Factor Analysis and Latent Variable Models in Personality Psychology

Ralf Schulze, Anna-Lena Hilger, and Philipp M Engelberg, Bergische Universität Wuppertal, Wuppertal, Germany

Abstract

A generalized model for latent structures is introduced as a general framework for the factor model to give an overview about modeling options in personality research. It is briefly described and it is shown that the factor model is one of many instances of the generalized linear latent variable model. Factor analysis as one of the most important multivariate methods in personality research is then presented with an emphasis on central issues in exploratory factor analysis. The presentation also addresses the role of confirmatory factor analysis as a method in personality research.

Introduction

Among the most central research questions of personality research are the following: What are the fundamental personality constructs and what are their relationships or what is their structure? Since constructs are commonly considered as unobserved or even unobservable variables (Bollen, 2002), personality psychology, like many other subdisciplines of psychology, can mainly be characterized as the investigation of stable psychological attributes that are not directly observed in empirical research. Nevertheless, it is always assumed that the constructs of personality psychology somehow relate to observable behavior, and that these relationships are the basis to make inferences about the existence and relations between the constructs.

The major models that are used in personality psychology to formalize the relationship between unobserved variables and observed indicators are described in this article. Unobserved variables are interchangeably called latent variables or factors in what follows. The focus will be put on basic statistical models with latent variables instead of (1) more philosophical approaches to the question of what the nature of the latent variables actually may be (cf Bollen, 2002; Borsboom et al., 2003) and (2) the many qualitatively different latent structures, like hierarchical and higher order models, for example, that are prominent in personality psychology and ability research (see Schulze, 2005).

Examples of latent variables are personality traits (e.g., extraversion) or abilities (e.g., fluid intelligence) which relate to observable or so-called manifest variables. Typical examples for manifest variables are the responses to the self-report items of a personality test. A basic premise of factor analytic models and the latent variable models considered in this article is that the manifest variables are influenced by one (or more) latent variable(s). That is, the factors are considered to be the cause for the responses. In this sense, the latent variables determine the responses and thereby the value of the indicators. If many indicators are determined by the same factor, which is a common design feature of measurement instruments in personality psychology, then the latent variable is called a common cause or common factor. Given that such common factors exist, the indicators share common variance, the extent of which depends on the strength of the relationships of the indicators with the common factors. A crucial function of factor analytic models therefore is the explanation of the shared variance (i.e., covariances) between manifest variables by the introduction of latent variables.

There are many different factor models that can be distinguished with respect to several theoretical and statistical features. To understand the differences and to be able to make an appropriate choice for a certain measurement purpose, it is helpful to examine the most important features. Among these features are the number and the properties of both latent and manifest variables.

In the statistical latent variable models, both manifest and latent variables can be defined as random variables, and they can be either discrete (categorical) or continuous. A common well-known classification scheme of variables is the concept of level of measurement (e.g., nominal, ordinal, and metric). However, since statistical methods do technically not depend on the level of measurement but mainly on distributional assumptions (Gaito, 1980), the distinction of discrete and continuous random variables is adopted in this article for the classification of both manifest and latent variables.

The distributional assumptions of the variables depend on this classification. A typical example for a discrete manifest variable is an item from an ability measure that can only be answered correctly or incorrectly. If this is the case, then the manifest variable is discrete with two possible values 0 and 1 only (i.e., dichotomous) and it follows a Bernoulli distribution. Also typical for personality research are Likert-type items that are ordinarily treated as categorized continuous variables and assumed to be sampled from a normal distribution.

The classification of latent variables as categorical or continuous depends entirely on theoretical assumptions. Intelligence, for example, is commonly considered to be a continuous trait following a normal distribution. On the other hand, latent variables can be conceptualized as categorical, like in theories of personality types or as is common in the area of personality disorders. In such cases, the latent variables follow a distribution for discrete random variables.

In the following passages, a general framework for latent variable models that includes both categorical and continuous manifest and latent variables will first be described. Next, a particular instance, the factor analytic model, will be presented in more detail as it is the most relevant in terms of frequency of use in personality research, and also because of its historical importance for the field.

A General Framework for Latent Variables

Latent variable models have a rather long and partly fragmented tradition as statistical models for psychological assessment. As a consequence, many psychometric models have been proposed that include different numbers and types of manifest and latent variables. Despite the partial fragmentation, the equivalence has been proven for some models from very different traditions (Takane and de Leeuw, 1987; Kamata and Bauer, 2008). Modern approaches integrate characteristics of diverse models and provide a general framework for latent variable modeling (McDonald, 1999; Skrondal and Rabe-Hesketh, 2004; Bartholomew et al., 2011). Among the benefits of using a general framework is that researchers are more flexible and can make sure that their model fits the data characteristics, instead of adjusting the data to fit the model requirements. It also facilitates insights into the commonalities of seemingly or traditionally different models of measurement. Even if special instances of the general model have been more common in personality research so far, the current development in personality research indicates that a wide range of models will be applied in future (Morizot et al., 2007).

The general framework for latent variables can best be described as a generalized linear model for latent variables (Skrondal and Rabe-Hesketh, 2004). Broadly speaking, the generalized linear framework allows us to model linear relations of both categorical and continuous dependent (manifest) and independent (latent) variables, or more generally, of variables sampled from distributions of the exponential family (see, e.g., Gill, 2000). The exponential family is a group of distributions whose functions can be rewritten in an exponential form. Exemplary members of the exponential family that are familiar to most researchers are the normal and the binomial distribution. Generalized linear latent models provide a framework for measurement models where either categorical or continuous observed responses are predicted by either categorical or continuous latent variables.

The Generalized Model for Latent Variables

The generalized model for latent variables consists of three components which will be very briefly introduced in the following passages. The first component of the model is a matrix of dependent observed variables Y (e.g., responses to items) sampled from a probability function of the exponential family. These random variables can be characterized by their respective expected value. For a normally distributed item (e.g., a self-report item) the expected value would simply be the mean for all respondents in the population. For a Bernoulli-distributed item (e.g., a dichotomous item scored as either correct or incorrect), the expected value would be the probability of choosing the correct option.

The second element of the model is the linear component ν that represents an interim dependent variable in the model that is predicted by the latent explanatory variables η using the corresponding weights Λ. The latent explanatory variable can be continuous or categorical.

The third component of the generalized model for latent variables is the link function. With the link function it is possible to link the expected value of the distribution of Y to the linear component ν. If Y is assumed to be normally distributed, the link function is the identity link. With the identity link, the expected value of Y can directly be linked with the linear component, resulting in a standard regression model with a latent predictor. If the dependent variable Y is not normally distributed because it is categorical, for example, different link functions have to be used. The expected value of the Bernoulli-distributed dichotomous variable, the probability of success can be linked with the linear component either with a logit link or a probit link. In other words, the link function is the part of the model where the generalization takes places. The use of different link functions provides a means to model not only normally distributed dependent variables but any dependent variables sampled from a distribution of the exponential family.

For the purpose of illustration assume that l respondents answered q items of a personality questionnaire which are determined by k latent variables. The response scale is treated as continuous and the responses are assumed to follow a multivariate normal distribution. The same assumption applies to the latent variables. In this case, the linear component ν can be written in matrix notation as

$$\nu = \Lambda\eta \qquad [1]$$

where η is a $k \times l$ vector of values of the latent variables for every respondent and Λ is a $q \times k$ matrix of weights for every item. Given the assumptions stated above, the linear component can directly be linked to the expected values of the observed variables as follows:

$$E(Y) = \mu = \Lambda\eta \qquad [2]$$

$$Y = \Lambda\eta + \Psi E \qquad [3]$$

Equation [2] is consistent with the generalized linear latent model presentation. In eqn [3] the values of the observed variables are modeled and, hence, the error term E weighted by Ψ is part of the equation.

In the dichotomous case, the observed variables follow a Bernoulli distribution. Nevertheless, an underlying latent variable may still be normally distributed but the above given link function would no longer be appropriate in this case. The logit link is one of two options to link the expected values of Y, the probability of success, to the linear component:

$$E(Y) = P(Y = 1) = \frac{\exp(\Lambda\eta)}{1 + \exp(\Lambda\eta)} \qquad [4]$$

Alternatively, the probit link can be used:

$$E(Y) = P(Y = 1) = \phi(\Lambda\eta) \qquad [5]$$

where ϕ is the standard normal cumulative distribution function. The two models in eqns [2] and [4] differ in type and distribution of the observed variables but the same distributional assumptions apply to the explanatory latent variables in both models.

In sum, the generalized linear model for latent variables provides a very flexible framework for measurement models in psychology. It encompasses many popular models for latent variables as special cases that have often been seen as distinct. For example, the model for dichotomous responses in eqns [4] or [5] is a two-parameter Item Response Theory (IRT) model or a factor model for categorical (dichotomous) variables (see Christoffersson, 1975; Takane and de Leeuw, 1987; Kamata and Bauer, 2008). Yet another example for a popular special case is the first model in eqns [2] and [3] that is known as a factor model, which will be presented in more detail in the next section. It should also be noted that, in addition to IRT models and factor models, the framework includes latent class models, longitudinal models, multilevel models, and structural equation models (Skrondal and Rabe-Hesketh, 2004).

Factor Analysis

Factor analysis is more than 100 years old and has gradually evolved over time into a multifarious statistical technique that is part of the generalized linear latent variable model class (see Bartholomew, 2007). It is probably not an overestimation of its impact on personality research in particular that without its invention personality theory would not be the same as we know it. Hence, at least a basic understanding of factor analysis is vitally important to understand current personality research. In the following, a brief overview of the most important steps is given. A far more detailed treatment of the technique and associated issues can be found in Mulaik (2010), for example.

Exploratory and Confirmatory Approaches

Factor analysis may be used with different research goals in mind. On the one end of a continuum from exploratory to confirmatory analyses, the goal may be to infer the number and relationships between latent variables underlying an empirical data set, free from any preconceived notions about the true latent structure. Exploratory factor analysis (EFA) is a form of factor analysis that is well suited for this research goal.

On the other end of the continuum, the goal of confirmatory factor analysis (CFA) is to empirically test or assess the tenability of a hypothesized latent structure for a set of observed variables. The major difference between CFA and EFA is that the latent structure has to be specified before any analyses can be done with CFA. Theoretical assumptions and previous empirical findings may be sources of knowledge to specify a factor model in advance before even the data collection begins.

Although the (common) factor model of both EFA and CFA is basically the same, there are many procedural differences that serve to accomplish the different goals of analysis. Hence, it appears to be useful to classify a factor analytic approach as EFA or CFA even when the basic model is the same and the specific goal may be located somewhere between the endpoints of the continuum. The focus of the following presentation will lean toward the EFA end because it has been the approach that is descriptively more characteristic for personality research, although this might well change in the proximate future. Terminology, issues, and problems of using the factor model in personality research will be highlighted in the subsequent sections.

Factor Model

Consistent with the presentation of the generalized model for latent variables above, the factor model can be stated in matrix notation as

$$\mathbf{Y} = \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\Psi}\mathbf{E} \qquad [6]$$

This function is called the fundamental equation of factor analysis. The vector $\boldsymbol{\eta}$ contains the common factors that are assumed to be causal for the manifest variables and, hence, for their covariances and correlations. $\mathbf{E}$ is a vector of unique factors with expected values of zero and variances of unity. The unique factors may actually be composed of more than one latent variable and/or random errors. It is essential, however, that any latent variable that is an element of $\mathbf{E}$ is causal for only one measured variable and cannot be considered to be common for two or more manifest variables.

Common factors and unique factors are weighted by their respective factor pattern matrices $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}$, where the former is the matrix of so-called loadings and the latter is a diagonal matrix because unique factors are assumed to be uncorrelated. Since common factors and unique factors are uncorrelated, the correlations between manifest variables ($\mathbf{R}_{YY}$) are only due to (1) correlations among the common factors, noted as $\mathbf{R}_{\eta\eta}$, (2) regression coefficients of common factors and manifest variables, noted as $\boldsymbol{\Lambda}$, and (3) the squared factor pattern vector of unique factors, $\boldsymbol{\Psi}^2$. Taken together, this equation is known as the *fundamental theorem of factor analysis*:

$$\mathbf{R}_{YY} = \boldsymbol{\Lambda}\mathbf{R}_{\eta\eta}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}^2 \qquad [7]$$

In factor analysis, the relations between the manifest variables are explained in terms of the latent variable structure. The correlation matrix of manifest variables that represents the relations entirely explained by common factors is the so-called reduced correlation matrix given in eqn [8]:

$$\mathbf{R}_C = \mathbf{R}_{YY} - \boldsymbol{\Psi}^2 \qquad [8]$$

Since $\boldsymbol{\Psi}^2$ is diagonal, the correlations between the manifest variables are unaffected by the subtraction in eqn [8]. The diagonal of $\mathbf{R}_C$ contains values that are called communalities. The communalities are the proportions of the variances of the q manifest variables that are determined by the k common factors. The communality of a single variable j is defined as the sum of its squared factor loadings in the case of uncorrelated (i.e., orthogonal) common factors, which is a starting assumption in EFA:

$$h_j = \lambda_{j1}^2 + \lambda_{j2}^2 + \ldots + \lambda_{jk}^2 \quad (j = 1, 2, \ldots, q) \qquad [9]$$

The reduced correlation matrix $\mathbf{R}_C$ is the basis for the extraction of common factors.

Extraction of Factors

The model stated in eqns [6] and [7] includes a number of unknown parameters like factor loadings, correlations between factors, and so forth. The first step in EFA is the estimation of the model parameters, also known as the extraction of common factors. The reduced correlation matrix $\mathbf{R}_C$ is the starting point for this estimation process. It should be noted that the communalities are already necessary at this point although they can be determined only when knowing at least the loadings of the model (see eqn [9]). As a result, the reduced correlation matrix $\mathbf{R}_C$ that is the starting point for the estimation process includes preliminary estimates for the communalities. The final communalities are determined after the model parameters are estimated.

The extraction of common factors can be done with a number of different statistical techniques. Only two of the most prominent approaches can be described briefly here: the principal axes and the maximum likelihood method.

The essence of the principal axes method in the common factor model is the calculation of factors that explain a maximum of the variance of the manifest variables, while minimizing the residuals. Summing up the squared factor loadings of the factor p gives the eigenvalue γ of this factor:

$$\gamma_p = \lambda_{1p}^2 + \lambda_{2p}^2 + \ldots + \lambda_{qp}^2 \quad (p = 1, 2, \ldots, k) \qquad [10]$$

When using the principal axes method, the reduced correlation matrix $\mathbf{R}_C$ is decomposed into its eigenvalues and eigenvectors, which serve as a basis to determine the factor loadings. A detailed description of the mathematical procedure can be found in the pertinent literature (e.g., Mulaik, 2010). Principal component analysis is another popular procedure at this stage of the process. There are a number of differences between principle component analysis and factor analysis as described here. Two of the important differences are that (1) principal component analysis is used with the goal to find a minimum number (rank) of dimensions that allow to reproduce a maximum of the variance in the raw scores of the variables, and, most importantly in the present context, and (2) the model of principal component is fundamentally different from the one presented in eqn [6] (for an elaborate comparison see, e.g., Widaman, 2007).

Another important method for the extraction of common factors is the maximum likelihood method. One of the fundamental differences to other methods of extraction in EFA, including the principal axes method, is that the maximum likelihood method provides statistical estimates of model parameters for a population in which the factor model is assumed to hold. When the parameters are estimated with this method, the likelihood of the observed correlations between the manifest variables, given estimates of the parameters, can be calculated. The procedure is repeated until the factorial structure for which the observed correlations have the maximum likelihood is found.

As Fabrigar et al. (1999) pointed out, both methods have several advantages and disadvantages. In contrast to the principal axes method, the maximum likelihood approach provides a statistical model test and the possibility to calculate several fit indices to assess the appropriateness of the factor model. The model test and indices may also be used for the comparison of models with different numbers of latent variables, for example. Statistical tests are also provided for the model parameters by using the maximum likelihood. One of the disadvantages of this method lies in the requirement to make the assumption of a multivariate normal distribution of the manifest variables. The principal axes method, in contrast, does not require distributional assumptions. Although the statistical tests may be rendered useless by the violation of the distributional assumption with maximum likelihood, the parameter estimates are not necessarily negatively affected in such a case.

Determination of the Number of Factors

One of the fundamental issues with regard to the latent structure is the question of how many factors there are. In the process of determining this number, two possible errors of misspecifications can occur. Firstly, overfactoring is given when the process results in a model with more factors than necessary. Secondly, underfactoring describes a result with a model that includes too few latent variables as common causes for the observed variables. In the latter case, the model is too parsimonious, that is, the preferred number of common factors is too small. With respect to the calculation or estimation of the model parameters, it appears to be advisable to avoid the error of underfactoring more strongly than overfactoring.

Arguably, the three most popular methods for the determination of the number of factors in EFA are:

- The eigenvalues-greater-than-one rule, also called the Kaiser–Guttman rule. This rule demands the extraction of as many factors as there are eigenvalues greater than one after the initial extraction of factors based on the correlation matrix $\mathbf{R}_{YY}$ instead of the reduced correlation matrix $\mathbf{R}_C$. In literature this method was often misapplied by referring to the wrong matrix (Fabrigar et al., 1999).
- A graphical procedure for determining the number of factors is the so-called scree test. Here, the size of the eigenvalues (based on $\mathbf{R}_{YY}$ or $\mathbf{R}_C$) is plotted against the number of eigenvalues in descending order, and adjacent points are connected by a straight line. The graph is inspected to identify the point where, from left to right, the slope of the connecting lines markedly levels off. The number of factors is determined as the number of eigenvalues to the left of the point identified.
- A simulation-based approach to determine the number of factors is parallel analysis. The idea is to repeatedly determine the eigenvalues in a simulation study that arise in a specific situation simply by chance. The situation in the simulation study is specified as identical in number of observed variables and respondents with the situation of interest. The average (or 0.95-quantile) of the simulated eigenvalues is computed for each ordered eigenvalue. The mean (or quantile) is then compared with the empirical eigenvalues. The number of factors to extract is given by the number of empirical eigenvalues that are greater than the simulated counterparts, because they are 'above chance.'

Because none of the methods is fail-safe, it appears to be reasonable to apply multiple criteria. Fabrigar et al. (1999) suggest using the scree test and parallel analysis in conjunction to find the appropriate number of factors. If the common factors are extracted with the maximum likelihood method, it is more natural to use the result of the model test and additional fit indexes to determine the number of latent variables. Of course, depending on where on the continuum from exploratory to confirmatory the specific approach is, theoretical plausibility may also be taken into consideration.

Identification and Rotational Indeterminacy

If a unique set of values can be determined for the model parameters, a model is said to be identified. When the number of common factors is larger than 1 in EFA, this is not the case. This so-called rotational indeterminacy alludes to fact that different sets of loadings in the model may be chosen that cannot be distinguished with respect to their quality of reproducing the correlation matrix $\mathbf{R}_{YY}$. It is also possible under rotational indeterminacy to choose a transformed set of loadings without changing the factor model equation fundamentally. Suppose that eqn [6] with extracted factors may be written as $\mathbf{Y} = \boldsymbol{\Lambda}^*\boldsymbol{\eta}^* + \boldsymbol{\Psi}\mathbf{E}$. This equation may be modified as

$$\begin{aligned} &\mathbf{Y} = \boldsymbol{\Lambda}^*\mathbf{M}\mathbf{M}^{-1}\boldsymbol{\eta}^* + \boldsymbol{\Psi}\mathbf{E} = \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\Psi}\mathbf{E} \\ &\text{where } \boldsymbol{\Lambda} = \boldsymbol{\Lambda}^*\mathbf{M} \\ &\text{and } \boldsymbol{\eta} = \mathbf{M}^{-1}\boldsymbol{\eta}^* \end{aligned} \quad [11]$$

Hence, there is an infinite number of $k \times k$ transformation matrices $\mathbf{M}$ that may possibly be used for adjustment of the common factors (see also Mulaik, 2010: 275 ff.). Prima facie, rotational indeterminacy may appear to be a fundamental problem of EFA but, in fact, it is actually productively exploited by most researchers when the matrix of loadings is transformed to arrive at a solution that is more interpretable.

In most cases this means to target a solution where (1) every common factor shows high factor loadings on a subset of manifest variables and low factor loadings on the remaining manifest variables, and (2) every manifest variable shows high loadings on a very small number of common factors. Such a structure would roughly be called a simple structure.

A fundamental question that has to be answered when rotating factor loadings is whether the factors are forced to remain uncorrelated or not, that is, whether $\mathbf{R}_{\eta\eta}$ is forced to remain diagonal or not. Orthogonal rotation techniques provide factor models whose factors are forced to be uncorrelated whereas oblique rotation techniques allow correlations between common factors.

There are numerous available rotation techniques, both orthogonal and oblique (for a systematic overview, see Browne, 2001). Fabrigar et al. (1999) pointed out that oblique rotation techniques do not necessarily provide high correlations between common factors. If latent variables are uncorrelated in the population, oblique rotations will give solutions with correlations close to zero. For that reason, the use of oblique rotations seems advisable in EFA.

The last step in EFA after rotation is the interpretation of the common factors. The interpretation is based on the mathematical properties of the factor model and on the meaning of the set of manifest variables that show high loadings on the same factor. For the purpose of interpretation the researcher may refer to the factor pattern matrix and to the factor structure matrix. Both matrices contain information about the relationships between the factors and the manifest variables of the model. The factor loadings $\boldsymbol{\Lambda}$ are regression weights that apply when predicting the observed variables from the common factors. They can be found in the factor pattern matrix. The factor structure matrix contains the covariances between the manifest variables and the factors. If all variables are standardized, as is customary in EFA, this matrix shows the correlations between indicators and latent variables. If the final factor model evinces a simple structure, this will be reflected especially in the factor pattern matrix. Only if an orthogonal rotation was conducted or an oblique rotation results in orthogonal factors, the factor pattern and the factor structure matrix would be identical.

Blurring the Distinction between EFA and CFA: Recent Developments

A vast number of useful extensions and enhancements of the basic factor model as presented has been published, including nonlinear, longitudinal, and even meta-analytic factor analysis, for example. Two of the more recent developments are highlighted here to illustrate how progress with the method has been fruitful for personality research. Both developments are indicative of the gradually becoming more fuzzy distinction between confirmatory and exploratory factor analytic approaches.

Firstly, stepwise exploratory factor analysis (SEFA; Kano and Harada, 2000) is a method that enhances the use of EFA for scale construction in personality research by providing a means for variable selection that helps to improve the fit of the factor model. SEFA is focused on the model fit, an aspect that is of utmost importance in confirmatory approaches, but it helps to improve fit by the elimination and addition of observed variables in a stepwise fashion. An example of the use of this method in personality research is given by Schulze and Roberts (2006).

Secondly, exploratory structural equation models (ESEM; Asparouhov and Muthén, 2009) is a method that helps to overcome the need for strict loading restrictions in CFA that often lead to factor models with an inadequate fit to personality item data. In contrast to SEFA, the data basis is not altered with ESEM. Instead, the loadings as model parameters in particular are more flexibly determined than is common in CFA by reintroducing rotational techniques into the confirmatory approach. An example of the application of ESEM using the big five model of personality is given by Marsh et al. (2010).

Overall, both approaches exemplify different methods between the endpoints of purely EFA and CFA. It appears that for personality research as it has been done for more than a century, such approaches may lead the way to enhanced instruments and models.

Synopsis

The (exploratory) factor model serves as an example for a special latent variable model which can be integrated in the framework of the generalized linear model for latent variables. For all models that include latent variables, the following steps have to be considered: (1) specifying the mathematical model, (2) estimating the model parameters, (3) (when possible) evaluating the fit of the model, and (4) interpreting the results. For EFA, these steps were briefly described above while taking into account that the EFA framework has special characteristics. That is, based on the choice of estimation method, the fit of the model is not always evaluated (e.g., with principal axes extraction). In addition, due to the issue of rotational indeterminacy, the factor loadings can be transformed to satisfy the need for a better interpretation of the loading pattern. Used with great care and understanding of the statistical theory, EFA can identify the structure of latent variables that explain the correlation of manifest variables. For more information about factor analysis see, for example, Mulaik (2010). For more information about the generalized linear model for latent variables, see Skrondal and Rabe-Hesketh (2004).

See also: Factor Analysis and Latent Structure Analysis: Confirmatory Factor Analysis; Factor Analysis and Latent Structure Analysis: Overview; Factor Analysis and Latent Structure: IRT and Rasch Models; Five Factor Model of Personality, Assessment of; Intelligence: Assessments of; Intelligence: Central Conceptions and Psychometric Models; Personality Assessment; Psychometrics.

Bibliography

Asparouhov, T., Muthén, B., 2009. Exploratory structural equation modeling. Structural Equation Modeling 16, 397–438.

Bartholomew, D.J., 2007. Three faces of factor analysis. In: Cudeck, R., MacCallum, R.C. (Eds.), Factor Analysis at 100: Historical Developments and Future Directions. Erlbaum, Mahwah, NJ, pp. 9–21.

Bartholomew, D.J., Knott, M., Moustaki, I., 2011. Latent Variable Models and Factor Analysis: A Unified Approach, third ed. Wiley, New York.

Bollen, K.A., 2002. Latent variables in psychology and the social science. Annual Review of Psychology 53, 605–634.

Borsboom, D., Mellenbergh, G.J., van Heerden, J., 2003. The theoretical status of latent variables. Psychological Review 110, 203–219.

Browne, M.W., 2001. An overview of analytic rotation in exploratory factor analysis. Multivariate Behavioral Research 36, 111–150.

Christoffersson, A., 1975. Factor analysis of dichotomized variables. Psychometrika 40, 5–32.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J., 1999. Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods 4, 272–299.

Gaito, J., 1980. Measurement scales and statistics: resurgence of an old misconception. Psychological Bulletin 87, 564–567.

Gill, J., 2000. Generalized Linear Models: A Unified Approach. Sage, Thousand Oaks, CA.

Kamata, A., Bauer, D.J., 2008. A note on the relation between factor analytic and item response theory models. Structural Equation Modeling 15, 136–153.

Kano, Y., Harada, A., 2000. Stepwise variable selection in factor analysis. Psychometrika 65, 7–22.

Marsh, H.W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A.J.S., Trautwein, U., Nagengast, B., 2010. A new look at the big five factor model through exploratory structural equation modeling. Psychological Assessment 22, 471–491.

McDonald, R.P., 1999. Test Theory: A Unified Treatment. Erlbaum, Mahwah, NJ.

Morizot, J., Ainsworth, A.T., Reise, S.P., 2007. Toward modern psychometrics: application of item response theory models in personality research. In: Robins, R.W., Fraley, R.C., Krueger, R.F. (Eds.), Handbook of Research Methods in Personality Psychology. Guilford Press, New York, pp. 407–423.

Mulaik, S.A., 2010. Foundations of Factor Analysis, second ed. Chapman & Hall, London.

Schulze, R., 2005. Modeling structures of intelligence. In: Wilhelm, O., Engle, R.W. (Eds.), Handbook of Understanding and Measuring Intelligence. Sage, Thousand Oaks, CA, pp. 241–263.

Schulze, R., Roberts, R.D., 2006. Assessing the big five: development and validation of the openness conscientiousness extraversion agreeableness neuroticism index condensed (OCEANIC). Zeitschrift für Psychologie 214, 133–149.

Skrondal, A., Rabe-Hesketh, S., 2004. Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models. Chapman & Hall/CRC, Boca Raton, FL.

Takane, Y., de Leeuw, J., 1987. On the relationship between item response theory and factor analysis of discretized variables. Psychometrika 52, 393–408.

Widaman, K.F., 2007. Common factors versus components: principals and principles, errors and misconceptions. In: Cudeck, R., MacCallum, R.C. (Eds.), Factor Analysis at 100: Historical Developments and Future Directions. Erlbaum, Mahwah, NJ, pp. 177–203.