitude of the violation. In the present paper two clarifications of the results by Gignac are discussed. First, it is noted how the misfit relates to the magnitude of the violation in the simulation study by Gignac. Second, it is argued that such a pattern of distortion of model fit as found by Gignac is generally expected and not unique to the bifactor versus higher-order factor model comparison.

Studying the construct of intelligence requires a suitable psychometric model in which intelligence is operationalized in terms of its measures (commonly, the subtest scores of an intelligence test battery). Interestingly, there is yet no consensus with respect to the choice of the psychometric model for intelligence. The discussion seems to focus around two classes of models, non-hierarchical models like the correlated factor model (e.g., Horn, 1968) and hierarchical models like the bifactor model (e.g., Gignac, 2008) and the higher-order factor model (e.g., Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004). The hierarchical models separate general intelligence, or g (Jensen, 1998) from specific cognitive abilities like working memory, perceptual organization, and verbal comprehension. The bifactor model and the higher-order factor model differ in the statistical specification by which this distinction is enforced. That is, in the bifactor model, g is modeled as an additional cognitive ability underlying all subtests, while in the higher-order factor model, g is modeled as a second-order factor underlying the first-order specific cognitive abilities.

Ideally, the decision on which of the models is the most appropriate conceptualization of intelligence is guided primarily by theoretical consideration for which subsequent empirical support is gathered. Concerning the first, no consensus has yet been reached about which of the models has a better theoretical justification. See for instance Horn (1968) for arguments in favor of the correlated factor model, Jensen (1998, p.78) and Johnson and Bouchard (2004) for arguments in favor of the higher-order factor model, and Gignac (2008) and Beaujean (2015) for arguments in favor of the bifactor model.

With respect to the empirical support for the different models, generally, empirical studies seem to focus on the comparison of the bifactor model and the higher-order factor model. These studies tend to show that the bifactor model is better fitting in terms of various goodness-of-fit measures (e.g., Brunner, Nagy, & Wilhelm, 2012; Gignac, 2008; Golay & Lecerf, 2011; Keith, 2005; Watkins & Kush, 2002). However, Murray and Johnson (2013), Morgan, Hodge, Wells, and Watkins (2015) and Maydeu-Olivares and Coffman (2006) pointed out that such fit indices might be biased in favor of the bifactor model. These studies cast doubt over the statistical comparison of the bifactor model and the higher-order factor model.

The statistical difference between the bifactor model and the higher-order factor model is well understood. That is, the higher-order factor model can be seen as a special case of the bifactor model: the higher-order factor model is obtained from a bifactor model by imposing a so-called proportionality constrained on the g-loadings in the bifactor model (Yung, Thissen, & McLeod, 1999). To increase our understanding about the proportionality constraint, Gignac (2016) showed that if the constraint is violated, model misfit is strongly related to the magnitude of the violation. In the present paper two clarifications of the results by Gignac are discussed. First, it is noted how the misfit is related to the magnitude of the violation in the simulation study by Gignac. Second, it is argued that such a pattern of distortion in model fit as found by Gignac is generally expected and not unique to the bifactor versus higher-order factor model comparison.

1. The Gignac study

Gignac (2016) created 12 population correlation matrices according to a model with g and 3 specific cognitive abilities. In these matrices,

☆ The research by Dylan Molenaar was made possible by a grant from the Netherlands Organization for Scientific Research (NWO VENI- 451-15-008).

E-mail address: D.Molenaar@uva.nl.

http://dx.doi.org/10.1016/j.intell.2016.03.007
0160-2896/

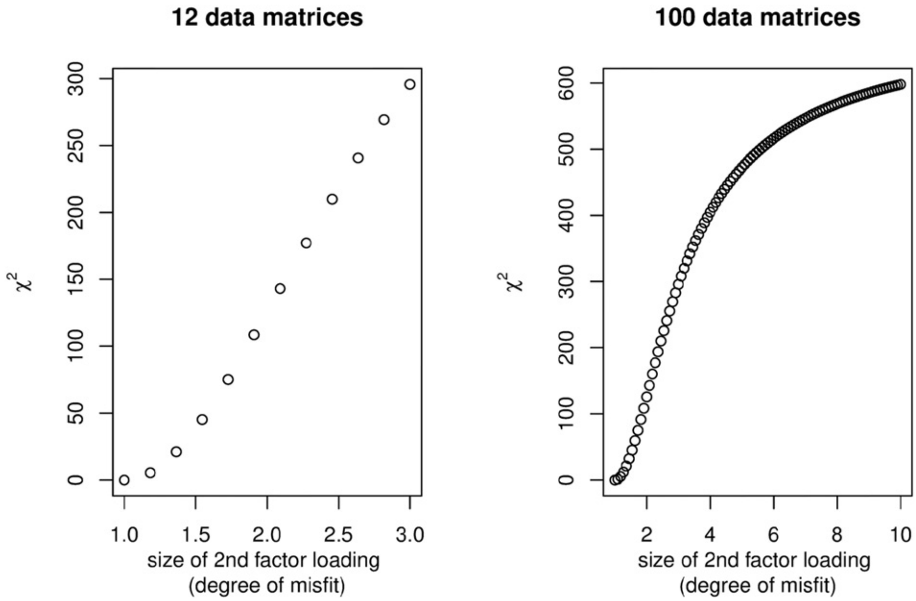

Fig. 1. The χ^2 for the model with an equal factor loading of the first and second subtest based on 12 covariance matrices with the second factor loading increasing from 1 to 3 (left), and based on 100 covariance matrices with the second factor loading increasing from 1 to 10 (right).

violations of the proportionality constraint were introduced with an increasing magnitude such that first correlation matrix did not contain a violation of the proportionality constraint (i.e., the matrix confirms a higher-order factor model) and the twelfth matrix contained a relatively large violation of the proportionality constraint. To these matrices, both the higher-order factor model and the bifactor model were fit and the χ^2, TLI, AIC, and BIC fit indices were determined. In Fig. 2 of the Gignac paper, it is shown that the relation between the difference in these fit measures between the higher-order factor model and the bifactor model tends to be highly linear.

1.1. The relation between magnitude of violation and model fit

As the TLI, AIC, and BIC fit indices are all linear transformations of the χ^2, the four plots in Fig. 2 of Gignac (2016) are identical in essence as they only differ in their mean and variance but not in their configuration (the patterns are exactly identical). Therefore, focus will be only on the χ^2. The χ^2-statistic of a model is obtained by multiplying the so-called 'fit function' with N-1, where N is the number of subjects in the data. Thus for the higher-order factor model:

$$\chi^2{}_{\text{higher-order}} = (N-1) \times f_{\text{higher-order}}(\text{data}|\text{higher-order model parameter estimates}),$$

and for the bifactor model:

$$\chi^2{}_{\text{bifactor}} = (N-1) \times f_{\text{bifactor}}(\text{data}|\text{bifactor model parameter estimates}).$$

As can be seen, the χ^2 depends on N, the data, and the model parameter estimates. If the higher-order factor model and the bifactor model are being compared, one can take the difference in χ^2 of the two models, that is

$$\chi^2{}_{\text{higher-order vs bifactor}} = \chi^2{}_{\text{higher-order}} - \chi^2{}_{\text{bifactor}}$$

This difference -which is referred to as the χ^2 difference test or likelihood ratio test- can then be used to infer whether the model fit of the higher-order factor model significant deteriorates relative to the model fit of the bifactor model.

In the study of Gignac (2016), all the population correlation matrices were generated to confirm to a bifactor model as the correlation matrices contained violations of the higher-order factor model, but no violations of the bifactor model. As population correlation matrices do not include sampling error, the bifactor model fits the correlation matrices perfectly. That is, $\chi^2{}_{\text{bifactor}}$ is 0 in all cases (indicating perfect fit). Thus, the difference in χ^2 between the higher-order factor model and the bifactor model equals the χ^2 of the higher-order factor model, that is

$$\chi^2{}_{\text{higher-order vs bifactor}} = \chi^2{}_{\text{higher-order}} - \chi^2{}_{\text{bifactor}} = \chi^2{}_{\text{higher-order}} - 0 = \chi^2{}_{\text{higher-order}}$$

This can also be seen from Tables 1, 8, and 9 of the Gignac paper as in these tables the χ^2 values for the bifactor model are all 0. As a result, the χ^2 difference between the higher-order factor model and the bifactor model will depend solely on N and the fit function, that is

$$\chi^2{}_{\text{higher-order vs bifactor}} = \chi^2{}_{\text{higher-order}} = (N-1) \times f_{\text{higher-order}}(\text{data}|\text{higher-order model parameter estimates}),$$

Two important properties about this difference follow: 1) the difference in χ^2 between the higher-order factor model and the bifactor model correlates perfectly with the sample size (N); and 2) the difference in χ^2 between the higher-order factor model and the bifactor model will be linearly related with the magnitude of the violations of the proportionality constraint only if the fit function is a linear function. However, the fit function is a non-linear function.[1] Therefore, even if the magnitude of the correlations in the data are increased linearly and the model parameters increase linearly (which is not necessarily the case), the difference in χ^2 between the bifactor model and higher-order factor model will not increase linearly due to the nonlinear nature of the fit function.

[1] The fit function is given by: $f_{\text{higher-order}} = \log(\mathbf{S}) - \log(\mathbf{\Sigma}) + \text{trace}(\mathbf{S}\mathbf{\Sigma}^{-1}) - p$, where $\mathbf{S}$ is the observed covariance matrix, $\mathbf{\Sigma}$ is the model implied covariance matrix evaluated at the parameter estimates, and p is the number of subtests.

1.2. Violations of nesting constraint will generally result in increased misfit

To illustrate the above, that is, to illustrate that the relation between the magnitude of misfit and the χ^2 is non-linear, and more importantly, to show that this relation is not unique to the violation of the proportionality constraint, a small data simulation is conducted similarly to Gignac (2016). Specifically, the focus will be on the violation of equality of two factor loadings in a one-factor *g* model for $N = 500$. To this end, population covariance matrices are generated according to a one-factor *g* model with four subtests. In the first covariance matrix (no violation), all unstandardized factor loadings are equal to 1.0. In the subsequent covariance matrices, the factor loading of the second subtest is increased. In a first example, the loading of the second subtest is increased from 1.0 to 3.0 in 12 steps. In a second example, the loading of the second subtest is increased from 1.0 to 10 in 100 steps.

To test whether the factor loading of the second subtest is equal to the factor loading of the first subtest the χ^2 difference test can be considered. That is,

$$\chi^2_{\text{equal vs unequal}} = \chi^2_{\text{equal}} - \chi^2_{\text{unequal}}$$

As the χ^2 for the model with unequal factor loadings is the true model, and will thus fit perfectly, the χ^2 of this model (χ^2_{unequal}) will be 0 as discussed above. Therefore:

$$\chi^2_{\text{equal vs unequal}} = \chi^2_{\text{equal}}$$

Thus, one only needs to fit the model with the equal first and second factor loading to obtain the difference in χ^2. For this illustration, Lavaan (Rosseel, 2012) is used. The R-script to generate the data and fit the models can be found in the Appendix A.

In Fig. 1, the relation is depicted between the size of the factor loadings of the second subtest (i.e., the magnitude of the violation) and the χ^2 of the model with an equal first and second factor loading. Note that in the left plot of Fig. 1 the magnitude of the violation (the size of the second factor loading) ranges from 1 to 3 in 12 steps (i.e., 12 covariance matrices are analyzed). As can be seen, this pattern of results resembles the pattern of results in Gignac (2016) closely. That is, the dots seem to suggest a near perfect correlation. However, as can be seen in the left plot in Fig. 1, where the size of the second factor loading ranges from 1 to 10 in 100 steps, the relation is non-linear which illustrates the above point. More importantly, Fig. 1 illustrates how the pattern of distortion in model fit as found by Gignac is not unique to the bifactor versus higher-order factor model comparison.

2. Discussion

This paper focused on the distortion of fit that Gignac (2016) found for the bifactor model if the proportionality constraint was violated. In the present paper it was shown that 1) there is a non-linear relationship between the magnitude of the proportionality constraint violation and the misfit; and 2) that this pattern of distortion is not unique to the bifactor versus second-order factor model comparison. That is, a distortion of model fit will occur whenever there is misfit.

In the paper by Gignac (2016) it is explained in an accessible way how the higher-order factor model and the bifactor model are related through the proportionality constraint. The present paper does not detract from the finding by Gignac that the AIC, BIC, and TLI show the expected statistical properties when the data are generated by a bifactor model. In that sense, they are unbiased. However, these results do not rule out that the bifactor model may be overparameterised in practice as suggested by the results of Maydeu-Olivares and Coffman (2006) and Murray and Johnson (2013).

Appendix A

```
library(lavaan)

x1=seq(1,3,length=12)

lambda1=cbind(1,x1,1,1)


x2=seq(1,10,length=100)

lambda2=cbind(1,x2,1,1)


theta=diag(c(1,1,1,1))


fact='f1 =~ y1 + y2 + y3 + y4'

fact_eq='f1 =~ l*y1 + l*y2 + y3 + y4'


nms=paste("y",1:4,sep="")

fit1=c()


for(j in 1:nrow(lambda1)){

 l=matrix(lambda1[j,])

 cov=l%*%t(l)+theta

 dimnames(cov)[[1]]=dimnames(cov)[[2]]=nms

 r=cfa(model=fact_eq,sample.cov=cov,sample.nobs=500)

 fit1[j]=r@Fit@test[[1]]$stat

}


fit2=c()


for(j in 1:nrow(lambda2)){

 l=matrix(lambda2[j,])

 cov=l%*%t(l)+theta

 dimnames(cov)[[1]]=dimnames(cov)[[2]]=nms

 r=cfa(model=fact_eq,sample.cov=cov,sample.nobs=500)

 fit2[j]=r@Fit@test[[1]]$stat

}


par(mfrow=c(1,2),mar=c(5, 5, 4, 2) + 0.1)

plot(x1,fit1,xlab="size of 2nd factor loading \n(degree of misfit)",ylab=expression(chi^2),main="12 data matrices")

plot(x2,fit2,xlab="size of 2nd factor loading \n(degree of misfit)",ylab=expression(chi^2),main="100 data matrices")
```

References

Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher-order models. *Journal of Intelligence*, *3*(4), 121–136.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, *80*, 796–846.

Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: *g* as superordinate or breadth factor? *Psychology Science*, *50*, 21–43.

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: that is why the bifactor model tends to fit better. *Intelligence*, *55*, 57–68.

Golay, P., & Lecerf, T. (2011). Orthogonal higher-order structure and confirmatory factor analysis of the French Wechsler adult intelligence scale (WAIS-III). *Psychological Assessment*, *23*, 143–152.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, *75*(3), 242–259.
Jensen, A. R. (1998). The g factor. *The science of mental ability*. Westport: Praeger.
Johnson, W., & Bouchard, T. J. (2004). The structure of human intelligence: it is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.
Johnson, W., Bouchard, T. J., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: consistent results from three test batteries. *Intelligence*, *32*, 95–107.
Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: theories, tests and issues* (2nd ed.). New York, NY: Guilford Press.
Maydeu-Olivares, A., & Coffman, D. L. (2006). Random intercept item factor analysis. *Psychological Methods*, *11*(4), 344.
Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are fit indices biased in favor of Bi factor models in cognitive ability research?: a comparison of fit in correlated factors, higher-order, and Bi-factor models via Monte Carlo simulations. *Journal of Intelligence*, *3*(1), 2–20.
Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, *41*(5), 407–422.
Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.
Watkins, M. W., & Kush, J. C. (2002). Confirmatory factor analysis of the WISC-III for students with learning disabilities. *Journal of Psychoeducational Assessment*, *20*, 4–19.
Yung, Y. F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*(2), 113–128.