Coefficients of Factor Score Determinacy for Mean Plausible Values of Bayesian Factor Analysis

André Beauducel and Norbert Hilger

Abstract
In the context of Bayesian factor analysis, it is possible to compute plausible values, which might be used as covariates or predictors or to provide individual scores for the Bayesian latent variables. Previous simulation studies ascertained the validity of mean plausible values by the mean squared difference of the mean plausible values and the generating factor scores. However, the mean correlation of sets of single plausible values of different factors were shown to be an adequate estimator of the correlation between factors. Using sets of single plausible values to compute a mean prediction in secondary analysis implies that their determinacy should be known. Therefore, a plausible value-based determinacy coefficient allowing for estimation of the determinacy of single plausible values was proposed and evaluated by means of two simulation studies. The first simulation study demonstrated that the plausible value-based determinacy coefficient is an adequate estimate of the correlation of single plausible values with the population factor. It is also shown that the plausible value-based determinacy coefficient of mean plausible values approaches the conventional, model parameter-based determinacy coefficient with increasing number of imputations. The second simulation study revealed that the plausible value-based determinacy coefficient and the model parameter-based determinacy coefficient yield similar results even for misspecified models in small samples. It also revealed that for small sample sizes and a small salient loading size, the coefficients of determinacy overestimate the validity, so that it is recommended to report the determinacy coefficients together with a bias-correction to estimate the validity of plausible values in empirical settings.

1University of Bonn, Germany

Corresponding Author:
André Beauducel, Institute of Psychology, Rheinische Friedrich-Wilhelms-Universität Bonn, Kaiser-Karl-Ring 9, Bonn 53111, Germany.
Email: beauducel@uni-bonn.de
Asparouhov and Mutheén (2010a) proposed plausible values in the context of Bayesian structural equation modeling (Mutheén & Asparouhov, 2012), which might be used as covariates or predictors or to provide individual scores for the Bayesian latent variables. Although there has been some critique of Bayesian structural equation modeling (Stromeyer et al., 2015), several critical issues were successfully addressed by means of a series of specifications (Asparouhov et al., 2015). However, there are cautionary notes on the use of default settings for convergence (Zitzmann & Hecht, 2019) or on the use of default prior distributions (Smid & Winter, 2020). Nevertheless, Bayesian factor analysis (BFA) is meanwhile regularly used (e.g., Bonafede et al., 2021; Wang et al., 2021; Weide et al., 2021). It follows from the regular use of BFA that plausible values as factor score estimates computed from BFA (Asparouhov & Mutheén, 2010a, 2010b) may also be used regularly. Asparouhov and Mutheén (2010a) have shown and discussed different applications and advantages of plausible value factor scores. However, the use of plausible values requires an indicator for the validity of plausible values.

As individual scores on factors can be of practical interest, it is not surprising that factor score estimates with different interesting properties have been proposed and evaluated in the framework of conventional factor analysis (Grice, 2001; McDonald, 1981, 2011; Mulaik, 2010). A well-known property of factor score estimates is their determinacy or validity, that is, the correlation of the factor score estimates with the factor (Beauducel & Hilger, 2015; Guttman, 1955). Sometimes, when the factor score means are of relevance, the mean squared error (MSE), that is, the mean squared difference between the factor scores and the true factor is also evaluated (Krijnen et al., 1996). The best mean-centered linear factor score estimate has the smallest MSE and the highest determinacy, that is, the highest correlation with the factor. It has been recommended to report the coefficient of determinacy whenever factor score estimates are computed in empirical studies, where the original factor scores are principally unknown (Ferrando & Lorenzo-Seva, 2018; Grice, 2001). It has also been noted that the squared coefficient of determinacy may even be reported as an indicator of the reliability of factor score estimates (Ferrando & Lorenzo-Seva, 2018). Accordingly, the first aim of this study was to investigate how well the coefficient of determinacy and the MSE indicate the validity of the mean plausible values as factor score estimates when computed from the Bayesian model parameters.

Asparouhov and Mutheén (2010a) have shown that the correlation between two factors averaged across the correlations of single plausible values for the two factors yields a more accurate estimate of the population factor intercorrelation than the correlation between the best linear factor score estimates for the two factors. Therefore, the mean prediction based on the predictions of sets of single plausible values may be used in secondary analysis. When single plausible values are used for further analysis,
their factor score determinacy may be of interest. Therefore, after some definitions, a coefficient allowing for the estimation of the determinacy of single plausible values will be proposed. However, the conventional determinacy coefficient is solely based on model parameters, the new coefficient is a plausible value-based determinacy coefficient. On this basis, the second aim of the study was the investigation of the determinacy of single plausible values and of the effect of the number of imputations and model misfit on the correlation of mean plausible values with the population factors and on the plausible value-based determinacy coefficient. The determinacy of single plausible values as well as the effect of the number of imputations on the determinacy of mean plausible values were investigated in a first simulation study. A second simulation study was performed to compare the effects of sampling error and model fit on the plausible value-based determinacy coefficient with the respective effects on the model parameter-based determinacy coefficient.

**Definitions**

The population common factor model can be defined as

\[ x = \Lambda f + \Psi e, \]  

(1)

where \( f \) is a normally distributed random vector representing \( q \) common factors, \( \Lambda \) is a \( p \times q \) matrix of common factor loadings, where \( p \) represents the number of observed variables, \( e \) is a normally distributed random vector of \( p \) linear independent unique factors, and \( \Psi \) is a \( p \times p \) positive definite matrix of unique or error factor loadings. It is assumed that \( \mathbb{E}(f) = 0 \), \( \mathbb{E}(ff) = \Phi \), \( \text{diag}(\Phi) = I_q \), \( \mathbb{E}(e) = 0 \), \( \mathbb{E}(fe) = 0 \), and \( \mathbb{E}(ee) = I \), resulting in

\[ \Sigma = \mathbb{E}(xx') = \Lambda \Lambda' + \Psi^2. \]  

(2)

It follows from the model assumptions that postmultiplication of equation (1) with the transpose of the generating factor yields

\[ \mathbb{E}(xf) = \Lambda \Phi, \]  

(3)

so that the covariance of the generating factor with the measured variables can be expressed in terms of the model parameters \( \Lambda \) and \( \Phi \).

**Coefficients of Determinacy and MSE**

The best linear factor score estimate, that is, the factor score estimate with the highest correlation with the factor, has already been proposed by Thurstone (1935) and can be written as

\[ \hat{f}_{BL} = \Phi \Lambda' \Sigma^{-1} x = B_{BL} x, \]  

(4)
where $\Phi$ represents the $q \times q$ matrix of factor intercorrelations, $\Lambda$ the $p \times q$ matrix of factor loadings, $\Sigma$ the $p \times p$ covariance matrix of observed variables, and $x$, the random vector of $p$ observed variables, $BBL$ is the coefficient matrix for the best linear factor score estimate. The correlation of $\hat{f}_{BL}$ with the factor $f$ is the coefficient of determinacy (Ferrando & Lorenzo-Seva, 2018), which is given by

$$CD_{BL} = \text{diag}(E(\hat{f}_{BL} \hat{f}_{BL}'))^{-1/2} \text{diag}(E(\hat{f}_{BL} f')) = \text{diag}(\Phi \Lambda' \Sigma^{-1} \Lambda \Phi)^{-1/2} \text{diag}(E(\Phi \Lambda' \Sigma^{-1} x'))$$

$$= \text{diag}(\Phi \Lambda' \Sigma^{-1} \Lambda \Phi)^{1/2}.$$  

(5)

Equation 5 is of interest when factor score estimates are computed in the context of empirical investigations because it is only based on model parameters because it is based on equation 3. Therefore, equation 5 allows to compute an indicator for the validity of factor score estimates even when the original factor scores are not available. Equation 5 follows from the population factor model. However, some amount of model misfit might occur because of sampling error, model misspecification, or model error of the factor model (MacCallum, 2003). As the amount of model fit depends on the estimation method (MacCallum et al., 2007) and as the effect of Bayesian estimation on the coefficient of determinacy has until now not been investigated, it is not yet known how well equation 5 represents the validity of mean plausible values. A related question is whether the mean plausible value factor scores $\hat{f}_{PV}$ of a large number of imputations is a proxy of the best linear factor score estimate when the loadings $\Lambda$ and factor intercorrelations $\Phi$ are estimated from BFA.

If $\hat{f}_{PV} \neq \hat{f}_{BL}$, as might occur when single plausible values are used instead of the mean of a set of plausible values $B_{PV}$ can be computed $a$ posteriori because $\hat{f}_{PV} = B_{PV}x$ implies

$$B_{PV} = \Sigma^{-1} x \hat{f}_{PV}.$$  

(6)

The factor score determinacy for $\hat{f}_{PV}$ is then

$$CD_{PV} = \text{diag}(E(\hat{f}_{PV} \hat{f}_{PV}'))^{-1/2} \text{diag}(E(\hat{f}_{PV} f')) = \text{diag}(E(\hat{f}_{PV} \hat{f}_{PV}'))^{-1/2} \text{diag}(E(\hat{f}_{PV} x' \Sigma^{-1} \Lambda \Phi)).$$  

(7)

Although one would expect very similar results for the computation of factor score determinacy according to equations (5) and (7) when the mean plausible values are computed from a large set of plausible values, the differences between these equations can be substantial when single plausible values are used or when mean plausible values are computed from a small set of plausible values. One would expect that factor score determinacy computed from equation (7) approaches the value computed from equation (5) for mean plausible values based on an increasing number of imputations.
Luo and Dimitrov (2018) computed the correlation of the mean plausible values of BFA with the generating factors, that is, factor score determinacy in the context of a simulation study. They found sufficiently high correlations of the mean plausible values with the original factors when the mean plausible values are based on at least 20 imputations. However, they did not compare the correlation of the mean plausible values with the factor scores with the coefficient of determinacy as it is computed from the model parameters by means of equation (5) and (7). Therefore, a first simulation study was performed to investigate the determinacy of single plausible values and to compare the effect of the number of imputations on the coefficient of determinacy computed from equation (7) with the correlation of the mean plausible values with the factor scores and with the coefficient of determinacy computed from equation (5).

Moreover, Budescu (1982) discussed the coefficient of determinacy in the context of multiple correlation and proposed to perform bias-correction similarly to the computation of the adjusted multiple correlation. Accordingly, Budescu (1982, p. 973, equation 3) proposed the following equation for the computation of bias for the squared coefficient of determinacy

$$\text{Bias}(CD^2) = \left( \frac{p - 2}{n - p - 1} \right) (1 - CD^2) + \left( \frac{2(n - 3)}{(n - p)^2 - 1} \right) (1 - CD^2)^2. \quad (8)$$

Accordingly, a bias-corrected coefficient of determinacy can be computed as

$$CD_{corr} = (CD^2 - \text{Bias}(CD^2))^{1/2}.$$

Although the bias correction can be applied for the coefficients computed according equation (5) and (7), it will be performed only for the conventional coefficient of determinacy because Budescu (1982) proposed this correction for this coefficient.

The MSE may provide different results than the coefficient of determinacy when the mean or standard deviation of the mean plausible values are different from the mean and standard deviation of the factor scores. Asparouhov and Muthén (2010a) compared the MSE of mean plausible values based on 500 imputations with the MSE of the best linear factor score estimate computed from maximum likelihood confirmatory factor analysis and found an extreme similarity for 10,000 cases. Luo and Dimitrov (2018) found that even fewer imputations may lead to an appropriate MSE for mean plausible values. As Asparouhov and Muthén (2010a) and Luo and Dimitrov (2018) used the difference between the mean plausible values and the factor scores to compute the MSE, they did not use an MSE that is based on model parameters.

Although the MSE can be computed from the difference between the mean plausible values and the original factor scores in the context of simulation studies, this is impossible in empirical studies. It is therefore of interest to compute the MSE in a way that does not require the original factor scores. When the MSE is computed from the model parameters and from mean-centered $\hat{f}_{BL}$, it follows from Krijnen et al. (1996, p. 3016, Equation 5) that it can be written as
MSE(\(\hat{f}\)) = \left\{ (\Sigma B - \Lambda \Phi)' \Sigma^{-1} (\Sigma B - \Lambda \Phi) \right\} + (\Phi - \Phi \Lambda' \Sigma^{-1} \Lambda \Phi). \quad (9)

As the matrix in braces is positive definite, Krijnen et al. (1996) conclude that MSE(\(\hat{f}\)) is minimal over all matrices \(B\) if \((\Sigma B - \Lambda \Phi) = 0\), which is given for \(B = \Sigma^{-1} \Lambda \Phi = B_{BL}\) so that
\[
MSE(\hat{f}_{BL}) = \Phi - \Phi \Lambda' \Sigma^{-1} \Lambda \Phi = \Phi - CD_{BL}^2. \quad (10)
\]

As the original factors are usually z-standardized, equation 10 implies that for corresponding factors one can compute
\[
diag(MSE(\hat{f}_{BL})) = I - CD_{BL}^2. \quad (11)
\]

If \(\hat{f}_{PV} \neq \hat{f}_{BL}\), as might occur in the sample, it follows from Krijnen et al. (1996, p. 3016, Equation 4) that MSE(\(\hat{f}_{PV}\)) for corresponding \(\hat{f}_{PV}\) and \(f\) can be computed as
\[
diag(MSE(\hat{f}_{PV})) = diag(E(\hat{f}_{PV}' \hat{f}_{PV}) - E(\Phi \Lambda' \Sigma^{-1} x \hat{f}_{PV}) - E(\hat{f}_{PV}' x' \Sigma^{-1} \Lambda \Phi) + \Phi). \quad (12)
\]

As for equations (2) and (4), the similarity of the results from equations (11) and (12) depends on the similarity of \(\hat{f}_{BL}\) and \(\hat{f}_{PV}\), which should increase with sample size.

Simulation study 1 was performed to compare the coefficient of determinacy computed from equation (7) with the correlation of the plausible values for a different number of imputations (starting from single imputation, i.e., a single plausible value to a large number of aggregates as a basis for mean plausible values). Simulation study 2 was performed to compare the effect of sample size and model fit on the coefficient of determinacy computed from equations (5) and (7) and on the MSE computed from equations (10) and (11). As the generating factor scores are available in a simulation study, the results of equations (5) and (7) can be compared with the correlation of \(\hat{f}_{PV}\) with \(f\) and the results of equations (11) and (12) can be compared with the difference of \(\hat{f}_{PV}\) and \(f\). As a related issue, the correlation of \(\hat{f}_{PV}\) with \(\hat{f}_{BL}\) will also be computed.

### Simulation Study 1: Number of Imputations

**Model Specification and Estimation**

Data generation was performed with IBM SPSS, Version 26 and BFA was performed with Mplus 8.6. Population models were based on \(q = 2\) common factors, three low \((l = .30)\) and three high \((l = .50)\) salient loadings per factor. All nonsalient loadings were zero (independent clusters model). We used population models with uncorrelated factors as well as population models with correlated common factors, \(\phi \in \{.00, .40, .50, .60\}\). As the effect of sampling error was not investigated in this study, for
each of the four populations a single sample with the maximum possible sample size for plausible values based on BFA with Mplus 8.6, that is, \( n = 10,000 \) was generated. Observed variables were computed from normally distributed common factors \( f \sim \mathcal{N}(0,1) \) and error factors \( e \sim \mathcal{N}(0,1) \). To investigate the effect of model misspecification, the models were correctly specified with freely estimated salient loadings, and they were misspecified by means of an equality constraint for the salient loadings of each factor. Nonsalient loadings were estimated with normally distributed priors with a zero mean and a variance of \( \sigma^2 = 0.01 \). The number of imputations for mean plausible values was \( i = 1, 2, 4, 8, 16, 32, 64, 128, 256, \) and 512. A total of 512 single plausible values was used for \( i = 1 \), so that the 512 determinacies of the single plausible values were averaged. For \( i = 2 \), two-plausible values were averaged to compute the mean plausible values, so that 256 determinacies were computed and averaged. Accordingly, 128 determinacies of mean plausible values based on \( i = 4 \) were averaged. Finally, for \( i = 512 \), all plausible values were averaged to form the mean plausible value. Determinacies were computed as the correlation of the plausible values with generating factors (GDetPV), the coefficient of determinacy according to equation 7 based on model parameters and parameters of the plausible values (CDPV), and as a reference value, which does not depend on the number of imputations, the coefficient of determinacy according to equation 5 (CDBL).

**Results**

The results of the model fit as well as the factor-intercorrelations based on the population, model estimates, the mean correlation of single plausible values, and the best linear predictor are given in Table 1. As intended, the fit of the misspecified models was moderate to acceptable to provide models that would generally be accepted as a basis for the computation of factor score estimates of plausible values for secondary analysis (see Table 1). Like in Asparouhov and Muthén (2010a) the intercorrelation of the factor in the model could be estimated by the mean correlation of single plausible values, whereas the best linear predictor resulted in a substantial overestimation of the correlation. This effect occurred in correctly specified models and in misspecified models which indicates that the effect does not depend on model fit.

The effect of the number of imputations on the correlation of the plausible values with the generating factor (GDetPV), the related determinacy coefficient (CDPV), and the coefficient of determinacy computed for the best linear predictor (CDBL) is given in Figure 1A to 1D. In all conditions, more than 64 imputations were necessary for GDetPV and CDPV to approach CDBL. The values of GDetPV and CDPV were very similar across all conditions. There was no relevant effect of model misfit and interfactor correlations indicating that CDPV can be used as an estimate for GDetPV in empirical studies, when the generating factor scores are not available.
Figure 1. Correlations of Plausible Values with Population Factors (GDetPV), Coefficient of Determinacy for Plausible Values (CDPV), and Coefficient of Determinacy for the Best Linear Predictor (CDBL) for an Increasing Number of Imputations (i) in Factor Models Based on Uncorrelated Population Factors (A), Population Factor Intercorrelations of $\phi = .40$ (B), Population Factor Intercorrelations of $\phi = .50$ (C), and Population Factor Intercorrelations of $\phi = .60$ (D).
Simulation Study 2: Sample Size and Model Misfit

Model Specification and Estimation

In the first condition, population models were based on \( p \in \{3, 5\} \) common factors, low and high salient loadings \( l \in \{.40, .80\} \), and six variables with salient loadings per factor \( p/q = 6 \). Moreover, all nonsalient loadings were zero (independent clusters model), and we used population models with uncorrelated factors as well as population models with correlated common factors, \( \phi \in \{.00, .40\} \). Small and large sample sizes were investigated, \( n \in \{200, 800\} \). For each of these \( 2(p) \times 2(l) \times 2(\rho) \times 2(n) = 16 \) conditions, 1,000 samples of normally distributed observed variables computed from normally distributed common factors \( f \sim N(0,1) \) and error factors \( e \sim N(0,1) \) were drawn and submitted to BFA.

Again, IBM SPSS, Version 26 was used for data generation, and Mplus 8.6 was used for BFA. In the first condition, all salient loadings were equal and the model specification of all models was in line with the population models, that is, there was no effect of model misfit. In this condition, a single BFA for each data set was specified. The models were based on unit-variances of the factors, freely estimated salient loadings and on nonsalient loadings estimated with normally distributed priors with a zero mean and a variance of \( \sigma^2 = 0.01 \) (see Appendix for Mplus input example). In a second condition, we investigated the effect of model misfit for \( q = 3 \), low and moderate salient loadings \( l \in \{.30, .50\} \), \( n \in \{200, 800\} \), and \( \phi \in \{.00, .40\} \). The loading pattern of these models is given in the Appendix (Table A1). Correctly specified models were again based on unit-variances of the factors, freely estimated salient loadings and on nonsalient loadings estimated with normally distributed priors with a

| Model | \( \phi_{pop} \) | \( \phi_{mod} \) | \( \phi_{PV} \) | \( \phi_{BL} \) |
|-------|----------------|----------------|----------------|----------------|
| Correctly specified models |
| 1     | 1.000 \( \times \) .000 | 1.000 .000 | .00 .00 .00 .00 |
| 2     | 1.000 .000 | 1.000 .40 | .40 .39 .54 |
| 3     | 1.000 .000 | 1.000 .50 | .50 .49 .67 |
| 4     | 1.000 .000 | 1.000 .60 | .60 .59 .78 |
| Misspecified models |
| 5     | \( <.001 \) .033 | .904 .00 | .00 .00 .00 .00 |
| 6     | \( <.001 \) .026 | .946 .40 | .56 .57 .70 |
| 7     | \( <.001 \) .023 | .960 .50 | .65 .65 .79 |
| 8     | \( <.001 \) .019 | .974 .60 | .72 .72 .87 |

Note. * \( \chi^2 \)-based posterior predictor \( p \) value. RMSEA = root mean square error of approximation; CFI = comparative fit index.
zero mean and a variance of $\sigma^2 = 0.01$. Misspecified models were based on an equality constraint for the salient loadings as well as for the error terms (see Appendix for Mplus input example).

Zitzmann and Hecht (2019) emphasized the relevance of the precision of BFA estimates. They propose the effective sample size (ESS) as an indicator for the precision of parameter estimation. As Mplus 8.6 does not calculate ESS, but the potential scale reduction (PSR), the PSR has to be transformed into ESS. Here, we use an ESS of 100 samples resulting in a PSR of 1.01 (Zitzmann and Hecht, 2019) and in a BCONVERGENCE option of 0.01, which is smaller than the Mplus default of 0.05. To ensure convergence, we used the BITERATIONS option to perform 100,000 iterations when necessary. The mean plausible values were computed from 1,000 imputations, which according to the previous simulation study should be sufficiently large to approach optimal determinacy of mean plausible values.

We computed the following dependent variables to evaluate the coefficients of determinacy: The expected value of determinacy, that is, correlation of the mean plausible values with the generating factor scores ($GDet_{PV}$), the coefficient of determinacy according to equation 7 based on model parameters and parameters of the mean plausible values ($CD_{PV}$), the coefficient of determinacy according to equation 5 based solely on model parameters ($CD_{BL}$), and expected value for the determinacy of the best linear predictor, that is, the correlation of the best linear predictor computed from equation 4 with the generating factor scores ($GDet_{BL}$). The mean squared difference of $GDet_{PV}$ and $CD_{PV}$ and the mean squared difference of $GDet_{BL}$ and $CD_{BL}$ was calculated to evaluate the precision of the coefficients when compared to the expected determinacies computed from the generating factor scores. The bias-correction proposed by Budescu (1982) was performed on the means of $CD_{BL}$.

The following dependent variables were computed to evaluate the MSE of the mean plausible values: The expected value for the MSE, that is, the mean squared difference of the mean plausible values and the generating factor scores ($GMSE_{PV}$), the MSE according to equation 12 based on model parameters and parameters of the mean plausible values ($MSE_{PV}$), the MSE according to equation 11 based solely on model parameters ($MSE_{BL}$), and the MSE as the mean squared difference between the best linear predictor computed from equation 1 and the generating factor scores ($GMSE_{BL}$). The mean squared difference of expected value $GMSE_{PV}$ and the estimated $MSE_{PV}$ and the mean squared difference of the expected value $GMSE_{BL}$ and $MSE_{BL}$ was calculated to evaluate the precision of the estimated MSE.

**Results**

The average fit of the correctly specified models was very good (comparative fit index [CFI] $\geq .98$). The precision of the coefficients of determinacy computed from plausible value parameters and model parameters ($CD_{PV}$, equation 7) and solely from model parameters ($CD_{BL}$, equation 5) was similar (see Table 2). In line with this result, the mean correlation of $f_{PV}$ with $f_{BL}$ was greater than .99 in this condition. It
should be noted that the mean determinacy is substantially overestimated by CDPV as well as CDBL in the conditions based on small sample sizes and small salient loadings. Overall, the correction proposed by Budescu (1982) reduced the bias of CDBL considerably although it does not completely eliminate the bias for \( q = 3 \) (see Table 2).

The results are similar for MSE, where precision of the computation from plausible values parameters and model parameters (MSEPV, equation 12) and the computation solely from model parameters (MSEBL, equation 11) was similar (see Table 3). The mean MSE is underestimated in the condition based on small sample sizes and small salient loadings.

The results for the condition for the comparison of correctly specified and misspecified models are given in Table 4. The misspecified models had a substantially reduced model fit. Some misspecified models had even an unacceptably low model fit. The results were nevertheless of interest because they allow to explore extreme effects of model misfit on determinacy. The correlation of the mean plausible values with the factors (\( GDet_{PV} \)) was slightly reduced when misfit occurred as were the plausible value-based determinacy coefficient (CDPV) and the model parameter-based determinacy coefficient (CDBL). Although model misfit had a negative impact on the determinacy of the mean plausible values, its effect on the two coefficients was similar. Even when the determinacy was very low, CDPV and CDBL were

Table 2. Expected Values and Mean Squared Errors of Coefficients of Factor Score Determinacy.

| \( \phi \) | \( q \) | \( l \) | \( n \) | \( GDet_{PV} \) | \( CD_{PV} \) | MSE (CDPV) | \( GDet_{BL} \) | \( CD_{BL} \) | MSE(CDBL) | \( CD_{BL, \text{cor}} \) |
|---|---|---|---|---|---|---|---|---|---|---|
| .0 | 3 | .4 | 200 | .707 | .755 | .004 | .708 | .748 | .003 | .719 |
| 800 | .722 | .738 | .001 | .723 | .739 | .001 | .733 |
| .8 | 200 | .955 | .950 | .000 | .955 | .959 | .000 | .956 |
| 800 | .956 | .953 | .000 | .956 | .957 | .000 | .957 |
| 5 | .4 | 200 | .702 | .768 | .006 | .704 | .756 | .004 | .706 |
| 800 | .719 | .743 | .001 | .720 | .745 | .001 | .734 |
| .8 | 200 | .955 | .948 | .000 | .955 | .961 | .000 | .955 |
| 800 | .956 | .952 | .000 | .956 | .959 | .000 | .957 |
| .4 | 3 | .4 | 200 | .732 | .768 | .003 | .732 | .762 | .002 | .736 |
| 800 | .745 | .755 | .000 | .745 | .757 | .000 | .750 |
| .8 | 200 | .956 | .952 | .000 | .956 | .960 | .000 | .956 |
| 800 | .957 | .955 | .000 | .957 | .958 | .000 | .957 |
| 5 | .4 | 200 | .738 | .782 | .003 | .738 | .772 | .003 | .726 |
| 800 | .753 | .766 | .000 | .754 | .768 | .000 | .758 |
| .8 | 200 | .956 | .950 | .000 | .956 | .962 | .000 | .955 |
| 800 | .957 | .954 | .000 | .957 | .959 | .000 | .958 |

Note. Population = independent clusters model with six salient loadings per factor, \( q \) = number of factors, \( l \) = salient loading size, \( n \) = sample size; \( \phi \) = factor inter-correlation; 1,000 samples per condition; \( PV \) = mean plausible values of 1,000 imputations per sample; \( BL \) = best linear factor score estimate; \( GDet \) = correlation between estimates and true scores; \( CD \) = coefficient of determinacy (cf. equations 5 and 7); \( MSE \) = mean squared error, \( CD_{BL, \text{cor}} \) = bias-corrected, conventional coefficient of determinacy.
### Table 3. Expected Values and Mean Squared Errors of MSE Estimates.

| $\phi$ | $q$ | $l$ | $n$ | $\text{GMSE}_{PV}$ | $\text{MSE}_{PV}$ | $\text{MSE} (\text{GMSE}_{PV})$ | $\text{GMSE}_{BL}$ | $\text{MSE}_{BL}$ | $\text{MSE} (\text{GMSE}_{BL})$ |
|--------|------|-----|-----|-------------------|------------------|-----------------|------------------|------------------|-------------------|
| .0     | 3    | .4  | 200 | .500             | .432             | .010            | .501            | .440             | .009              |
|        |      |     | 800 | .479             | .456             | .002            | .478            | .453             | .002              |
| .8     | 200  | .095| .098| .000             | .094             | .080            | .091            | .088             | .083              |
|        |      |     | 800 | .088             | .091             | .000            | .088            | .083             | .000              |
| 5      | .4   | 200 | .508| .416             | .014             | .511            | .427            | .013             |
|        |      |     | 800 | .483             | .449             | .002            | .482            | .444             | .003              |
| .8     | 200  | .098| .104| .000             | .094             | .076            | .080            | .081             | .000              |
|        |      |     | 800 | .089             | .095             | .000            | .088            | .081             | .000              |
| .4     | 3    | .4  | 200 | .464             | .412             | .007            | .466            | .418             | .007              |
|        |      |     | 800 | .445             | .430             | .001            | .445            | .427             | .001              |
| .8     | 200  | .093| .096| .000             | .092             | .079            | .086            | .082             | .000              |
|        |      |     | 800 | .086             | .088             | .000            | .086            | .082             | .000              |
| 5      | .4   | 200 | .457| .393             | .008             | .459            | .403            | .008             |
|        |      |     | 800 | .434             | .414             | .001            | .432            | .410             | .001              |
| .8     | 200  | .095| .100| .000             | .092             | .075            | .085            | .080             | .000              |
|        |      |     | 800 | .086             | .090             | .000            | .086            | .080             | .000              |

Note. Population = independent clusters model with six salient loadings per factor, $q =$ number of factors, $l =$ salient loading size, $n =$ sample size; $\phi =$ factor inter-correlation; 1,000 samples per condition; $PV =$ mean plausible values of 1,000 imputations per sample; $BL =$ best linear factor score estimate; $MSE =$ mean squared error; $\text{MSE}_{PV}, \text{MSE}_{BL} =$ estimated MSE of score estimates (cf. equations 11 and 12).

### Table 4. Expected Values and Mean Squared Errors of Coefficients of Factor Score Determinacy for Correctly Specified and Corresponding Misspecified Models (Underlying Population Matrix with $q = 3$ Factors and Salient Loadings $l_1 = .5, l_2 = .3$).

#### Correctly specified

| $\phi$ | $n$ | RMSEA | CFI  | $GDet_{PV}$ | $CD_{PV}$ | $\text{MSE}(\text{CD}_{PV})$ | Det(BL) | $CD_{BL}$ | $\text{MSE}(\text{CD}_{BL})$ | $CD_{BL, \text{cor}}$ |
|--------|-----|-------|------|-------------|-----------|-------------------------------|---------|-----------|-------------------------------|---------------------|
| .0     | 200 | .010  | .980 | .761        | .795      | .002                          | .762    | .792      | .002                          | .770                |
|        | 800 | .004  | .995 | .772        | .782      | .000                          | .773    | .785      | .000                          | .780                |
| .4     | 200 | .009  | .982 | .780        | .806      | .001                          | .780    | .804      | .001                          | .783                |
|        | 800 | .004  | .996 | .790        | .796      | .000                          | .790    | .799      | .000                          | .794                |

#### Misspecified

| $\phi$ | $n$ | RMSEA | CFI  | $GDet_{PV}$ | $CD_{PV}$ | $\text{MSE}(\text{CD}_{PV})$ | Det(BL) | $CD_{BL}$ | $\text{MSE}(\text{CD}_{BL})$ | $CD_{BL, \text{cor}}$ |
|--------|-----|-------|------|-------------|-----------|-------------------------------|---------|-----------|-------------------------------|---------------------|
| .0     | 200 | .027  | .892 | .754        | .764      | .001                          | .757    | .776      | .002                          | .752                |
|        | 800 | .026  | .915 | .762        | .763      | .000                          | .763    | .769      | .000                          | .764                |
| .4     | 200 | .026  | .883 | .773        | .778      | .001                          | .775    | .789      | .001                          | .766                |
|        | 800 | .027  | .917 | .781        | .779      | .000                          | .782    | .784      | .000                          | .779                |

Note. $\phi =$ factor inter-correlation; $n =$ sample size; 1,000 samples per condition; $PV =$ mean plausible values of 1,000 imputations per sample; $BL =$ best linear factor score estimate; $GDet =$ correlation between estimates and true scores; $CD =$ coefficient of determinacy (cf. equations 5, 7); $MSE =$ mean squared error, $CD_{BL, \text{cor}} =$ bias-corrected, conventional coefficient of determinacy.
similar. This indicates that both coefficient of determinacy can be used under conditions of model misfit to estimate determinacy. It should also be noted that determinacy can also be low when model fit is excellent.

**Discussion**

As plausible values have been proposed in the context of BFA (Asparouhov & Muthén, 2010a), they might be used as covariates or predictors in research settings or to provide individual scores for the BFA factors in applied settings. Accordingly, studies of the validity of plausible values are needed. Studies on plausible values ascertained the validity of the scores by means of the correlation or squared difference of the plausible values and the generating factor scores (Asparouhov & Muthén, 2010a; Luo & Dimitrov, 2018). However, the generating factor scores are unknown in empirical studies so that the coefficient of determinacy has to be calculated to evaluate the validity of factor score estimates (Ferrando & Lorenzo-Seva, 2018). The coefficient of determinacy is solely based on model parameters and can be computed whenever BFA is performed in an empirical setting. Therefore, the first aim of this study was to compare the coefficient of determinacy based on BFA model parameters with the correlation of mean plausible values with the generating factors.

As the mean correlation of single plausible values of a factor with the single plausible values of another factor is a helpful estimator of the population factor intercorrelation (Asparouhov & Muthén, 2010a), the validity of single plausible values as predictors could be of interest. Therefore, a coefficient of determinacy that can be used for single plausible values as well as for mean plausible values was proposed. Accordingly, the second aim of the study was the investigation of the determinacy of single plausible values and of the effect of the number of imputations and model misfit on the correlation of mean plausible values with the population factors and on the plausible value-based determinacy coefficient.

It was investigated by means of a first simulation study whether the plausible value-based coefficient of determinacy can be used to estimate the determinacy of single plausible values as well as mean plausible values when the population factors are not available. Therefore, the first simulation study compared the correlation of single plausible values and of mean plausible values based on a different number of imputations with the population factor with the plausible value-based coefficient of determinacy. It was found that the plausible value-based determinacy coefficient was an acceptable estimator of the correlation of single plausible values with the respective factor and for the effect of the number of imputations on the correlation of mean plausible values with the respective factor. With an increasing number of imputations, the determinacy of the mean plausible values increased and approximated the determinacy of the best linear predictor. With less than 20 imputations, the determinacy of the mean plausible values was substantially smaller than the determinant of the best linear predictor. It follows that the comparison of the plausible value-based determinacy coefficient with the conventional determinacy coefficient may be used
to determine the number of imputations needed to reach acceptable validity of the mean plausible values. Moreover, the precision of the plausible value-based determinacy coefficient was also acceptable for misspecified models with some amount of misfit and for models based on different factor intercorrelations.

The first simulation study showed that the plausible value-based determinacy coefficient approximates the conventional, model parameter-based determinacy coefficient with a large number of imputations. As the first simulation study did not investigate the effect of sampling error, the second simulation study investigated whether the plausible value-based determinacy coefficient based on 1,000 imputations approximates the model parameter-based determinacy coefficient in small ($n = 200$) and large ($n = 800$) samples. Both determinacy coefficients were very similar even in small samples, even in conditions with model misspecification. This indicates that the plausible value-based determinacy coefficient can be used for the estimation of determinacy in empirical studies. It was, however, also found that the correlation of the mean plausible values with the generating factors, the plausible value-based determinacy coefficient and the model parameter-based determinacy coefficient were slightly reduced in the misspecified models compared to the corresponding correctly specified models. This indicates that determinacy may be reduced when it is based on substantially misspecified models. Nevertheless, both determinacy coefficients correctly indicated the reduction of determinacy. It should also be noted that determinacy was also very low for some models with excellent fit. This underlines that whenever plausible values are used, their determinacy should be reported.

It should be noted that for small sample sizes and a small salient loading size, both coefficients of determinacy slightly overestimate the validity of mean plausible values. The correction proposed by Budescu (1982) reduced the bias of the conventional coefficient of determinacy considerably. It might therefore be recommended to use the bias-correction to estimate the validity of mean plausible values especially when sample sizes are small. However, it is recommended to report the uncorrected as well as the bias-corrected values so that the amount of bias can also be considered.

Since the similarity of factor scores and plausible values has already been evaluated by means of the MSE, the MSE was also considered. It was shown that for mean-centered best linear predictors, the MSE can directly be computed from the conventional coefficient of determinacy. However, an MSE version based on the mean plausible values was also proposed. The results of the simulation study reveal that there was no substantial difference between the two versions of MSE. A bias was also found for small samples and small salient loadings resulting in an underestimation of MSE in this condition. One may consider to calculate the conventional MSE from bias-corrected determinacy coefficients to reduce the bias of MSE.

As a limitation, we note that the means in the population of the scores were zero which has enhanced the similarity of the coefficients of determinacy and the MSE. Effects of non-zero means remain to be investigated. A further limitation refers to the conditions of the simulation study. A large number of factors, different numbers of
variables with salient loadings per factor, cross-loadings, and different prior variances for the zero-loadings are aspects that remain to be investigated. It should, however, be noted that we followed the recommendation of Zitzmann and Hecht (2019) in enhancing the precision of BFA by means of a smaller convergence criterion and a larger number of iterations than the default options of Mplus 8.6. This resulted in a substantial increase of computation time.

Nevertheless, the tentative conclusion of this study is that the plausible value-based determinacy coefficient can be used for the estimation of determinacy of single plausible values and for the determinacy of mean plausible values. When the plausible value-based determinacy coefficient is substantially smaller than the model parameter-based determinacy coefficient one may consider to increase the number of imputations for mean plausible values. A further conclusion is that the bias-correction proposed by Budescu (1982) may help to improve the precision of determinacy coefficients. These conclusions are relevant for the use of sets of single plausible values as predictors for secondary analysis and for the use of mean plausible values in empirical studies when the generating factor scores are unknown.

Appendix

Table A1. Sample-Based Simulation, Second Condition: Population Loading Matrix With $l_1 = .5$ and $l_2 = .3$.

| Item | $F_1$ | $F_2$ | $F_3$ |
|------|-------|-------|-------|
| x1   | .5    | —     | —     |
| x2   | .5    | —     | —     |
| x3   | .5    | —     | —     |
| x4   | .5    | —     | —     |
| x5   | .3    | —     | —     |
| x6   | .3    | —     | —     |
| x7   | —     | .5    | —     |
| x8   | —     | .5    | —     |
| x9   | —     | .5    | —     |
| x10  | —     | .5    | —     |
| x11  | —     | .3    | —     |
| x12  | —     | .3    | —     |
| x13  | —     | —     | .5    |
| x14  | —     | —     | .5    |
| x15  | —     | —     | .5    |
| x16  | —     | —     | .5    |
| x17  | —     | —     | .3    |
| x18  | —     | —     | .3    |
Example for Mplus input for the sample-based simulation (correctly specified models):

DATA: FILE = "sample.csv";
VARIABLE: NAMES ARE x1-x18;
USEVARIABLES ARE x1-x18;
ANALYSIS: ESTIMATOR = BAYES;
BITERATIONS = 100,000;
BCONVERGENCE = 0.01;
MODEL: F1 by x1*x2-x18 (f1p1-f1p18);
F2 by x1*x2-x18 (f2p1-f2p18);
F3 by x1*x2-x18 (f3p1-f3p18);
F1@1; F2@1; F3@1;
MODEL PRIORS: f1p7-f1p18~N(0,0.01);
f2p1-f2p6~N(0,0.01);
f2p13-f2p18~N(0,0.01);
f3p1-f3p12~N(0,0.01);
OUTPUT: STANDARDIZED (STDXY);
SAVEDATA: FILE = BAY_PV.dat;
SAVE = FSCORES (1000);
RESULTS ARE BAY_RESULTS.dat;

Example for Mplus input for the sample-based simulation (misspecified models):

DATA: FILE = "sample.csv";
VARIABLE: NAMES ARE x1-x18;
USEVARIABLES ARE x1-x18;
ANALYSIS: ESTIMATOR = BAYES;
BITERATIONS = 100000;
BCONVERGENCE = 0.01;
MODEL: F1 by x1@1;
F1 by x2*x3-x6 (Lf1);
F1 by x7*x8-x12 (sLf1_2);
F1 by x13*x14-x18 (sLf1_3);
F2 by x1*x2-x6 (sLf2_1);
F2 by x7@1;
F2 by x8*x9-x12 (Lf2);
F2 by x13*x14-x18 (sLf2_3);
F3 by x1*x2-x6 (sLf3_1);
F3 by x7*x8-x12 (sLf3_2);
F3 by x13@1;
F3 by x14*x15-x18 (Lf3);
x2-x6 (e1);
x8-x12 (e2);
x14-x18 (e3);
MODEL PRIORS: sLf1_2~N(0,0.01);
sLf1_3~N(0,0.01);
sLf2_1~N(0,0.01);
sLf2_3~N(0,0.01);
sLf3_1~N(0,0.01);
sLf3_2~N(0,0.01);
OUTPUT: STANDARDIZED (STDYX);
SAVEDATA: FILE = BAY_PV.dat;
SAVE = FSCORES (1000);
RESULTS ARE BAY_RESULTS.dat;

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
André Beauducel https://orcid.org/0000-0002-9197-653X

References
Asparouhov, T., & Muthén, B. (2010a). Plausible values for latent variables using Mplus [Technical Report]. www.statmodel.com/download/Plausible.pdf
Asparouhov, T., & Muthén, B. (2010b). Multiple imputation with Mplus [Technical Report]. www.statmodel.com/download/Imputations7.pdf
Asparouhov, T., Muthén, B., & Morin, A. J. S. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances: Comments on Stromeyer et al. Journal of Management, 41, 1561–1577. https://doi.org/10.1177/0149206315591075
Beauducel, A., & Hilger, N. (2015). Extending the debate between Spearman and Wilson 1929: When do single variables optimally reproduce the common part of the observed covariances? Multivariate Behavioral Research, 50, 555–567. https://doi.org/10.1080/00273171.2015.1059311
Bonafede, M., Chiorri, C., Azzolina, D., Marinaccio, A., Migliore, E., Mensi, C., Chellini, E., Romeo, E., Grosso, F., Franzoi, I.G., Granieri, A., & Guglielmucci, F. (2021). Preliminary validation of a questionnaire assessing psychological distress in caregivers of patients with malignant mesothelioma: Mesothelioma psychological distress tool-caregivers. Psycho-Oncology, 31(1), 122–129. https://doi.org/10.1002/pon.5789
Budescu, D. V. (1982). The estimation of factor indeterminacy. Educational and Psychological Measurement, 43, 971–976. https://doi.org/10.1177/001316448304300404
Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement, 78*, 762–780. https://doi.org/10.1177/0013164417719308

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6*, 430–450. https://doi.org/10.1037/1082-989X.6.4.430

Guttman, L. (1955). The determinacy of factor score matrices with applications for five other problems of common factor theory. *British Journal of Statistical Psychology, 8*, 65–82. https://doi.org/10.1111/j.2044-8317.1955.tb00321.x

Krijnen, W. P., Wansbeek, T. J., & Ten Berge, J. M. F. (1996). Best linear predictors for factor scores. *Communications in Statistics, 25*, 3013–3025. https://doi.org/10.1080/03610929608831883

Luo, Y., & Dimitrov, D. M. (2018). A short note on obtaining point estimates of the IRT ability parameter with MCMC estimation in Mplus: How many plausible values are needed? *Educational and Psychological Measurement, 79*, 272–287. https://doi.org/10.1177/0013164418777569

MacCallum, R. C. (2003). 2001 Presidential address: Working with imperfect models. *Multivariate Behavioral Research, 38*, 113–139. https://doi.org/10.1207/S15327906MBR3801_5

MacCallum, R. C., Browne, M. W., & Cai, L. (2007). Factor analysis models as approximations. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100* (pp. 153–176). Taylor and Francis (Kindle-Version).

McDonald, R. P. (1981). Constrained least squares estimators of oblique common factors. *Psychometrika, 46*, 337–341. https://doi.org/10.1007/BF02293740

McDonald, R. P. (2011). Measuring latent quantities. *Psychometrika, 76*, 511–536. https://doi.org/10.1007/S11336-011-9223-7

Mulaik, S. A. (2010). *Foundations of factor analysis* (2nd ed.). CRC Press.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods, 17*, 313–335. https://doi.org/10.1037/a0026802

Smid, S. C., & Winter, S. D. (2020). Dangers of the defaults: A tutorial on the impact of default priors when using Bayesian SEM with small samples. *Frontiers in Psychology, 11*, Article 611963. https://doi.org/10.3389/fpsyg.2020.611963

Stromeyer, W. R., Miller, J. W., Sriramachandramurthy, R., & DeMartino, R. (2015). The prowess and pitfalls of Bayesian structural equation modeling: Important considerations for management research. *Journal of Management, 41*, 491–520. https://doi.org/10.1177/0149206314551962

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

Wang, Y., Wang, Y., Stein, D., Liu, Q., & Chen, W. (2021). The structure of Chinese beginning online instructors’ competencies: Evidence from Bayesian factor analysis. *Journal of Computers in Education, 8*, 411–440. https://doi.org/10.1007/s40692-021-00186-9

Weide, A. C., Scheuble, V., & Beauducel, A. (2021). Bayesian and Maximum-Likelihood Modeling and Higher-Level Scores of Interpersonal Problems with Circumplex Structure. *Frontiers in Psychology, 12*, 4947. https://doi.org/10.3389/fpsyg.2021.761378

Zitzmann, S., & Hecht, M. (2019). Going beyond convergence in Bayesian estimation: Why precision matters too and how to assess it. *Structural Equation Modeling, 26*, 646–661. https://doi.org/10.1080/10705511.2018.1545232