

# How Does the Elimination of Group Mean-Differences Affect Factor Score Determinacy?

André Beauducel<sup>1</sup>, & Norbert Hilger<sup>1</sup>

<sup>1</sup> Department of Psychology, University of Bonn, Bonn, Germany

Correspondence: André Beauducel, Department of Psychology, University of Bonn, Kaiser-Karl-Ring 9, 53111 Bonn, Germany

Received: December 4, 2023 Accepted: January 29, 2024 Online Published: February 28, 2024

doi:10.5539/ijsp.v13n1p1

URL: <https://doi.org/10.5539/ijsp.v13n1p1>

## Abstract

The present study investigates to what degree the common variance of the factor score predictor with the original factor, i.e., the determinacy coefficient or the validity of the factor score predictor, depends on the mean-difference between groups. When mean-differences between groups in the factor score predictor are eliminated by means of covariance analysis, regression, or group specific norms, this may reduce the covariance of the factor score predictor with the common factor. It is shown that in a one-factor model with the same group mean-difference on all observed variables, the common factor cannot be distinguished from a common factor representing the group mean-difference. It is also shown that for common factor loadings equal or larger than .60, the elimination of a  $d = .50$  mean-difference between two groups in the factor score predictor leads to only small decreases of the determinacy coefficient. A compensation-factor  $k$  is proposed allowing for the estimation of the number of additional observed variables necessary to recover the size of the determinacy coefficient before elimination of a group mean-difference. It turns out that for factor loadings equal or larger than .60 only a few additional items are needed in order to recover the initial determinacy coefficient after the elimination of moderate or large group mean-differences.

**Keywords:** factor score predictor, determinacy coefficient, effect size, group mean-difference

## 1. Introduction

### 1.1 Factor Score Predictors and Group Mean-Differences

Factor score predictors are computed in order to provide individual scores for a factor resulting from factor analysis (DiStefano, Zhu, & Mindrila, 2009). These scores can be used in applied settings, e.g., job-selection. The validity of factor score predictors depends on their correlation with the corresponding factors, which is usually termed determinacy coefficient (Gorsuch, 1983; Grice, 2001). For orthogonal factors and block-diagonal patterns of factor loadings, the reliability of the regression factor score predictor also corresponds to the determinacy coefficient (Beauducel, Harms, & Hilger, 2016). However, under more general conditions, the correlation of the factor score predictor with the corresponding factor is an indicator of validity.

Several group mean-differences may co-occur with the individual differences represented by a common factor. Some group mean-differences may be estimated a priori in the context of multiple-group factor analysis, which leads to the issue of measurement invariance of factors across groups. Measurement invariance refers to the similarity of model parameters across groups, an assumption required for the interpretation of group differences on the factors (Vandenberg, & Lance, 2000). New methods that may allow for the improvement of measurement invariance have meanwhile been proposed (Asparouhov, & Muthén, 2023). However, not all relevant group mean-differences may be considered right from the start when factor analysis is performed. A reason for this could be that sample size leads to a limitation of the number of groups that can be considered in multiple-group factor analysis. Another reason could be that a grouping of individuals is obtained afterwards (e.g., by means of latent class analysis). A third reason could be that group mean-differences that appear of irrelevant size in the observed variables, can be of relevant size in the factor score predictors. Group mean-differences that are not represented by a model parameter for the means, may affect the covariances of observed variables, and may be present in the factor score predictor. In these cases, an elimination of group mean-differences may be performed a posteriori by means of regression analysis, covariance analysis, or by means of group-specific test norms or test standardization (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). The elimination of the group mean-differences may affect the validity of the factor score predictor, i.e., the factor score determinacy. Therefore, the

present study investigates the effect of eliminating group mean-differences on the resulting determinacy coefficient.

### 1.2 Specifications of the Study

In principle, different types of factor score predictors, different numbers of groups, different factor models, and different score distributions might be of interest in this context. However, in order to provide initial results that could be the starting point for further investigations, the present study is based on only two groups and on the determinacy coefficient computed for the best-linear factor score predictor (Krijnen, Wansbeek, & Ten Berge, 1996), which is sometimes termed regression factor score predictor. Moreover, the study is based on one factor models with equal factor loadings. These simplifications allow for a direct investigation of effects of the size of group mean-differences in terms of different effect sizes  $d$  (Cohen, Cohen, West, & Aiken, 2003) and the number of variables on the determinacy coefficient.

First, a population factor model for the description of the effect of group mean-differences on factors and model parameters is defined. On this basis, the effect of factor loadings and the number of variables on the determinacy coefficient is investigated. Benchmarks for factor loadings combined with a given number of observed variables resulting in reasonable determinacy coefficients are provided. Using these benchmarks, the effect of group mean-differences and determinacy coefficients is analyzed. It is proposed to calculate the number of additional items necessary to compensate for reduced determinacy resulting from eliminating group mean-differences. Finally, theoretical and practical implications of the results are discussed.

## 2. Definitions

In the population of individuals, the common factor model with an additional factor representing the means between two groups can be defined as

$$\mathbf{x} = \Lambda_{\xi} \xi + \Lambda_{\mathbf{g}} \mathbf{g} + \Psi \boldsymbol{\varepsilon}, \tag{1}$$

where  $\mathbf{x}$  is a standardized random vector of  $p$  observed variables,  $\xi$  is a random vector of  $q$  common factors, with  $E(\xi\xi') = \Phi$  and  $diag(\Phi) = \mathbf{I}_{q \times q}$  and  $\Lambda_{\xi}$  is the  $p \times q$  matrix of common factor loadings. Moreover,  $\mathbf{g}$  is a random vector representing the means of two groups, with  $E(\mathbf{g}\mathbf{g}') = \mathbf{1}$  for  $q = 1$  (for  $q > 1$  we get a  $q \times q$  identity matrix),  $E(\mathbf{g}\xi') = \mathbf{0}$ , and  $\Lambda_{\mathbf{g}}$  is the  $p \times 1$  matrix of loadings of the observed variables on the factor representing the standardized group mean-difference  $d$  (e.g., Cohen et al., 2003). For equal group size  $E(\mathbf{g}\mathbf{g}') = \mathbf{1}$  and  $E(\mathbf{g}) = \mathbf{0}$   $\mathbf{g}$  corresponds to an unweighted effect-coded indicator variable. For two groups, unweighted effect-coding means that the resulting indicator variable takes the value +1 for one group and -1 for the other group (Cohen et al., 2003). The loadings  $\Lambda_{\mathbf{g}}$  represent the correlation of the observed variables with  $\mathbf{g}$ . Therefore, the effect size  $d$  is converted into the correlation  $r$  according to Ruscio (2008, Table 1). For groups of equal size, the conversion of  $d$  into a correlation  $r$  is  $r = d(4 + d^2)^{-1/2}$ , so that the loadings are given by  $\Lambda_{\mathbf{g}} = d(4 + d^2)^{-1/2} \mathbf{1}$ , where  $\mathbf{1}$  is a  $p \times 1$  unit-vector. For unequal group sizes  $\Lambda_{\mathbf{g}}$  can be computed by  $\Lambda_{\mathbf{g}} = d(p_1^{-1}p_2^{-1} + d^2)^{-1/2} \mathbf{1}$ , where  $p_1$  is the base rate (i.e., the number of individuals in a group, divided by the number of individuals in the total finite population) of group 1 and  $p_2$  is the base rate of group 2. For unequal group size, the group membership is coded into the effect-coded variable  $\boldsymbol{\theta}$ , so that  $\mathbf{g} = \boldsymbol{\theta}E(\boldsymbol{\theta}\boldsymbol{\theta}')^{-1/2}$  with  $E(\mathbf{g}\mathbf{g}') = \mathbf{1}$ . The random vector  $\boldsymbol{\varepsilon}$  represents  $p$  unique factors, with  $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \mathbf{I}_{p \times p}$ ,  $E(\boldsymbol{\varepsilon}\xi') = \mathbf{0}$ ,  $E(\boldsymbol{\varepsilon}\mathbf{g}') = \mathbf{0}$ , and  $\Psi$  is a  $p \times p$  diagonal, positive definite, unique factor loading matrix with  $p$  unique factor loadings. This implies that

$$E(\mathbf{x}\mathbf{x}') = \Sigma = \Lambda_{\xi} \Phi \Lambda_{\xi}' + \Lambda_{\mathbf{g}} \Lambda_{\mathbf{g}}' + \Psi \tag{2}$$

is a correlation matrix.

For  $q = 1$  and when  $d$  is the same for all observed variables, it is impossible to disentangle  $\xi$  and  $\mathbf{g}$ , so that only a single common factor  $\xi_{\mathbf{g}}$  combining  $\xi$  and  $\mathbf{g}$  can empirically be identified when the mean-difference between the groups occurs consistently across all observed variables. A model based on a single detectable common factor  $\xi_{\mathbf{g}}$  comprising equal group mean-differences on the observed variables and individual differences can be written as

$$\mathbf{x} = \Lambda_{\xi_{\mathbf{g}}} \xi_{\mathbf{g}} + \Psi \boldsymbol{\varepsilon}, \text{ with } \Lambda_{\xi_{\mathbf{g}}} = (\Lambda_{\xi}^2 + \Lambda_{\mathbf{g}}^2)^{1/2} = (\Lambda_{\xi}^2 + d^2(4 + d^2)^{-1} \mathbf{1})^{1/2}. \tag{3}$$

The determinacy coefficient is the correlation of a factor score predictor with the corresponding factor, i.e.,

$$\mathbf{P} = diag \left( E \left( diag \left( \mathbf{w}' \mathbf{x} \mathbf{x}' \mathbf{w} \right)^{-1/2} \mathbf{w}' \mathbf{x} \xi_{\mathbf{g}}' \right) \right) = diag \left( diag \left( \mathbf{w}' \Sigma \mathbf{w} \right)^{-1/2} \mathbf{w}' \Lambda_{\xi_{\mathbf{g}}} \right). \tag{4}$$

Inserting  $\Sigma^{-1} \Lambda_{\xi_{\mathbf{g}}} \Phi$ , the weights for the best-linear factor score predictor (Grice, 2001), for  $\mathbf{w}$  in Equation 4 yields  $\mathbf{P} = diag \left( \Phi \Lambda_{\xi_{\mathbf{g}}} \Sigma^{-1} \Lambda_{\xi_{\mathbf{g}}} \Phi \right)^{1/2}$  and

$$\mathbf{P} = \text{diag}\left(\Lambda_{\xi g}' \Sigma^{-1} \Lambda_{\xi g}\right)^{1/2} \tag{5}$$

for  $q = 1$ .

### 3. Results

#### 3.1 Effect of Factor Loadings and $p$ on the Determinacy Coefficient

For  $q = 1$  and all loadings  $\Lambda_{\xi g}$  being equal, it follows from  $\mathbf{w} = \Sigma^{-1} \Lambda_{\xi g} \Phi$  that all weights  $\mathbf{w}$  are equal so that one can write  $\mathbf{W} = w\mathbf{1}$ , where  $w$  is a scalar with  $w \neq 0$  and  $\mathbf{1}$  is a  $p \times 1$  unit-vector, so that

$$\mathbf{P} = \text{diag}\left(w^2 \mathbf{1}' \Sigma \mathbf{1}\right)^{-1/2} w \mathbf{1}' \Lambda_{\xi g} = \text{diag}\left(\mathbf{1}' \Sigma \mathbf{1}\right)^{-1/2} \mathbf{1}' \Lambda_{\xi g}. \tag{6}$$

When all weights are equal and greater zero and all loadings are equal and greater zero, and  $q = 1$ , it is possible to write  $\Lambda_{\xi g} = \lambda_{\xi g} \mathbf{1}$ , so that

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \lambda_{\xi g}^2 & \cdots & \lambda_{\xi g}^2 \\ \lambda_{\xi g}^2 & \sigma_x^2 & \ddots & \vdots \\ \vdots & \ddots & \sigma_x^2 & \lambda_{\xi g}^2 \\ \lambda_{\xi g}^2 & \cdots & \lambda_{\xi g}^2 & \sigma_x^2 \end{bmatrix} \tag{7}$$

and  $\Psi_{\xi g} = \psi_{\xi g} \mathbf{1}_{p \times p}$ . As for  $q = 1$  and equal factor loadings the weights cancel out, Equation 6 can be written as

$$\mathbf{P} = \left(\sigma_x^2 + (p-1)\lambda_{\xi g}^2\right)^{-1/2} \lambda_{\xi g} p^{1/2}, \tag{8}$$

and because  $\sigma_x^2 = 1$ , the common variance of the scores with the corresponding factor is

$$\mathbf{P}^2 = \left(\lambda_{\xi g}^{-2} + p - 1\right)^{-1} p. \tag{9}$$

#### 3.2 Benchmarks of Factor Loadings for Reasonable Determinacy

Factor score predictors should at least have 50% of common variance with the corresponding factor. Entering  $\lambda_{\xi g}^2 = (p+1)^{-1/2}$  into Equation 9, yields

$$\mathbf{P}^2 = \left(\lambda_{\xi g}^{-2} + p - 1\right)^{-1} p = (2p)^{-1} p = 0.50. \tag{10}$$

However, 80% or 90% of common variance of the factor score predictor with the corresponding factor will be a more reasonable benchmark. Entering  $\lambda_{\xi g}^2 = 4(p+4)^{-1}$  into the slightly transformed Equation 10 yields

$$\mathbf{P}^2 = \left(\lambda_{\xi g}^{-2} + p - 1\right)^{-1} p = \left(\frac{5}{4}p\right)^{-1} p = 0.80, \tag{11}$$

and entering  $\lambda_{\xi g}^2 = 9(p+9)^{-1}$  yields  $\mathbf{P}^2 = 0.90$ . The intended loading size for a given  $p$  and  $\mathbf{P}^2$  can be computed by

$$\lambda_{\xi g}^2 = \left(\frac{p}{\mathbf{P}^2} - p + 1\right)^{-1}. \tag{12}$$

For 90% of common variance of the factor score predictor with the factor and  $p = 4$  a loading size of .69 is required, which is attainable in several areas (see Figure 1).

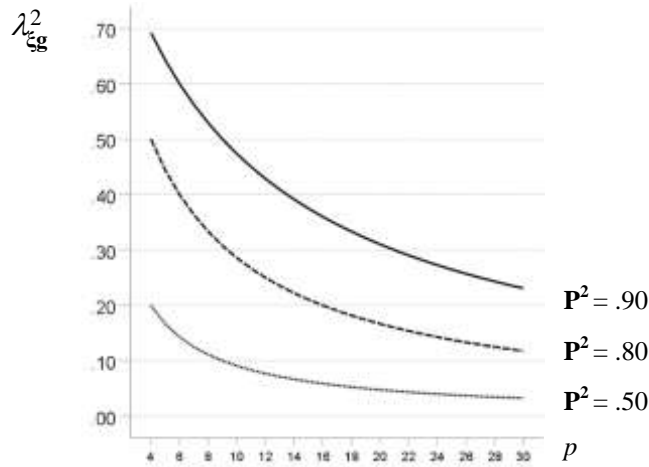


Figure 1. Loading size and  $p$  for 50%, 80%, and 90% common variance of the factor score predictor with the factor

### 3.3 Loading Benchmark, Group Differences and Determinacy

In the following, the effect of eliminating group differences in the factor score predictor on the determinacy coefficient is considered for  $q = 1$  and equal loading sizes. Eliminating group mean-differences implies that  $\Lambda_g = d(4 + d^2)^{-1/2} \mathbf{1} = \mathbf{0}$ . It follows from Equation 3 that eliminating group mean-differences reduces to  $\Lambda_\xi$ . It follows from  $\Lambda_{\xi g}$   $\Lambda_{\xi g}^2 = \Lambda_\xi^2 + d^2(4 + d^2)^{-1} \mathbf{1}$  and for all loadings being equal in  $\Lambda_{\xi g}$  that

$$\lambda_{\xi g}^2 = \lambda_\xi^2 + \frac{d^2}{4 + d^2}. \tag{13}$$

Entering  $\lambda_{\xi g}^2$  computed from Equation 13 and when the item variance is written as  $\sigma_x^2 = 1 = \lambda_\xi^2 + d^2(4 + d^2)^{-1} + \psi_{\xi g}^2$  Equation 9 yields

$$\mathbf{P}^2 = \left( \frac{\lambda_\xi^2 + \frac{d^2}{4 + d^2} + \psi_{\xi g}^2}{\lambda_\xi^2 + \frac{d^2}{4 + d^2}} + p - 1 \right)^{-1} p. \tag{14}$$

Entering  $d = 0$  into Equation 14 yields

$$\mathbf{P}^2 = \left( \frac{1 - \frac{d^2}{4 + d^2}}{\lambda_\xi^2} + p - 1 \right)^{-1} p. \tag{15}$$

This condition is a baseline for Figure 2 A-D, where the relationship of  $d$  and  $p$  with  $\mathbf{P}^2$  is presented for different levels of  $\lambda_\xi$ . Whereas the effect of eliminating a mean-difference of  $d = 1.00$  on  $\mathbf{P}^2$  is quite substantial for  $\lambda_\xi = .20$  and  $\lambda_\xi = .40$ , it is rather small for  $\lambda_\xi = .60$  and beyond.

However, according to Cohen et al. (2003)  $d = 0.80$  is a large effect, whereas  $d = .50$  represents a moderate effect, and  $d = .20$  represents a small effect. Eliminating a moderate effect of  $d = .50$  may also be considered. As an example, consider the condition  $p = 10$  and  $\lambda_\xi = .60$ . For this condition and  $d = 0.50$  Equation 14 yields

$$\mathbf{P}^2 = \left( \frac{1}{.60^2 + \frac{.50^2}{4 + .50^2}} + 10 - 1 \right)^{-1} 10 = .88. \tag{16}$$

For  $p = 10$ ,  $\lambda_\xi = .60$ , and  $d = 0.00$  Equation 14 yields

$$P^2 = \left( \frac{1 - \frac{.50^2}{4 + .50^2}}{.60^2} + 10 - 1 \right)^{-1} = .861. \tag{17}$$

Thus, in this condition, eliminating a moderate effect size in the factor score predictor will reduce the common variance of the factor score predictor with the original factor by only 2%.

It might be possible to compensate a decrease of  $P^2$  resulting from eliminating a group mean-difference by means of an increase of  $p$ . The increase of  $p$  necessary for compensating for the elimination of a group mean-difference can be calculated by means of the compensation-factor  $k$  as a multiplier of  $p$ . The compensation-factor  $k$  can be found by means of equating Equation 14 comprising the effects of  $d$  with Equation 14 for  $d = 0$  and  $k$  as a multiplier of  $p$ . This yields

$$\left( \frac{1}{\lambda_{\xi}^2 + \frac{d^2}{4 + d^2}} + p - 1 \right)^{-1} p = \left( \frac{1 - \frac{d^2}{4 + d^2}}{\lambda_{\xi}^2} + pk - 1 \right)^{-1} pk \tag{18}$$

and after some transformation

$$k = \left( \frac{1 - \frac{d^2}{4 + d^2}}{\lambda_{\xi}^2} - 1 \right) \left( \frac{1}{\lambda_{\xi}^2 + \frac{d^2}{4 + d^2}} - 1 \right)^{-1}. \tag{19}$$

The difference  $p^+ = kp - p$  might be rounded to the nearest integer, which indicates the number of observed variables that might be added in order to compensate for a decrease of  $P^2$  resulting from eliminating  $d$ . If  $kp - p \geq 1$  it might be reasonable to add observed variables in order to reach the same  $P^2$  that occurred before elimination of the group mean-difference.

As an example, consider a factor score predictor based on  $p = 5$ ,  $\lambda_{\xi} = .60$ , and  $d = .81$ , this results in  $P^2 = .83$  according to Equation 14. Now,  $d$  is eliminated by means of separate test-norms, or something similar, so that  $P^2 = .74$ . When  $p = 5$  and  $\lambda_{\xi} = .60$  are entered into Equation 19, the resulting  $k$  is

$$k = \left( \frac{1 - \frac{.81^2}{4 + .81^2}}{.60^2} - 1 \right) \left( \frac{1}{.60^2 + \frac{.81^2}{4 + .81^2}} - 1 \right)^{-1} = 1.38. \tag{20}$$

This implies that  $p^+ = 1.38 \cdot 5 - 5 = 1.9$ . When 2 items are added the resulting  $P^2$  will be about .83. So, a compensation for eliminating a mean-difference of about .80 is feasible. The compensation of the elimination of a group mean-difference  $d \geq .20$  by means of additional items  $p^+ = kp - p$  is plotted for different levels of  $P^2 \geq .70$  and different levels of  $\lambda_{\xi}$  in Figure 3. For  $\lambda_{\xi} = .40$  a compensation needs more than one item (Figure 3 A), but for  $\lambda_{\xi} = .60$  and  $\lambda_{\xi} = .80$  the effect of eliminating moderate or large mean-differences on  $P^2$  can be compensated by a few additional items (Figure 3, B and C).

When a group mean-difference is eliminated and when it is impossible to compute a multiple-group factor analysis for the respective groups, it might be of interest to compute the reduced factor loading. Equation 13 can be transformed in order to correct the initial factor loadings, when the eliminated group mean-difference  $d$  is known. This yields

$$\lambda_{\xi g}^2 - \frac{d^2}{4 + d^2} = \lambda_{\xi}^2. \tag{21}$$

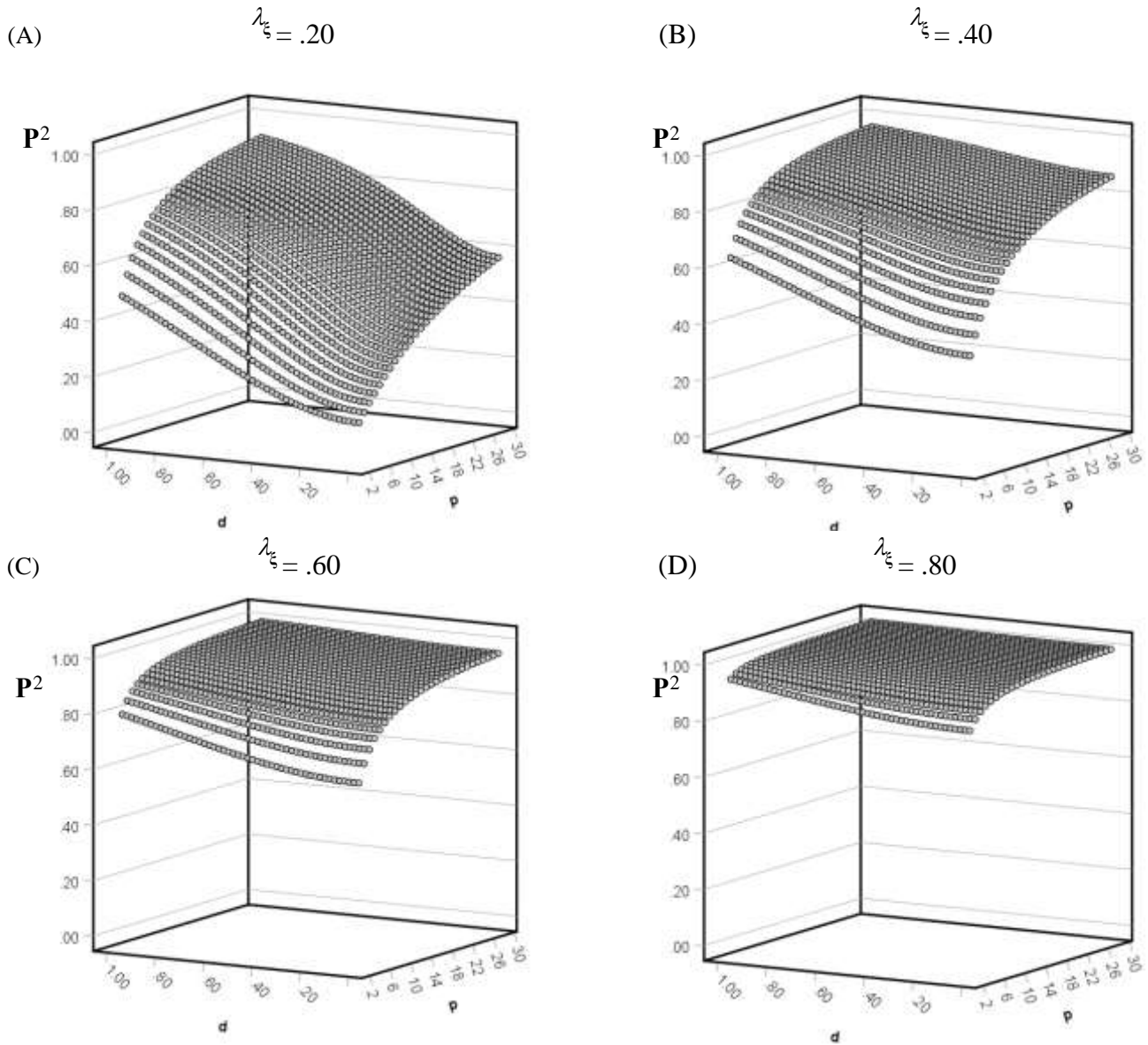
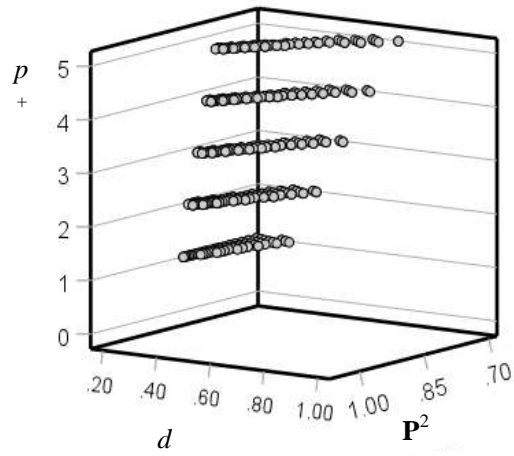
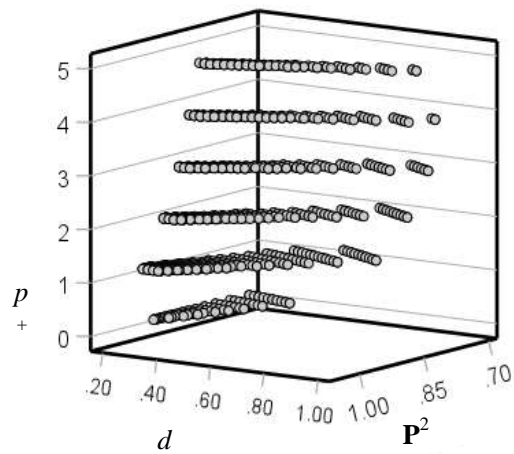


Figure 2. Effect of  $0 \leq d \leq 1$  and  $2 \leq p \leq 30$  on  $P^2$  for (A)  $\lambda_{\xi} = .20$ , (B)  $\lambda_{\xi} = .40$ , (C)  $\lambda_{\xi} = .60$ , and (D)  $\lambda_{\xi} = .80$

(A)  $\lambda_{\xi} = .40$



(B)  $\lambda_{\xi} = .60$



(C)  $\lambda_{\xi} = .80$

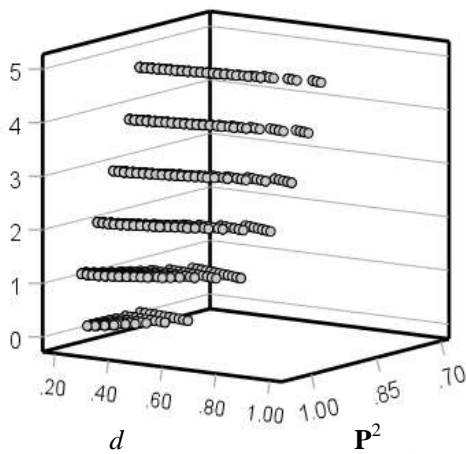


Figure 3. Number of additional items  $p_+$  that is necessary maintain a given  $P^2 \geq .70$ , when  $d \geq .20$  is eliminated, (A) for  $\lambda_{\xi} = .40$ , (B)  $\lambda_{\xi} = .60$ , and (C)  $\lambda_{\xi} = .80$

#### 4. Empirical Example

Data from 6892 German-speaking participants (2343 females, 4549 males; age in years:  $M = 25.93$ ,  $SD = 9.79$ ), who completed voluntarily the numerical tasks of the intelligence-structure-test 5 (IST 5; Beauducel, Brocke, & Liepmann, 2024) were used as an example. The tasks were numerical calculation, number series, and signs. Participants worked 10 minutes on each task. Each task comprised 20 items. The number of correct responses on each task (raw data) were entered into principal axis factor analysis (IBM SPSS Version 29.0.0.0). The resulting factor loadings are given in Table 1 (left column). The z-standardized best linear factor score predictor was computed and the corresponding determinacy coefficient was  $P^2 = .874$ . When a (Welch)  $t$ -test for gender differences was performed, the mean-difference was significant ( $t_{4843.35} = 12.98$ ,  $p < .001$ , two-tailed), with a larger mean for males ( $M = 0.11$ ,  $SE = 0.02$ ) than for females ( $M = -0.21$ ,  $SE = 0.02$ ), with  $d = 0.32$ . In the next step, the tasks were residualized for gender and entered into principal axis factor analysis (Table 1, right column), so that the mean difference between males and females was zero for the corresponding z-standardized best linear factor score predictor. The determinacy coefficient of the residualized best linear predictor was  $P^2 = .872$ .

Table 1. Factor loadings based on raw data and on data residualized for gender (right column)

Task	loadings (raw data)	loadings (residualized data)
numerical calculation	.83	.82
number series	.79	.79
signs	.87	.87
Eigenvalues	2.37, .35, .27	2.36, .36, .28

#### 5. Discussion

The determinacy coefficient describes the common variance of a factor score predictor with the respective factor and is therefore an indicator of the validity of factor score predictors. The present study investigates to what degree the validity of factor score predictor may depend on a given mean-difference between two groups. This question is relevant when mean-differences between groups are eliminated by means of covariance analysis, regression, or group specific norms.

It is shown that in a one-factor model comprising the same group mean-difference on all observed variables, the common factor cannot be distinguished from a common factor representing the group mean-difference. This occurs even when the common factor and the group mean-difference have a zero covariance. Moreover, the effect of eliminating a group mean-difference  $d$  on the common variance of the factor score predictor with the factor (determinacy coefficient) is investigated under the condition of equal common factor loadings. It is shown that for common factor loadings equal or larger than .60 the elimination of group mean-difference of a moderate effect size ( $d = .50$ ) leads to only small decreases of the determinacy coefficient. It is, nevertheless, important to estimate the reduced factor score determinacy when group mean-differences in the factor score predictors are eliminated. In this context, an estimation of the reduced factor loadings can also be of interest.

A compensation-factor  $k$  is proposed allowing for the estimation of the number of additional observed variables that are needed to recover the size of the determinacy coefficient before elimination of the group mean-difference. It turns out that for factor loadings equal or larger than .60 only a few additional items are needed in order to recover the initial determinacy coefficient after the elimination of moderate or large group mean-differences.

An empirical example based on numerical intelligence tasks demonstrates that a significant mean difference between males and females of about  $d = 0.30$  can be partialled out without any relevant effect on the determinacy coefficient. This implies that group-specific norms that might eliminate relevant group-differences in the factor score predictor do not necessarily reduce the validity of the scores. The empirical study shows that the algebraic results found for equal factor loadings can also be found for unequal factor loadings.

As a limitation of the study, it should be noted that no information on the educational level was available in the empirical data set. Therefore, the mean difference found in these data for numerical intelligence cannot be generalized to the population. Moreover, the investigation of the elimination of group differences on determinacy coefficients may also be performed for a larger number of factors. This would be of special interest because the elimination of group differences may affect the rotational position of the factors and the related pattern of factor inter-correlations. A further limitation of the study is that it is only based on the best-linear (regression) factor score predictor (Krijnen et al., 1996).



An investigation of the elimination of group mean-differences on the determinacy coefficients for other factor score predictors (Beauducel, & Hilger, 2022; Grice, 2001; Krijnen et al., 1996) may also be of interest. Other lines for further research are the investigation of the effect of unbalanced group-size, mean-differences between more than two groups, the effect of categorical observed variables, the effect of non-normal score distributions, and the effect of methods for the estimation of model parameters.

### Acknowledgments

This study was funded by the German Research Foundation (DFG), BE 2443/18-1.

### References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Asparouhov, T., & Muthén, B. (2023). Multiple group alignment for exploratory and structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 30, 169-191.  
<http://dx.doi.org/10.1080/10705511.2022.2127100>
- Beauducel, A., Harms, C., & Hilger (2016). Reliability estimates for three factor score estimators. *International Journal of Statistics and Probability*, 5, 94-107. <http://dx.doi.org/10.5539/ijsp.v5n6p94>
- Beauducel, A., & Hilger, N. (2022). Correlation-preserving mean plausible values as a basis for prediction in the context of Bayesian structural equation modeling. *International Journal of Statistics and Probability*, 11, 1-11. <http://dx.doi.org/10.5539/ijsp.v11n6p1>
- Beauducel, A. Brocke, B., & Liepmann, D. (2024). *Intelligenz-Struktur-Test 5*. Digitale Version. Göttingen: Hogrefe.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Statistical power analysis for the behavioral sciences*. (3rd ed.) Hillsdale, NJ: Erlbaum.
- DiStefano, C., M. Zhu, & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation* 14. <http://pareonline.net/getvn.asp?v=414&n=420>
- Gorsuch, R. L. (1983). *Factor analysis (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum.
- Grice, J. W. (2001). Computing and evaluation factor scores. *Psychological Methods*, 6, 430-450. <http://dx.doi.org/10.1037/1082-989X.6.4.430>
- Krijnen, W. P., Wansbeek, T. J., & Ten Berge, J. M. F. (1996). Best linear predictors for factor scores. *Communications in Statistics: Theory and Methods*, 25, 3013-3025. <http://dx.doi.org/10.1080/03610929608831883>
- Ruscio, J. (2008). A probability-based measure of effect size: robustness to base rates and other factors. *Psychological Methods*, 13, 19–30. <http://dx.doi.org/10.1037/1082-989X.13.1.19>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations of organizational research. *Organizational Research Methods*, 3, 4–70. <http://dx.doi.org/10.1177/109442810031002>.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).