

University of Groningen

Dimensionality assesment with factor analysis methods

Barendse, Mariska

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Barendse, M. (2015). *Dimensionality assesment with factor analysis methods*. [Thesis fully internal (DIV), University of Groningen]. [S.n.].

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 7

Summary and Discussion

In this dissertation, I investigated approaches to assess dimensionality with factor analysis methods. Within this framework dimensionality was defined as the number of factors that is needed to describe all statistical dependencies in the data (Lord & Novick, 1968; Zhang & Stout, 1999). I examined to what extent factor analysis methods are suited to determine the dimensionality in non-standard situations that are often encountered in practice (i.e., discrete data, multilevel structure, and/or nonlinear relations). This dissertation focused on exploratory factor analysis (Chapters 2 and 3), the analysis of multilevel data (Chapter 3), measurement bias (Chapters 4 and 5), and discrete data (Chapters 2, 3, and 4). Some new models and practical guidance for the investigation of dimensionality were provided.

In Chapter 2, we investigated the usefulness of exploratory factor analysis (EFA) to assess the dimensionality of discrete data in a simulation study. Results of different estimation methods (maximum likelihood (ML), robust ML, ML of polychoric correlations, weighted least squares (WLS), and robust WLS) demonstrated that the degree to which the number of factors is correctly identified depends on data characteristics (factor loading, sample size, and complexity of the factor structure). It was striking that the residual based fit criteria related to the robust WLS yielded results that were opposite to the results of residual based fit criteria in the ML related estimation methods. Robust WLS turned out to be

the best estimation method to assess the dimensionality of discrete data. It appeared to be more efficient than WLS, and thus requires smaller sample sizes. The robust WLS also showed less convergence problems than other investigated estimation methods. This is not surprising, because from a theoretical perspective robust WLS is more suitable for discrete data than ML related estimation methods. In general, it was difficult to identify all simulated factors even with exact fit measures. To identify major factors one can rely on the root mean square error of approximation or residual based fit indices. We encountered convergence problems with all estimation methods, especially with large numbers of common factors and small sample sizes. In practice, when we want to establish the dimensionality of a single data set, we may be able to solve convergence problems by choosing alternative identification constraints or by trying other start values. It would be interesting to examine the convergence with estimation methods other than the ones described in Chapter 2.

In Chapter 3, we addressed the issue of dimensionality in multilevel (two-level) EFA of discrete data with an empirical example. In regular multilevel EFA (i.e., without restrictions across levels) the numbers of factors can be different in within-level and between-level factor solutions (see, e.g., D'Haenens, Van Damme & Onghena, 2010). In this way latent factors have different interpretations across clusters. To identify the number of factors in a multilevel EFA, we compared a procedure without across level restrictions to a procedure with across level restrictions. The latter procedure is based on the assumption of an equal number of factors at the within-level and the between-level. Herewith, we considered three different ways of restricting the multilevel EFA structure. Across level restrictions ensure that the factors have the same interpretation across all clusters (i.e., measurement invariance across all clusters). Our empirical example analysis showed that the procedure with additional across level restrictions yielded results that were easier to interpret than the procedure without across level restrictions. As we only used one empirical dataset, further research with simulated data as well as with other empirical data is needed to investigate whether this approach is generally applicable.

In Chapters 2 and 3, we considered the evaluation of dimensionality of discrete data. As with continuous item responses, it is difficult to offer general rules of thumb (e.g., Marsh et al., 2004). Chapter 2 only offered some practical guidelines for model selection to determine the dimensionality of discrete data. Additionally, Chapter 3 showed that the dimensionality cannot be evaluated on the basis of statistics with cut-off levels only, but that substantive arguments are also important. In Chapter 3 we therefore preferred the better interpretable three-factor model over the better fitting four-factor model. The evaluation of dimensionality of multilevel discrete data is more complicated than the evaluation of single level data. This is so because the fit of multilevel EFA expresses the combined (mis)fit at multiple levels. In Chapter 3 we only investigated the residual based fit indices on each level separately. However, we could extend this to other fit criteria. The research of Ryu & West (2009) and Boulton (2011) might bring additional knowledge to take the within- and between-level sample sizes into account in calculating level specific fit criteria.

In Chapter 4, we presented the pairwise maximum likelihood (PML) estimation method (Jöreskog & Moustaki, 2001) to analyze discrete data. We compared the PML to the robust WLS that appeared to be the best estimation method in Chapter 2. As there were no readily available fit criteria for the PML method, we introduced three fit indices that were based on likelihood ratios, namely C_F (comparing the model-implied proportions of response patterns with the observed proportions of full response patterns), C_M (comparing the model-implied proportions of response patterns with the expected proportions under the assumption of multivariate normality), and C_P (comparing the model-implied proportions of pairs of item responses to the observed proportions of pairs of item responses). Results showed that the parameter estimates of the PML and the robust WLS were very similar and that the C_M and C_P fit criteria were useful in model selection. However, the C_F fit indices suffered from empty (i.e., not observed) response patterns and seemed not useful in practice. Although the PML did not outperform the robust WLS, the PML method might be advantageous from a theoretical and

from a practical point of view. From a theoretical view, the PML is an one-step method while the WLS is a multiple-step method that relies on the polychoric correlations that need to be fixed in further steps. From a practical view, we expect single step PML to perform better than WLS in large datasets and incomplete data. Further research is needed to examine whether PML is suited in other conditions encountered in empirical practice. In addition, it might be interesting to relate and compare the new PML fit criteria with other implementable fit criteria, such as the limited information fit criteria (Maydeu-Olivares, 2006; Maydeu-Olivares & Joe, 2006).

In Chapters 5 and 6 we addressed the issue of measurement bias, which is closely related to dimensionality. Originally, the restricted factor analysis (RFA) was suited to detect uniform bias only. In Chapter 5 we successfully extended the RFA method with interaction effects, both through the random slope parameterization (Muthén & Asparouhov, 2003) and through moderated structural equations (Klein & Moosbrugger, 2000), so that the RFA method can also detect nonuniform bias. In this study, the parameter estimates that represent uniform and nonuniform bias were reasonably efficient and accurate. Unexpectedly, the results of the two conceptually different methods to estimate nonuniform bias were very similar. This is probably due to the fact that they are based on the same information and that they used the same robust maximum likelihood estimation method. Results of the bias detection showed that the restricted factor analysis (RFA) methods clearly outperformed the multi-group factor analysis (MGFA) method in the detection of both uniform and nonuniform bias when the violator is continuous. With a dichotomous violator there was no noteworthy difference between the MGFA and the RFA. Results of our simulation study suggested the use of an iterative procedure instead of a single run procedure in bias detection, to keep the Type I error rate under control.

In Chapter 6 we discussed a Bayesian approach to the detection of both uniform and nonuniform measurement bias. Results of simulated data indicated that highly informative priors as well as weakly informative priors (i.e., priors that do not assume measurement bias) resulted in

accurate and efficient parameter estimates, when the violator was continuous. With a dichotomous observed violator we found less accurate results due to a loss of information and a reduction of the effect size. Overall, a multiple run procedure produced better bias detection results than a single run procedure. To detect bias, we recommend a liberal deviance information criterion (DIC) cut-off value, as small DIC values are not informative in model selection (see also Lee, 2007). We used the Gelman & Rubin (1992) convergence statistic that appeared to be an appropriate tool to determine convergence. The convergence of Bayesian RFA estimation is subject of concern, as we encountered severe convergence problems. Although we used the Raftery & Lewis (1992) diagnostic to determine the number of iterations, we expect that even more iterations are needed to avoid convergence problems. To avoid non-convergence in practice we recommend to choose more Monte Carlo chains with different initial values for each of the parameters, trying more iterations, and/or changing the initial values of the chains in empirical data. Convergence properties for a variety of settings, including a variety of sample sizes, can be an important topic for future research.

In Chapters 5 and 6 we showed that measurement bias can be efficiently investigated with both the frequentist and the Bayesian approach. We described the strengths and weaknesses of both approaches, so that researchers can make an informed decision.

In conclusion, I presented factor analysis approaches to assess dimensionality. Classical FA methods are only appropriate for linear relationships between variables with multivariate normally distributed continuous item scores. In this dissertation, I investigated the usefulness of both classical FA and generalizations thereof to assess the dimensionality in complicated data structures. Under the conditions studied here, the factor analysis models seemed to be able to handle discrete responses, nonlinear relations, and multilevel structures.

In this dissertation I only studied approaches within a factor analysis framework. It would be interesting to compare the results of Chapter 2 with those of classical methods, such as the parallel analysis, or other

exploratory methods such as non-parametric item response theory (IRT) models. It would also be interesting to compare the nonlinear models presented in Chapters 5 and 6 with the IRT models as presented by Molenaar (2012), who tested nonnormality (including interaction effects). In addition, one could compare the multilevel models studied in Chapter 3 with the Bayesian or frequentist IRT oriented multilevel models (e.g., Verhagen, 2012).