11-1-2011

Indeterminacy of Factor Score Estimates In Slightly Misspecified Confirmatory Factor Models

André Beauducel

University of Bonn, Bonn, Germany, beauducel@uni-bonn.de

Indeterminacy of Factor Score Estimates
In Slightly Misspecified Confirmatory Factor Models

André Beauducel
University of Bonn,
Bonn, Germany

Two methods to calculate a measure for the quality of factor score estimates have been proposed. These methods were compared by means of a simulation study. The method based on a covariance matrix reproduced from a model leads to smaller effects of sampling error.

Key words: Confirmatory factor analysis, structural equation modeling, indeterminacy, factor score estimates.

Introduction

Factor score estimates are computed when individual scores representing the factors of a model are interesting. This can be the case in personnel selection or in educational settings where individuals are to be compared with respect to their scores. Thus, although latent variables might be of interest in factor analysis and structural equation modeling, some applications are still based on the concrete scores of individuals; it is for this reason that factor score estimates are of interest for applied researchers. It should be noted that although factor score estimates are termed estimates, they are not estimates in the usual sense because there are no true values that may be approximated by the estimates (Schönemann & Steiger, 1976).

The term factor score estimates denotes the aim to construct scores that represent the unknown factors in an optimal way. It follows from this reasoning that it is necessary to evaluate the quality of the factor score estimates (Gorsuch, 1983). There are two well-known

indices that allow for an evaluation of factor score indeterminacy: The multiple correlation ρ or the squared multiple correlation ρ^2 of the factor with the measured variables and the minimum correlation between two sets of factor score estimates of the same solution, $2\rho^2 - 1$ (Grice, 2001; Green, 1976; Guttman, 1955; Schönemann, 1971). Additional interesting possibilities for the evaluation of different factor score estimates with respect to their determinacy can be found in Krijnen (2006).

Although the computation of factor score estimates is also possible for confirmatory factor analysis (CFA) and specific methods have been developed for this purpose (Beauducel & Rabe, 2009), most applications and discussions of factor score indeterminacy occur in the context of exploratory factor analysis. Beauducel and Rabe (2009) present a new type of factor score estimate representing specific aspects of a CFA model (e.g., parts of a loading matrix), whereas this present study investigates two different methods to calculate factor score indeterminacy.

A difference between exploratory factor analysis and CFA is that in CFA the loadings of the variables and the correlations between factors can be specified according to theoretical assumptions. When the model assumptions are correct, fit indices would indicate that the model fits the data. However, small amounts of model-misspecification do not lead to model rejection according to many general rules (Barrett, 2007;

André Beauducel is a Professor at the Institute of Psychology in the Department of Philosophy. Email him at: beauducel@uni-bonn.de. Address for correspondence: Kaiser-Karl-Ring 9, 53111 Bonn, Germany.

Fan & Sivo, 2007; Beauducel & Wittmann, 2005; Marsh, Hau & Wen, 2004; Hu & Bentler, 1999). As a consequence, model parameters can be over- and/or under-estimated not only because of sampling error, but also because of a difference between the model parameters and the population parameters.

There is a discussion on the size of difference between model and data that might be regarded as acceptable (Marsh, et al., 2004; Barrett, 2007), but a small difference between the covariance matrix implied by the model and the empirical covariance matrix is accepted by many researchers in structural equation modeling. A difference between model and data could also occur in exploratory factor analysis, but the only way to obtain model misspecification in this context is over- or under-extraction of factors. Nevertheless, this article focuses on factor score indeterminacy as it is calculated from CFA with correctly and misspecified model parameters, because indeterminacy has rarely been evaluated in this context. A simulation study was performed in order to investigate the effects of sampling error and model misspecification on factor score indeterminacy.

The Calculation of ρ or ρ^2

It should be noted that there are two different ways to calculate indeterminacy, often referred to as ρ, the correlation between the variables and the factor (Grice, 2001). In order to present the calculation of ρ or ρ^2, the common factor model is described first. The common factor model assumes that the observations are generated by

$$X = \Lambda F + E, \qquad (1)$$

where X is the random vector of observations of order p, F the random vector with factor scores of order q, E the unobservable random error vector of order p, and Λ the factor pattern matrix of order p by q. The observations X, the factor scores F, and the error vectors E are assumed to have an expectation zero ($\varepsilon[X] = 0$, $\varepsilon[F] = 0$, $\varepsilon[E] = 0$). The covariance between the factor scores and the error scores is assumed to be zero ($\mathrm{Cov}[F, E] = 0$). The standard deviation of F is

one, the expectation of the covariance of the observed variables is Σ ($\varepsilon[XX'] = \Sigma$). The covariance matrix Σ can be decomposed by

$$\Sigma = \Lambda\Phi\Lambda' + \Psi^2, \qquad (2)$$

where Φ represents the q by q factor correlation matrix and Ψ^2 the p by p covariance matrix between the observed variables X and the error scores E ($\mathrm{Cov}[X, E] = \Psi^2$) and Ψ^2 also represents the covariance matrix of the error scores E ($\mathrm{Cov}[E, E] = \Psi^2$). Ψ^2 is generally assumed to be a diagonal matrix and it will be assumed herein that it contains only positive values. In order to investigate CFA modelling as it often occurs in empirical research, it was, however, decided also to allow for some non-diagonal elements of Ψ^2.

The factor score indeterminacy ρ, the multiple correlation of the variables with the factor can be described on the basis of Thurstone's (1935) regression score estimate, which is the best linear factor score estimate (Krijnen, Wansbeek & Ten Berge, 1996). The covariances of the factors with the best linear factor score estimates are given by

$$\mathbf{diag}(F\,\hat{F}') = \mathbf{diag}(FX'\Sigma^{-1}\Lambda\Phi). \quad (3)$$

It follows from equation 1 that it is possible to insert $\Phi\Lambda'$ for FX' into equation 3. Moreover, it is possible to standardize the covariances of the factors with the best linear factor score estimates in order to obtain the correlations. This yields

$$\begin{aligned}
&\mathbf{diag}(F\,\hat{F}') \\
&= \mathbf{diag}(\Phi\Lambda'\Sigma^{-1}\Lambda\Phi)\mathbf{diag}(\Phi\Lambda'\Sigma^{-1}\Lambda\Phi)^{-1/2} \\
&= \mathbf{diag}(\Phi\Lambda'\Sigma^{-1}\Lambda\Phi)^{1/2}
\end{aligned}$$

$$(4)$$

so that the diagonal elements in the left hand side of equation 4 contain the correlations of the best linear factor score estimates with the factors. Standardizing F is not necessary, because it has by definition a standard deviation of one. Because the best linear factor score estimate is the best linear combination of the

measured variables in order to estimate the factor, the correlations in equation 4 also represent the multiple correlations of the measured variables with the factors.

When a factor model has a perfect fit, Σ, the expectation of the covariance matrix of observed variables, which is calculated as the covariance matrix reproduced from the model parameters, and S, the empirical covariance matrix of the observed variables, are equal. Nevertheless, in the context of CFA, small differences between S and Σ regularly occur. This is always the case when the Root Mean Square Residual (RMR) is greater than zero, because this index describes the difference between these two covariance matrices. When a relevant difference between S and Σ occurs, one has to choose between these two covariance matrices for the calculation of factor score indeterminacy. The choice is to calculate indeterminacy according to equation 4 or to use the empirical covariance matrix S as in

$$\mathbf{diag}(F\,\overset{\wedge}{F'}) = \mathbf{diag}(\mathbf{\Phi\Lambda'}S^{-1}\mathbf{\Lambda\Phi})^{1/2}. \quad (5)$$

The calculation of factor score indeterminacy by means of the sample covariance matrix S has been presented by Heermann (1963), Gorsuch (1983) and Grice (2001). The calculation of indeterminacy by means of the reproduced covariance matrix, which is based on the estimated population parameters of the model, is presented in Mulaik and McDonald (1978) and in McDonald (1981). Because both ways to calculate indeterminacy are referred to in the literature and no discussion of the possible differences is currently available, this study compares the two ways to calculate indeterminacy on the basis of a simulation study. The comparison of the coefficients of indeterminacy is especially relevant to CFA, where small amounts of model misspecification are sometimes accepted (Hu & Bentler, 1999). As in other studies (Grice, 2001), the results for the squared validity coefficients (ρ^2) were presented in the following, because ρ^2 can be interpreted as the common variance between the factor and the corresponding factor score estimate.

Methodology

The aim of the simulation study was to compare the two above-mentioned coefficients of indeterminacy (equations 4 and 5) with respect to model misspecification and effects of sampling error. Therefore, the two versions of ρ^2 were first compared for the population CFA models and then for the corresponding CFA models based on samples derived from the population.

Generation of Population CFA Models

Population models based on 2, 4 and 8 factors, moderate (0.40/0.60) and large (0.60/0.80) salient loadings, with orthogonal and oblique factors (with interfactor correlations of 0.30) were investigated. The population models were chosen in order to represent CFA models as they are often found in applied research. This explains why 2-, 4- and 8-factor models were investigated, as well as the size of the loadings and the moderate size of the interfactor correlations for the oblique models. In order to perform CFA modeling like in empirical research, it is necessary to investigate not only correctly specified models but also models with small amounts of model-misspecification. A common type of model-misspecification is the omission of correlated residuals (correlated error terms of observed variables). This type of model-misspecification is interesting in the present context, because it could be expected to have an impact on the loading size and thereby on the coefficients of indeterminacy.

In the first step, the parameters of the correctly specified population models including correlated residuals were fixed to their intended values, then the corresponding population covariance matrices were reproduced from the model parameters (according to equation 2). For simplicity, the size of the model parameters was chosen in a way that ensures that the reproduced covariance matrices were correlation matrices. Finally, the population covariance matrices were used for CFA modeling in order to estimate the misspecified model parameters. The CFA modeling was performed with Mplus 3.11 by means of maximum likelihood estimation. The salient loadings were freely estimated, the non-salient loadings were fixed to zero, the variances of the factors were fixed to one and the

correlations of all residuals were fixed to zero in the misspecified models (the variance of the residuals was freely estimated). For the orthogonal models the correlations between the factors were fixed to zero, for the oblique models they were freely estimated.

Table 1 contains the correctly specified and the misspecified population loadings for the 0.40/0.60 (moderate loadings) condition and for the 0.60/0.80 (large loadings) condition for the orthogonal two-factor models based on the population covariance matrices including correlated residuals (the correlations between the residuals are presented at the bottom of Table 1).

Table 2 contains the corresponding parameters for the oblique models. The misspecified models would be accepted according to conventional cut-off criteria for fit indices (e.g., Hu & Bentler, 1999). It was intended to generate small and generally accepted amounts of model-misspecification, so that even the misspecified models investigated here represent models as they might be published in empirical research. Nevertheless, the omission of the correlated residuals leads to small errors with respect to the loading size both in the orthogonal and in the oblique model (see Tables 1 and 2). The population parameters for the orthogonal and oblique four- and eight-factor models would be identical to the corresponding parameters presented in Table 1 and 2 so that they are not presented.

Another type of model misspecification with an impact on the loading size and thereby on the coefficients of indeterminacy occurs when equality constraints are imposed on loadings that are unequal in the population. In order to base the results of the present simulation study on more than one type of model misspecification, misspecifications resulting from equality constraints on the loadings were also investigated. Again, the parameters of the correctly specified models were fixed in the first step and then the corresponding population covariance matrices were calculated from the model parameters. Finally, these population covariance matrices were used for CFA modeling with misspecified parameters. Again, the model parameters were chosen in a way to ensure that the reproduced covariance matrices were correlation matrices.

The misspecified models were again estimated by means of maximum likelihood estimation.

The variances of the factors were constrained to be one, the non-salient loadings were fixed to zero, the unconstrained salient loadings were freely estimated, and the covariance matrix of the error terms was constrained to be diagonal (there were no correlated residuals in these models, but the variances of the residuals were freely estimated). For the orthogonal models the correlations between the factors were fixed to zero, for the oblique models they were freely estimated. The misspecification for the two-factor model was introduced by means of equality constraints for each of the smaller loadings of the variables v_1-v_4 on the first factor with each of the larger loadings v_{13}-v_{16} on the second factor. For the four- and eight-factor models, similar equality constraints were imposed on the loadings of each pair of factors.

Table 3 contains the correctly specified and the misspecified population loadings for the 0.40/0.60 (moderate loadings) condition and for the 0.60/0.80 (large loadings) condition for the orthogonal two-factor models. The equality of loadings resulting from the equality constraints was not perfect in the completely standardized solutions (it was perfect in the unstandardized solutions). Not surprisingly, the fit of the correctly specified population models was perfect, but even the misspecified models fit the data very well (see Table 3). The misspecified population model would not be rejected according to conventional fit criteria (Hu & Bentler, 1999). The population loadings were the same for the four- and eight-factor models and are therefore not presented.

The population loadings for the correctly specified and the misspecified oblique two-factor models are presented in Table 4. As before, the model misspecification was introduced by means of equality constraints on loadings that were not equal in the population (see Table 4). Again an evaluation of the model fit of the misspecified models would not lead to model rejection for conventional criteria (Hu & Bentler, 1999).

Table 1: Population Loadings for the Orthogonal Two-Factor Models
(Completely Standardized Solution)

	Moderate Loadings				Large Loadings			
	Without Model Misspecification [a]		With Model Misspecification [b]		Without Model Misspecification [a]		With Model Misspecification [c]	
	F1	F2	F1	F2	F1	F2	F1	F2
x_1	.400	-	.414	-	.600	-	.607	-
x_2	.400	-	.414	-	.600	-	.607	-
x_3	.400	-	.392	-	.600	-	.596	-
x_4	.400	-	.392	-	.600	-	.596	-
x_5	.600	-	.584	-	.800	-	.793	-
x_6	.600	-	.584	-	.800	-	.793	-
x_7	.600	-	.637	-	.800	-	.816	-
x_8	.600	-	.637	-	.800	-	.816	-
x_9	-	.400	-	.414	-	.600	-	.607
x_{10}	-	.400	-	.414	-	.600	-	.607
x_{11}	-	.400	-	.392	-	.600	-	.596
x_{12}	-	.400	-	.392	-	.600	-	.596
x_{13}	-	.600	-	.584	-	.800	-	.793
x_{14}	-	.600	-	.584	-	.800	-	.793
x_{15}	-	.600	-	.637	-	.800	-	.816
x_{16}	-	.600	-	.637	-	.800	-	.816
Correlated Residuals								
x_1 with x_2	.126		.000		.096		.000	
x_7 with x_8	.096		.000		.054		.000	
x_9 with x_{10}		.126		.000		.096		.000
x_{15} with x_{16}		.096		.000		.054		.000

Notes: [a] The model fit for the population model without misspecification is perfect by definition: $\chi^2(100) = 0.00$; [b] The χ^2-test for the misspecified model with moderate loadings is non-significant even for the largest sample size used in the simulation study (N=750): $\chi^2(104) = 50.93$; Comparative Fit Index = 0.99; Root Mean Square Error of Approximation = 0.026; Standardized Root Mean Square Residual = 0.012. [c] The χ^2-test for the misspecified model with large loadings is non-significant even for the largest sample size used in the simulation study (N=750): $\chi^2(104) = 51.18$; Comparative Fit Index = 0.99; Root Mean Square Error of Approximation = 0.026; Standardized Root Mean Square Residual = 0.012.

Table 2: Population Loadings for the Oblique Two-Factor Models
(Completely Standardized Solution)

	Moderate Loadings				Large Loadings			
	Without Model Misspecification [a]		With Model Misspecification [b]		Without Model Misspecification [a]		With Model Misspecification [c]	
	F1	F2	F1	F2	F1	F2	F1	F2
x_1	.400	-	.414	-	.600	-	.607	-
x_2	.400	-	.414	-	.600	-	.607	-
x_3	.400	-	.392	-	.600	-	.596	-
x_4	.400	-	.392	-	.600	-	.596	-
x_5	.600	-	.585	-	.800	-	.793	-
x_6	.600	-	.585	-	.800	-	.793	-
x_7	.600	-	.636	-	.800	-	.815	-
x_8	.600	-	.636	-	.800	-	.815	-
x_9	-	.400	-	.414	-	.600	-	.607
x_{10}	-	.400	-	.414	-	.600	-	.607
x_{11}	-	.400	-	.392	-	.600	-	.596
x_{12}	-	.400	-	.392	-	.600	-	.596
x_{13}	-	.600	-	.585	-	.800	-	.793
x_{14}	-	.600	-	.585	-	.800	-	.793
x_{15}	-	.600	-	.636	-	.800	-	.815
x_{16}	-	.600	-	.636	-	.800	-	.815
Interfactor-Correlation	.300		.289		.300		.297	
Correlated Residuals								
x_1 with x_2	.126		.000		.096		.000	
x_7 with x_8	.096		.000		.054		.000	
x_9 with x_{10}		.126		.000		.096		.000
x_{15} with x_{16}		.096		.000		.054		.000

Notes: [a] The model fit for the population model without misspecification is perfect by definition: $\chi^2(99) = 0.00$; [b] The χ^2-test for the misspecified model with moderate loadings is non-significant even for the largest sample size used in the simulation study (N=750): $\chi^2(103) = 51.40$; Comparative Fit Index = 0.97; Root Mean Square Error of Approximation = 0.026; Standardized Root Mean Square Residual = 0.017. [c] The χ^2-test for the misspecified model with large loadings is non-significant even for the largest sample size used in the simulation study (N=750): $\chi^2(103)= 51.35$; Comparative Fit Index = 0.99; Root Mean Square Error of Approximation = 0.026; Standardized Root Mean Square Residual = 0.012.

Table 3: Population Loadings for the Orthogonal Two-Factor Models
(Completely Standardized Solution)

	Moderate Loadings				Large Loadings			
	Without Model Misspecification [a]		With Model Misspecification [b]		Without Model Misspecification [a]		With Model Misspecification [c]	
	F1	F2	F1	F2	F1	F2	F1	F2
x_1	.400	-	**.491**	-	.60	-	**.668**	-
x_2	.400	-	**.491**	-	.60	-	**.668**	-
x_3	.400	-	**.491**	-	.60	-	**.668**	-
x_4	.400	-	**.491**	-	.60	-	**.668**	-
x_5	.600	-	.622	-	.80	-	.826	-
x_6	.600	-	.622	-	.80	-	.826	-
x_7	.600	-	.622	-	.80	-	.826	-
x_8	.600	-	.622	-	.80	-	.826	-
x_9	-	.400	-	.384	-	.60	-	.569
x_{10}	-	.400	-	.384	-	.60	-	.569
x_{11}	-	.400	-	.384	-	.60	-	.569
x_{12}	-	.400	-	.384	-	.60	-	.569
x_{13}	-	.600	-	**.535**	-	.80	-	**.765**
x_{14}	-	.600	-	**.535**	-	.80	-	**.765**
x_{15}	-	.600	-	**.535**	-	.80	-	**.765**
x_{16}	-	.600	-	**.535**	-	.80	-	**.765**

Notes: [a] The model fit for the population model without misspecification is perfect by definition: $\chi^2(104) = 0.00$; [b] The χ^2-test for the misspecified model without sampling error and moderate loadings is non-significant even for the largest sample size used in the simulation study (N=750): $\chi^2(108) = 40.13$; Comparative Fit Index = 1.00; Root Mean Square Error of Approximation = 0.000; Standardized Root Mean Square Residual = 0.051. [c] The χ^2-test for the misspecified model without sampling error and large loadings is non-significant even for the largest sample size used in the simulation study (N=750): $\chi^2(108) = 38.37$; Comparative Fit Index = 1.00; Root Mean Square Error of Approximation = 0.000; Standardized Root Mean Square Residual = 0.085. The loadings resulting from an equality constraint are given in bold face. The values in brackets at the bottom of the Table are the differences between ρ^2 based on the unbiased loadings and the corresponding ρ^2 based on the biased loadings from the misspecified model.

Table 4: Population Loadings for the Oblique Two-Factor Models
(Completely Standardized Solution)

	Moderate Loadings				Large Loadings			
	Without Model Misspecification [a]		With Model Misspecification [b]		Without Model Misspecification [a]		With Model Misspecification [c]	
	F1	F2	F1	F2	F1	F2	F1	F2
x_1	.400	-	**.491**	-	.60	-	**.668**	-
x_2	.400	-	**.491**	-	.60	-	**.668**	-
x_3	.400	-	**.491**	-	.60	-	**.668**	-
x_4	.400	-	**.491**	-	.60	-	**.668**	-
x_5	.600	-	.620	-	.80	-	.825	-
x_6	.600	-	.620	-	.80	-	.825	-
x_7	.600	-	.620	-	.80	-	.825	-
x_8	.600	-	.620	-	.80	-	.825	-
x_9	-	.400	-	.385	-	.60	-	.570
x_{10}	-	.400	-	.385	-	.60	-	.570
x_{11}	-	.400	-	.385	-	.60	-	.570
x_{12}	-	.400	-	.385	-	.60	-	.570
x_{13}	-	.600	-	**.534**	-	.80	-	**.765**
x_{14}	-	.600	-	**.534**	-	.80	-	**.765**
x_{15}	-	.600	-	**.534**	-	.80	-	**.765**
x_{16}	-	.600	-	**.534**	-	.80	-	**.765**
Interfactor-Correlation	.300		.295		.300		.293	

Notes: [a] The model fit for the population model without misspecification is perfect by definition: $\chi^2(103) = 0.00$; [b] The χ^2-test for the misspecified model without sampling error and moderate loadings is non-significant even for the largest sample size used in the simulation study (N=750): $\chi^2(107) = 41.53$; Comparative Fit Index = 1.00; Root Mean Square Error of Approximation = .000; Standardized Root Mean Square Residual = 0.050. [c] The χ^2-test for the misspecified model without sampling error and large loadings is non-significant even for the largest sample size used in the simulation study (N=750): $\chi^2(107) = 40.66$; Comparative Fit Index = 1.00; Root Mean Square Error of Approximation = 0.000; Standardized Root Mean Square Residual = 0.084. The loadings resulting from an equality constraint are given in bold face. The values in brackets at the bottom of the Table are the differences between ρ^2 based on the unbiased loadings and the corresponding ρ^2 based on the biased loadings from the misspecified model.

Generation of Populations of Cases

In order to generate populations of cases corresponding to the population correlation matrices implied by the correctly specified population models, four population data sets of variables each containing normally distributed, z-standardized random numbers for 375,000 cases were computed and aggregated with SPSS Version 14.

The first set of 375,000 cases was computed for the orthogonal models with correlated residuals and the second set was computed for the oblique models with correlated residuals. The third set was computed for the orthogonal models without correlated residuals and the fourth set for the oblique models without correlated residuals. In all population data sets, the random variables were orthogonalized by means of principal component analysis with subsequent Varimax-rotation before aggregation in order to exclude that even small sampling errors might affect the population parameters.

Eight orthogonal variables were fixed as orthogonal population factor scores f_i for the orthogonal models, 64 orthogonal variables were fixed as residual or error variances e_j and 16 variables were fixed as common variables c_k representing the correlated residuals. From these orthogonal random variables eight correlated variables per factor were generated. The generation of the variables x_1 and x_2 for the orthogonal models with moderate factor loadings can be described by means of

$$x_j = .40^{0.5}f_i + .60^{0.5}(.85e_j + .15c_k),$$
$$\text{for } i = 1; j = 1, 2, k = 1. \quad (6)$$

As observed from equation 6, variables x_1 and x_2 share the common variable c_1 and therefore have correlated residuals (the error term is in brackets). Moreover, the weights in equation 6 correspond to the square-root of the (moderate) factor loadings presented in Table 1. Thus, the population loadings presented in Table 1 are the (squared) weights for the aggregation of the population factor scores in order to compute the population variables. The corresponding weights of the population residuals were computed from the communalities (h^2) by means of $w = (1 - h^2)^{0.5}$; because each variable x_j has only one non-zero population loading on one factor f_i, the

weight for f_i in equation 6 represents h, the square-root of the communality. Accordingly, the weight w for the residual in equation 6 was calculated as $w = (1 - (0.40^{0.5})^2)^{0.5} = 0.60^{0.5}$. The generation of the variables x_3 and x_4 without correlated residuals can be described by means of

$$x_j = 0.40^{0.5}f_i + 0.60^{0.5}e_j,$$
$$\text{for } i = 1; j = 3, 4. \quad (7)$$

The equation for the generation of the variables x_5 and x_6 is

$$x_j = 0.60^{0.5}f_i + 0.40^{0.5}e_j,$$
$$\text{for } i = 1; j = 5, 6; \quad (8)$$

and the equation for the variables x_7 and x_8 is

$$x_j = 0.60^{0.5}f_i + 0.40^{0.5}(.85e_j + .15c_k),$$
$$\text{for } i = 1; j = 7, 8, k = 2. \quad (9)$$

Equations 6-9 describe the generation of the eight variables loading on the first factor (see Table 1). The equations for the remaining variables loading on factors 2-8 contain the same weights (and different subscripts) and are therefore not presented here. By this procedure 64 variables with moderate loadings on eight factors were generated. The equations describing the generation of variables with large loadings on orthogonal factors and variables with correlated residuals are

$$x_j = 0.60^{0.5}f_i + 0.40^{0.5}(0.85e_j + 0.15c_k),$$
$$\text{for } i = 1; j = 1, 2, k = 1, \quad (10)$$

$$x_j = 0.60^{0.5}f_i + 0.40^{0.5}e_j,$$
$$\text{for } i = 1; j = 3, 4, \quad (11)$$

$$x_j = 0.80^{0.5}f_i + 0.20^{0.5}e_j,$$
$$\text{for } i = 1; j = 5, 6, \quad (12)$$

and
$$x_j = 0.80^{0.5}f_i + 0.20^{0.5}(0.85e_j + 0.15c_k),$$
$$\text{for } i = 1; j = 7, 8, k = 2. \quad (13)$$

For the oblique models correlated factor scores were computed by means of aggregation of orthogonal random variables. The computation of the eight oblique population factor scores o_i from the z-standardized random variables z_i and

591

a z-standardized common random variable v can be described as

$$o_i = 0.30^{0.5} v + 0.70^{0.5} z_i, \quad \text{for } i = 1 \text{ to } 8.$$
$$(14)$$

Eight oblique population factor scores were computed as a basis for the oblique two-, four- and eight-factor models. It follows from equation 14 that the interfactor-correlations were 0.30 in the population, according to the weight of the common variable v (see Beauducel & Wittmann, 2005 for more details on the aggregation of random variables). The oblique factor scores o_i were inserted instead of the orthogonal factor scores f_i into equations 6-9 in order to generate the variables for the oblique factor models with moderate loadings and correlated residuals and in equations 10-13 in order to generate the variables for the oblique models with large loadings and correlated residuals.

The two-factor models were based on o_1 and o_2, the four-factor models on o_1-o_4, and the eight factor models on o_1-o_8. For the orthogonal models without correlated residuals, the 64 variables were generated only on the basis of f_i and e_i, without the common terms c_k, so that the equations for the models contained only the weights as in equations 7, 8, 11 and 12 (see Table 3, for the corresponding loadings). For the oblique models without correlated residuals the equations were based on the random variables o_i and e_i and they had also the same weights as equations 7, 8, 11 and 12 (see Table 4, for the corresponding loadings).

Subsamples of variables were analyzed for the two- and four-factor models. The two-factor models were based on the variables x_1-x_{16} (see Table 1), the four-factor models were based on the variables x_1-x_{32} and the eight-factor models were based on the 64 variables. The two types of models and their corresponding misspecifications (omitted correlations between residuals, specification of equal loadings) were analyzed separately, in order to allow for a separate interpretation of the results.

For the analysis of the correctly and misspecified models based on population data with correlated residuals, the results from the population data sets 1 and 2 were combined in order to allow for a combined analysis of orthogonal and oblique models. The conditions for this analysis were computation method of indeterminacy (according to equations 4 and 5), orthogonality (orthogonal versus oblique), number of factors (2, 4 and 8 factors), loading size (moderate versus large loadings), and number of cases or sample size (250, 500 and 750 cases).

For each of these 36 conditions 500 samples were analyzed by means of CFA so that the first simulation study was based on 18,000 samples. For each sample one CFA with correct model specification and one CFA with incorrect model specification was performed. For analysis of the correctly and misspecified models based on population data without correlated residuals, the population data sets 3 and 4 were combined in order to allow for a combined analysis of orthogonal and oblique models. The conditions (computation method, orthogonality, number of factors, loading size and number of cases) were exactly as in the analysis of the models with correlated residuals.

For the correctly specified models, the difference between the population ρ^2 of the correctly specified models and the samples ρ^2 of the corresponding correctly specified models (same number of factors, same loading size, etc.) was calculated and averaged across factors.

For the misspecified models, the difference between the population ρ^2 of the misspecified models and the samples ρ^2 of the corresponding misspecified models (same number of factors, same loading size, etc.) was calculated and averaged across factors. The ρ^2-differences were calculated for both computation methods (see equation 4 and 5) and entered into repeated measures ANOVA.

In order to limit the results to those that are interesting in the present context, only main-effects and interactions involving the factor Computation-method are reported. Due to the very large sample size (6,000 cases) all reported effects were significant at $p < 0.001$ and only effects with large effect sizes (partial $\eta^2 > 0.20$) are reported. The effect sizes of the within-subjects effects were based on Greenhouse-Geisser corrected univariate effects.

Results

Table 5 contains the mean coefficients of indeterminacy for the different population models. The coefficients of indeterminacy were averaged for the factors with odd and even numbers, because the model misspecification based on equality constraints imposed on the loading pattern had different effects on factors with odd and even numbers. The coefficients of indeterminacy were different for the correctly and the misspecified population models (see Table 5).

For the population models based on correlated residuals the coefficients of indeterminacy were larger for all misspecified models than for the correctly specified models. For these models, the effect of misspecification on ρ^2 was identical for factors with odd and even numbers. For the models without correlated residuals, the effects of model-misspecification on ρ^2 were different for factors with odd and even numbers: For factors with odd numbers ρ^2 was larger than in the correctly specified models and for factors with even numbers ρ^2 was

smaller than in the correctly specified models. Overall, the population models show some variation of ρ^2, which might be regarded as a basis for an investigation of ρ^2 in the samples.

The differences between the population ρ^2 and the corresponding samples ρ^2 for the models based on correlated residuals were entered into a repeated measures ANOVA with Computation method (two levels, based on equations 4 and 5), Misspecification (correctly specified versus misspecified) and Number of factors (three levels) as within-subjects factors and Number of cases (three levels), Loading-size (two levels), and Obliqueness (orthogonal versus oblique) as between subjects factors. Misspecification was considered as within-subjects factor, because the same data sets were used for the correctly specified models and for the misspecified models. It was decided to consider Number of factors as within-subjects factor, because the four-factor models include the two factors of the two-factor models and the eight-factor models include the four factors of the four-factors model. A large main effect

Table 5: Mean population ρ^2 for the Two Different Calculation Methods

Model Type	Loading Size	Specification	According To Equation 4		According To Equation 5	
			Odd Factors	Even Factors	Odd Factors	Even Factors
With Correlated Residuals	.40	Correctly Specified	.738	.738	.751	.751
		Misspecified	.761	.761	.761	.761
	.60	Correctly Specified	.897	.897	.903	.903
		Misspecified	.906	.906	.906	.906
Without Correlated Residuals	.40	Correctly Specified	.751	.751	.751	.751
		Misspecified	.791	.697	.904	.623
	.60	Correctly Specified	.903	.903	.903	.903
		Misspecified	.922	.883	1.011	.823

Notes: The column odd factors contains the mean ρ^2 for the factors with odd numbers, the column even factors contains the mean ρ^2 for the factors with even numbers.
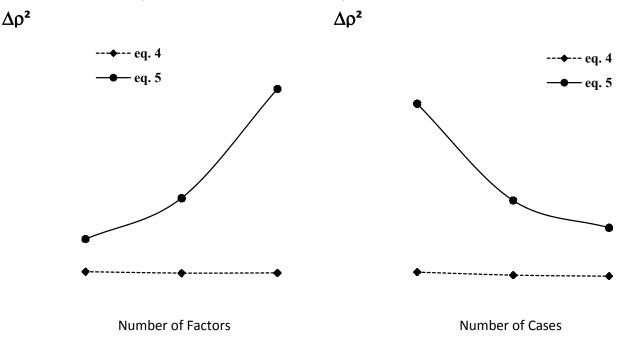
occurred for Computation method (η^2= 0.94). The mean ρ^2-difference was 0.081 (SD= 0.068) when based on equation 4 and 0.171 (SD= 0.094) when based on equation 5. Thus, the mean difference between ρ^2 in the population and in the samples was about twice as large when it was based on equation 5. This indicates that the empirical covariance matrix (used in equation 5) introduces a substantial amount of sampling error into ρ^2.

A large effect size occurred for the interaction between computation method and number of factors (η^2= 0.94). This interaction is mainly due to a larger increase of the ρ^2-difference with number of factors when ρ^2 is computed according to equation 5 (see Figure 1a). Another large effect size occurs for the interaction of computation method and number of cases (η^2= 0.81). This interaction is mainly due to a larger increase of the ρ^2-difference with Number of cases when ρ^2 is computed according to equation 5 (see Figure 1b). Moreover, a large three-way interaction computation method x number of factors x number of cases occurred

(η^2= 0.83). This three-way-interaction occurs because the size of the two-way interaction computation method x Number of factors is larger for the small samples (250 cases) than for the large samples (750 cases). In fact, the mean difference between the ρ^2-differences for the two computation methods is only 0.018 for the two-factor models based on 750 cases and it is 0.304 for the eight-factor models based on 250 cases.

Finally, the interaction of computation method with Obliqueness is of relevant size (η^2= 0.43). The difference between the computation methods is smaller for the orthogonal models than for the oblique models. Although there is a substantial main effect for misspecification (η^2= 0.47), the size of the interaction between computation method and misspecification is moderate (η^2= 0.16) and the interaction is extremely small in terms of mean differences: The difference between the ρ^2-differences for the two computation methods is 0.092 for the correctly specified models and it is 0.089 for the misspecified models; thus, misspecification had no relevant effect on the difference between the computation methods.

Figure 1: ρ^2-Differences for the Two Computation Methods Based on the Data Sets with Correlated Residuals: a) for 2-, 4-, and 8-factor models; b) for 250, 500, and 750 cases

The differences between the population ρ^2 and the corresponding samples ρ^2 based on the models without correlated residuals were entered into a repeated measures ANOVA with the same factors as the ρ^2-differences for the models based on correlated residuals. Again, a large main effect occurred for computation method ($\eta^2 = 0.97$), indicating that the mean ρ^2-difference between population and sample ρ^2 was considerably smaller when ρ^2 was computed according to equation 4.

The mean ρ^2-difference was only 0.01 (SD = 0.01) when ρ^2 was computed according to equation 4 and it was 0.14 (SD = 0.09) when ρ^2 was computed according to equation 5. A substantial interaction of computation method with number of factors occurred ($\eta^2 = 0.97$). An inspection of this interaction reveals that the computation methods had similar ρ^2-differences for the two-factor models, but that the computation method based on equation 5 yielded much larger ρ^2-differences in the eight-factor models (see Figure 2a). Another substantial interaction occurred for computation method and number of cases ($\eta^2 = 0.77$), indicating that the ρ^2-differences increased more with decreasing sample size when ρ^2 was computed according to equation 5 (see Figure 2b).

The effect size of the three-way interaction Computation method x Number of factors x Number of cases was also substantial ($\eta^2 = 0.83$). This relation of Number of factors and Number of cases with the Computation method can be described by the following result: The mean ρ^2-differences were rather similar for both Computation methods when based on the two-factor models with 750 cases (their difference was 0.033). The mean differences were, however, very different for the computation methods when based on the eight-factor models with 250 cases (their difference was 0.333). The ρ^2-differences based on equation 5 were larger than the ρ^2-differences based on equation 4 when the size of the loadings was larger (Computation method x Loading-size; $\eta^2 = 0.59$). The ρ^2-differences based on equation 5 were also larger than the ρ^2-differences ba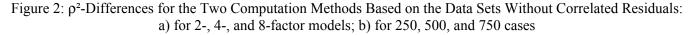sed on equation 4 for orthogonal models than for oblique models (Computation method x Obliqueness; $\eta^2 = 0.92$). The effect of model misspecification on the ρ^2-differences for the two methods was, however, moderate ($\eta^2 = 0.17$). For the correctly specified models the difference between the computation methods was slightly larger (0.125) than for the misspecified models (0.122).
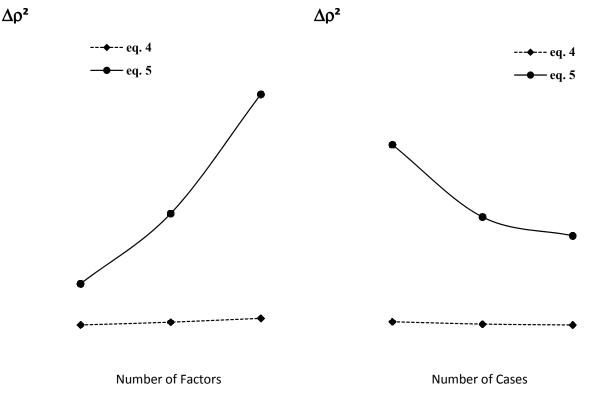
Conclusion

This study compared two calculation methods of the indeterminacy coefficient ρ^2 (or ρ) that allows for the evaluation of factor score estimates. Thereby it should be investigated which method should be preferred when a CFA model is slightly misspecified, as is often the case. Therefore, the two calculation methods for indeterminacy were compared in correctly and misspecified CFA models.

Correctly specified and misspecified models based on data sets with correlated residuals as well as on data sets without correlated residuals were investigated. For the models based on data sets with correlated residuals, the correlated residuals were not specified in order to generate misspecified models in addition to the correctly specified models. For the models based on data sets without correlated residuals misspecified models were generated by means of equality constraints imposed on unequal loadings.

Two computation methods for coefficients of indeterminacy were investigated: The first method is based on the correlations or covariances of the observed variables reproduced from the model (equation 4), the second method (equation 5) is based on the empirical correlations or covariances of the observed variables. Because both the computation of ρ^2 by means of the reproduced covariance matrix (McDonald, 1974; Mulaik & McDonald, 1978) and the computation of ρ^2 by means of the sample covariance matrix (Gorsuch, 1983; Grice, 2001; Heermann, 1963) have been proposed, an investigation of the differences between these methods was regarded as important. Moreover, in case of model misspecification, it is clear that the covariance matrix reproduced from the model (Σ) contains some error. The errors due to model

Figure 2: ρ²-Differences for the Two Computation Methods Based on the Data Sets Without Correlated Residuals:
a) for 2-, 4-, and 8-factor models; b) for 250, 500, and 750 cases

misspecification are not present in the empirical covariance matrix (**S**), so that the computation based on **S** might have been expected to work well for misspecified models. Therefore, the two computation methods were investigated both in correctly as well as in misspecified models. However, the model misspecifications were moderate in order to represent models that might be accepted according to conventional fit criteria (Hu & Bentler, 1999). The reason for the investigation of models with small amounts of misspecification was that this allows some insight into the effects of model misspecification on ρ² that might occur in empirical research with a given amount of accepted misfit. Sample size (250, 500, 750 cases), number of factors (2, 4, 8 factors), obliqueness (orthogonal versus correlated factors), and size of salient loadings (0.40/0.60 versus 0.60/0.80) were manipulated in the simulation study. The main limitations of the present simulation study are that only two types of model misspecification were explored and that the effects of severe model

misspecification were not investigated. Nevertheless, the results of the simulation study shed some light on the effects of sampling error on ρ² for different types of correctly and misspecified CFA models.

The difference between ρ² computed from the population and the samples was substantially smaller when ρ² was computed according to equation 4 (as can be seen from the main effect of Computation method). This result can be interpreted as a larger effect of sampling error on ρ² when computed according to equation 5, as might be expected from using the sample covariance matrix **S** in equation 5 instead of the population covariance matrix **Σ**.

The interpretation that the use of **S** for the computation of ρ² introduces some sampling error into the coefficient is also supported by the interaction of computation method with sample size, indicating that the difference between the population ρ² and the sample ρ² was larger for smaller sample sizes, especially when ρ² was

computed according to equation 5 (based on **S**). Even in the misspecified models, when Σ suffers from the misspecification, due to its being reproduced from the (misspecified) model parameters, the mean differences between the populations ρ^2 and the samples ρ^2 was smaller when ρ^2 was computed on the basis of Σ (equation 4).

Although the model misspecifications used in the present study were not very large, it is still possible that advantages of using **S** for the computation of ρ^2 (equation 5) might occur for extreme amounts of model misspecification. On the other hand, it seems rather unlikely that severely misspecified models would generally be accepted according to fit indexes and it might be regarded as problematic to base the results of a simulation study on models that should not occur in empirical research. The results of the present study are therefore taken as support for a computation of ρ^2 by means of the reproduced correlation or covariance matrix (equation 4). Moreover, it was found for the population models that effects of misspecification can result in serious over-estimation of ρ^2, so that the validity of factor score predictors might be over-estimated, just because the respective models were incorrectly specified.

Nevertheless, the effect of sampling error and model misspecification on ρ^2 found in this study should not discourage researchers to report indeterminacy coefficients when factor score estimates are computed from CFA models. It is necessary to report indeterminacy coefficients – otherwise the validity of the factor score estimates remains unknown. Of course, indeterminacy coefficients might be even more biased than reported here when a model is more seriously misspecified; the case of extreme misspecification was not investigated in this study because factor score estimates should not at all be computed for seriously misspecified CFA models, thus the question of the validity of such scores is irrelevant.

References

Barrett, P. (2007). Structural equation modeling: Adjudging model fit. *Personality and Individual Differences*, *42*, 815-824.

Beauducel, A., & Rabe, S. (2009). Model-related factor score estimates for confirmatory factor analysis. *British Journal of Mathematical and Statistical Psychology*, *62*, 489-506.

Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling*, *12*, 41-75.

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, *42*, 509-529.

Gorsuch, R. L. (1983). *Factor analysis* (*2nd Ed.*). Hillsdale, NJ: Lawrence Erlbaum Associates.

Green, B. F. (1976). On the factor score controversy. *Psychometrika*, *41*, 263-266.

Guttman, L. (1955). The determinacy of factor score matrices with applications for five other problems of common factor theory. *British Journal of Mathematical and Statistical Psychology*, *8,* 65-82.

Grice, J. W. (2001). Computing and evaluation of factor scores. *Psychological Methods*, *6*, 430-450.

Heermann, E. F. (1963). Univocal or orthogonal estimators of orthogonal factors. *Psychometrika*, *28*, 161-172.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Krijnen, W. P. (2006). Some results on mean square error for factor score prediction. *Psychometrika*, *71*, 395-409.

Krijnen, W. P., Wansbeek, T., & Ten Berge, J. M. F. (1996). Best linear predictors for factor scores. *Communications in Statistics: Theory and Methods*, *25*, 3013-3025.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320–341.

McDonald, R. P., & Burr, E. J. (1967). A comparison of four methods of constructing factor scores. *Psychometrika*, *32*, 381-401.

McDonald, R. P. (1974). The measurement of factor indeterminacy. *Psychometrika*, *39*, 203-222.

Mulaik, S., & McDonald, R. P. (1978). The effect of additional variables on factor indeterminacy in models with a single common factor. *Psychometrika*, *43*, 177-192.

Muthén, L. K., & Muthén, B. O. (2004). *Mplus user's guide (Version 3.11).* Los Angeles: Author.

Schönemann, P. H. (1971). The minimum average correlation between equivalent sets of uncorrelated factors. *Psychometrika*, *36*, 21-30.

Schönemann, P. H., & Steiger, J. H. (1976). Regression component analysis. *British Journal of Mathematical and Statistical Psychology*, *29*, 175-189.

SPSS. (2005). *SPSS for Windows, Release 14.0.0.* Chicago: Author.

Thurstone, L. L. (1935). *The vectors of mind.* Chicago: University of Chicago Press.