Dimensionality assesment with factor analysis methods

Barendse, Mariska

Chapter 1

Introduction

In scale construction and scale evaluation it is important to assess the dimensionality of item scores. The dimensionality of data is often defined as the minimum number of latent traits that is needed to describe all statistical dependencies in the data (Lord & Novick, 1968; Zhang & Stout, 1999). From a practical point of view, the determination of the dimensionality helps to understand the structure underlying the item responses. There are various methods to determine the dimensionality. Most of them are associated with item response models and factor analysis models. Factor analysis (FA) methods are most commonly used to assess dimensionality (e.g., Conway & Huffcutt, 2003; Ten Holt, van Duijn & Boomsma, 2010).

This dissertation will focus on FA methods to assess dimensionality. The key assumption is that the dimensionality underlying data equals the minimal number of common factors. In FA, the relationships between a large number of observed variables are described by a smaller number of common factors. The observed variance of each variable is decomposed into a part that the variable has in common with other variables and a part that is unique. A pattern matrix describes the structure by means of coefficients that express the regression of the observed variable on the unobserved common factors. The unique part is not (linearly) related to the other observed variables and it contains a specific part and a measurement error (see Bollen, 1989).

Classical FA methods are appropriate to describe linear relationships between variables with multivariate normally distributed continuous item scores. However, item scores are often discrete. Moreover, they may have a hierarchical structure (i.e., multilevel data). It is therefore important to investigate to what extent FA methods can deal with complicated data structures. The main goal of this dissertation is to investigate to what extent FA methods are suited to determine the dimensionality. In particular, I discuss exploratory factor analysis, discrete data, multilevel structures, and measurement bias. The latter is closely related to dimensionality.

Outline of This Thesis

In Chapter 2 we examine exploratory factor analyses (EFA) to determine the dimensionality underlying discrete data. EFA does not constrain the factor structure. Through maximizing the likelihood of the observed covariances, various overall goodness–of–fit measures can be considered to determine the minimal number of common factors. In this way, the dimensionality can be assessed. There are several studies that investigate dimensionality underlying continuous data (e.g., Conway & Huffcutt, 2003; Fabrigar, Wegener, MacCallum & Strahan, 1999; Preacher & MacCallum, 2003), but literature on the assessment of dimensionality underlying discrete data is limited (see Tate, 2003; Holgado-Tello, Chacón-Moscoso, Barbero-García & Vila-Abad, 2010; Gaskin & Happell, 2014). The assessment of dimensionality underlying discrete data is an important research topic, because psychological tests usually consist of item scores that are discrete (e.g., "correct"/"incorrect", or Likert scales ranging from "definitely does not apply" to "definitely does apply"). Previous research showed that treating discrete variables as continuous and maximizing the likelihood of covariances yields biased parameter estimates, incorrect standard errors, and incorrect fit statistics (e.g., Dolan, 1994; Johnson & Creech, 1983; Muthén & Kaplan, 1985, 1992; Rhemtulla, Brosseau-Liard & Savalei, 2012).

In Chapter 2 we investigate the usefulness of various fit criteria to assess the dimensionality with factor analysis methods for discrete data.

In a simulation study, we vary conditions that are often encountered in empirical data, such as the magnitude of the factor loadings, sample sizes, and factor structures, and we also investigate various fit criteria. To analyze these simulated data we use estimation methods that are often used in practice, although they are not well–suited for discrete data (maximum likelihood (ML) of covariances, ML of polychoric correlations, robust ML). Additionally, we use estimation methods that account for the discrete nature of the data (weighted least squares (WLS) and robust WLS of polychoric correlations). To determine the minimum number of common factors we used various fit criteria, both exact measures (e.g., chi–square fit criteria), and approximate measures of fit criteria (e.g., root mean square error of approximation). To test both exact and approximate measures of fit, we generate a model with major factors and minor factors. We expect that a measure of exact fit will be most suitable to indicate all factors, and that an approximate fit index is most appropriate for indicating the number of major factors.

In Chapter 3 we give an example of EFA with discrete multilevel data. Data obtained from clustered subjects complicate the assessment of the dimensionality, as the hierarchical structure of the data should be taken into account. Multilevel data are often encountered in psychological and educational data, such as data from pupils within schools or children within families. In this chapter we fit a multilevel EFA with robust WLS to determine the dimensionality in discrete two–level data. To determine the dimensionality we propose two different procedures. In the first procedure we use a model without across level restrictions, leaving room for different within–level and between–level factor solutions (possibly with varying numbers of common factors). In this case factors may not have the same interpretation across clusters. In the second procedure we impose across level restrictions, leaving both the numbers of factors and the interpretation of the factors equal across levels. This implies that the second procedure assumes measurement invariance across clusters.

In Chapter 4 we compare the performance of the – often used – robust WLS with – the relatively new – pairwise maximum likelihood (PML; Jöreskog & Moustaki, 2001) estimation method for discrete data. The

PML seems an attractive estimation method as it maximizes the sum of the likelihoods of the bivariate response patterns. In this way, it avoids maximizing the likelihood of full multivariate response patterns which is computationally very intensive. The PML is a one–step estimation method that is computationally more efficient than the full information maximum likelihood estimation method. A comparison of the PML estimation method with the stepwise robust WLS estimation method is interesting as the robust WLS is the most commonly used estimation method with discrete data. For the PML method there are no readily available goodness–of–fit statistics. In Chapter 4 we therefore suggest new fit statistics for model selection and compare the performance of the PML estimation methods to the robust WLS in a simulation study.

In Chapters 5 and 6 we discuss methods to assess measurement bias (antonym of measurement invariance), which is closely related to the determination of the dimensionality. In measurement bias research, it is examined whether observed differences in item scores do validly represent actual differences in what the test or questionnaire is supposed to measure (Oort, 1996, after Mellenbergh, 1989). A variable with respect to which an item may be biased is then referred to as "violator". From a modeling perspective a variable that is associated with a violator may be considered as an additional dimension in a measurement model. Consider a unidimensional factor model, in which the items are supposed to be conditionally independent of each other. If the observed variance of this model is partly explained by an additional biasing factor, we have to reject a unidimensional factor model in favour of a multidimensional factor model. So, measurement bias implies multidimensionality. The other way around, misspecification in the measurement model may lead to findings of bias. Consequently, if multidimensionality is not correctly accounted for in the measurement model it may show up as measurement bias (e.g., Meredith, 1993; Jak, Oort & Dolan, 2010).

Measurement bias can be investigated through multigroup factor analysis (MGFA; Meredith, 1993) with respect to a variable indicating group membership, or with restricted factor analysis (RFA; Oort, 1992, 1998) with respect to a group membership variable and/or other types of variables. In RFA, the factor model is extended with exogenous variables that represent the violator that is correlated with the factor(s) of interest. Uniform bias is indicated by direct effects of violators on the item scores. Nonuniform bias is indicated by interaction effects of violators and the constructs of interest on the item scores (Barendse, Oort & Garst, 2010). When investigating measurement bias, conceivable advantages of RFA over MGFA are that in RFA biasing variable violator can be continuous or discrete, observed or latent, and measurement bias can be examined with respect to multiple violators simultaneously.

In Chapter 5 we investigate two approaches to assess nonuniform bias, namely the random slope parameterization (Muthén & Asparouhov, 2003) and latent moderated structural equations (Klein & Moosbrugger, 2000). These RFA methods are compared to the MGFA method in detecting measurement bias. In a simulation study we vary the type of bias (uniform, nonuniform), the type of violator (dichotomous, continuous), and its relationship with the construct of interest (independent, dependent). To detect bias we use both a single run procedure and an iterative procedure that guards against Type I error (Navas-Ara & Gómez-Benito, 2002; Oort, 1998). In MGFA, continuous violators must be dichotomized. We therefore expect that the RFA method may outperform the MGFA in detecting bias when the violator is continuous.

In Chapter 6 we explore a Bayesian approach to estimate the RFA method. A Bayesian approach to detect measurement bias in MGFA has been explored (e.g., Lee, 2007; Muthén & Asparouhov, 2012), but not with RFA. Possible advantages of the Bayesian RFA over the frequentist RFA is that it may handle the estimation of the interaction term easier and that it enables the researcher to incorporate prior substantive knowledge in the analysis. The purpose of Chapter 6 is to investigate a Bayesian approach to detect both uniform and nonuniform bias with RFA models, with interaction effects as implemented by Lee (2007). In a simulation study we vary relevant conditions such as the type of bias, the type of violator, and the correlation between the trait and the violator. We investigate the accuracy of the parameter estimates and different bias detection procedures that are based on the deviance information criterion fit statistic.

Each of the chapters discussed above is self–contained and can be read independently of the other chapters. The final chapter concludes with a summary and a short discussion of the main findings in this dissertation.