Psychological Test and Assessment Modeling, 2026.68:32-52
DOI: https://doi.org/10.2440/001-0025

Considering Several Method Effects in Confirmatory Factor Analysis for Improving Model Fit in Investigating the Structure of Optimism Items

Karl Schweizer[1], Andreas Gold[1], Wolfgang Rauch[2], Tengfei Wang[3]

[1] Goethe University Frankfurt
[2] Pädagogische Hochschule Ludwigsburg
[3] Zhejiang University

Abstract:
The contribution of method-effect latent variables to accounting for the common systematic variation of optimism data in confirmatory factor analysis was investigated using an example correlation matrix. The considered method effects included the item-position effect, the wording effect and the item-level method effect. The analyses models were the congeneric, bifactor and original as well as basically tau-equivalent measurement models. In all cases, considering additional method-effect latent variables besides the attribute latent variable improved model fit, and in all cases, the virtually same sum of scaled factor variances emerged. But the bifactor model allowed for only two additional method-effect latent variable while the original and basically tau-equivalent models for four of them. In all multi-factor measurement models the optimism latent variable accounted for over two thirds of the common systematic variation. It turned out that the tau-equivalent models are better suited for revealing the method-effect structure of data than the other models.

Keywords: *measurement model, congeneric, tau-equivalent, confirmatory factor analysis, basically tau-equivalent*

Correspondence:

Karl Schweizer, Institute of Psychology, Goethe University Frankfurt;
k.schweizer@psych.uni-frankfurt.de

Introduction

This paper describes the investigation of the structure of a psychometric scale measuring optimism by confirmatory factor analysis (CFA) aiming at a complete as possible account of the common systematic variation of the data by considering several latent variables representing the attribute and well-documented method effects. Method effects refer to common systematic variation of data that stems from the measurement method or measurement context instead of the targeted attribute (Maul, 2013; Sechrest et al., 2000; Sechrest, Davis, Stickle, & McKnight, 2000). In a standard one-factor CFA measurement model, the latent variable has to account for the attribute-related common systematic variation that is assumed to be the only kind of common systematic variation, an assumption that may not be correct. Method-effect latent variables integrated into the CFA measurement model can avoid model misfit because of unaccounted additional common systematic variation of method effects. A complication of such a situation is that it is also necessary to assure that the account of common systematic variation by the attribute latent is sufficient, that is, it must be large enough with respect to what is accounted by the method latent variables.

This investigation addresses three method effects: the item-position effect, the wording effect and the item-level method effect. The item-position effect is a frequently observed method effect that can hardly be avoided since it arises inherently from the serial processing of several related items. This effect has been extensively studied using different research methods in different content areas (e.g., Campbell & Mohr, 1950; Debeer & Janssen, 2013; Hartig & Buchholz, 2012; Hohensinn, Kubinger, Reif, Holocher-Ertl, Khorramdel, & Frebort, 2008; Knowles, 1988; Kubinger, 2009; Lozano & Revuelta, 2020; Ren, Wang, Altmeyer, & Schweizer, 2014; Verguts & De Boeck, 2000. It typically manifests itself as a gradual increase in item means, item variances, and factor loadings from the first to the last items. While various explanations have been proposed, the most widely endorsed one is the learning explanation (e.g., Embredson, 1991; Verguts & De Boeck, 2000; von Gugelberg, Schweizer, & Troche, 2025).

Another method effect that may complicate attempts to control further effects like the item-position effect is the wording effect, which also has a long history of research (e.g., Messick, 1966). This effect is rooted in a special structure of the scale: the use of both positively and negatively worded items. It consists in construct-irrelevant variation of data due to the change of the item wording. This effect is mainly ascribed to acquiescence, the tendency to agree with the items regardless of content (Baumgartner & Steenkamp, 2001; DiStefano, Morgan, & Molt, 2012; Weijters, Baumgartner, & Schillewaert, 2013; Weijerts et al., 2013), but can also alternatively be explained by additional processing due to the negativity of items or additional variability due to individually different reactions to the change of wording (Arias, Ponce, & Martinez-Molina., 2022). Using CFA, the systematic variation of this effect is captured by a separate latent variable (DiStefano & Motl, 2006).

The item-level method effect has long been referred to as correlated error or correlated residual and ascribed to random influences (Landis, Edwards,& Cortina, 2009). Thorough analyses reveal that such an effect can arise when a pair of items shares a larger number of common cues than the remaining items of the scale among each other or through paraphrasing, leading to an overly large correlation (Bandalos, 2021). Such excessive item similarity reveals a construction failure that requires correction. In CFA, this effect can be controlled by introducing an additional latent variable into the measurement model to account for the extra common systematic variation (Schweizer, Gold, Krampen, & Troche, 2024).

Accounting for the common systematic variation of these method effects (Maul, 2013; Sechrest et al. 2000) requires the inclusion of additional latent variables into CFA measurement models. A similar approach is employed in multitrait-multimethod (MTMM) research, where CFA is extended to include not only additional trait latent variables, but also method-effect latent variables (Byrne, 2016). However, there is also a difference between CFA associated with the MTMM approach and CFA considering the item-position effect: this effect is not restricted to a few manifest variables, as is typical of method effects investigated by MTMM designs, but extends to all items requiring cross-loadings of all manifest variables on two latent variables (with one exception). Further, as it is hardly possible to create its contrary, it cannot be part of MTMM designs that typically display a symmetric structure.

The regular congeneric measurement model (Graham, 2016) can be extended to the bifactor model (Reise, 2012). In a bifactor model, each manifest variable loads on two latent variables. But the cross-loadings of the manifest variables (with one exception) require the estimation of a large number of parameters. This restricts the consideration of additional latent variables unless the set of manifest variables is fractionated into subset associated with extra latent variables. In contrast, the tau-equivalent model (Graham, 2016), also considered in method effects research, estimates only one parameter with respect to one latent variable that is the general factor loading on the latent variable since the other factor loadings are restricted to the same size. Extending this model by including additional latent variables is possible because of the small number of additional parameters. However, parsimony regarding parameter estimation is achieved at the expense of the capability of this model to reproduce complex empirical covariance matrices.

In the following, we describe the details of the models. Furthermore, the application to an example correlation matrix that can be expected to include not only variation due to the attribute but also due to method effects is described.

The CFA measurement models

The core of CFA is the measurement model since it includes the assumptions regarding the structure of data. We start with describing the *congeneric CFA measurement model* that is given by

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\lambda}\xi + \boldsymbol{\delta} \tag{1}$$

with $\mathbf{X}$ as the $p \times 1$ vector of manifest variables, $\boldsymbol{\mu}$ as the $p \times 1$ vector of intercepts, $\boldsymbol{\lambda}$ as the $p \times 1$ vector of factor loadings quantifying the relationships between latent and manifest variables, ξ as the latent variable representing the attribute of interest, in the case of a two-factor model the considered method effect as another component ($\boldsymbol{\lambda}\xi$), and $\boldsymbol{\delta}$ as the $p \times 1$ vector of residuals that are unique for each manifest variable.

The relationship between the CFA measurement model (Equation 1) and the input to CFA that is the $p \times p$ empirical covariance matrix, $\mathbf{S}$, is established by the corresponding $p \times p$ covariance matrix model, $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \boldsymbol{\lambda}\phi\boldsymbol{\lambda}' + \boldsymbol{\theta} \tag{2}$$

with the $p \times 1$ vector of factor loadings on the latent variable of the measurement model, $\boldsymbol{\lambda}$, the variance parameter, ϕ, and the $p \times p$ diagonal matrix of residual variances, $\boldsymbol{\Theta}$. Product $\boldsymbol{\lambda}\phi\boldsymbol{\lambda}$ accounts for the common systematic variation of the attribute.

An alternative specification of the second component of the right-hand side of Equation 1 leads to the *original tau-equivalent CFA measurement model*. The vector of factor loadings of Equation 1 is replaced by the product of the general factor loading, λ, and the $p \times 1$ vector of factor loading constants, $\boldsymbol{\beta}$, also referred to as loading pattern so that

$$\mathbf{X} = \boldsymbol{\mu} + (\lambda\boldsymbol{\beta})\xi + \boldsymbol{\delta}\ . \tag{3}$$

While the general factor loading requires estimation, the loading pattern consists of fixed numbers of the same size ($\boldsymbol{\beta} = \mathbf{1}$). Extending this model to comprise further latent variables requires further $\boldsymbol{\beta}$s including numbers that characterize the effects that are to be represented.

The modification of Equation 2 to fit with Equation 3 yields

$$\boldsymbol{\Sigma} = (\lambda\boldsymbol{\beta})\phi(\lambda\boldsymbol{\beta})' + \boldsymbol{\theta}\ . \tag{4}$$

Again, $\boldsymbol{\lambda}$ is replaced by $\lambda\boldsymbol{\beta}$. The product to the right of the equality sign accounts for the common systematic variation of the attribute.

The scalar λ of Equation 4 can be moved to the spot in between the vectors and associated with ϕ so that $\lambda\phi\lambda$. Further ϕ^* can replace $\lambda\phi\lambda$ for estimation so that

$$\boldsymbol{\beta}(\phi^*)\boldsymbol{\beta}' = (\lambda\boldsymbol{\beta})\phi(\lambda\boldsymbol{\beta})'. \tag{5}$$

Under the assumption of scaling according to the reference-group method (setting ϕ = 1 in Equations 4 and 5), ϕ^* can be treated as the square of λ.

Furthermore, there is the specification as *basically tau-equivalent CFA measurement model* (Schweizer & Wang, 2026) that differs from the original tau-equivalent model by the relaxation of tau-equivalence, that is, strict tau equivalence as correspondence of factor loadings is replaced by basic tau equivalence realized as zero slope of the regression line describing the factor loadings or loading pattern. It includes the replacement of the ingredients of $\boldsymbol{\beta}$ of Equations 3, 4 and 5 by the outcomes of transforming the factor loadings estimated according to the one-factor congeneric CFA measurement model (Equation 1). This can be perceived as rotation of the vector of factor loadings around the mid-point of this vector. Rotation brings the estimated factor loadings in line with basic tau-equivalence (zero slope) while maintaining item specificity.

The $p \times 1$ vector of estimates of congeneric factor loadings, $\boldsymbol{\lambda}$ (Equation 1), is transformed into the $p \times 1$ loading pattern, $\boldsymbol{\beta}^*$ (Equation 3). The transformation includes the computation of the $p \times 1$ vector of expected values of the regression line characterizing the observed factor loadings, $\boldsymbol{\nu}$, and the specification of the $p \times 1$ goal vector, $\boldsymbol{\varphi}$, reflecting tau equivalence achievable by assigning the mean of the observed factor loadings to all of its elements ($\varphi_i = \varphi_j$).

Given $\boldsymbol{\lambda}$, $\boldsymbol{\nu}$ and $\boldsymbol{\varphi}$, the elements of $\boldsymbol{\beta}^*$ are obtainable as

$$\beta_i^* = (\varphi_i \,/\, \nu_i)\, \lambda_i \tag{6}$$

(i = 1, …, p) where β_i^* is the ith transformed factor loading serving as constant included in the loading pattern, φ_i is the ith element taken from the goal vector, ν_i is the ith element of the vector representing the regression line (expected values), and λ_i is the ith observed congeneric factor loading (see second method for creating the basically tau-equivalent CFA measurement model, Equation 18, Schweizer & Wang, 2026). Advice for performing basically tau-equivalent CFA is provided in the Appendix A.

The implementation of method effects

The item-position effect. In the original tau-equivalent CFA measurement model and the basically tau-equivalent CFA measurement model, the item-position effect is realized as described in (Schweizer, Gold, & Krampen, 2022). The loading pattern on the additional position-effect latent variable include fixations that are defined as

$$\beta_{\text{Position-effect}}\,(i) = (i - 1) / (p - 1) \tag{7}$$

where $\beta_{\text{Position-effect}}\,(i)$ is the fixation with respect to manifest variable X ($i = 1, \ldots, p$) of $\boldsymbol{\beta}_{\text{Position-effect}}$ associated with the position-effect latent variable, $\xi_{\text{Position-effect}}$, and p is the number of manifest variables.

The wording effect. The representation of the wording effect requires the assignment of the manifest variables to one of two sets: $C_{\text{Positive-wording}}$ and $C_{\text{Negative-wording}}$. If $X_i \in C_{\text{Negative wording}}$ then β_i of $\boldsymbol{\beta}_{\text{Negative-wording}}$ loads on the wording-effect latent variable, $\xi_{\text{Wording-effect}}$, and is defined as

$$\beta_{\text{Wording-effect}}\,(i) = c \tag{8}$$

with $0 < c \leq 1$. Otherwise c is zero. We prefer $c = 1/n$ with n equal to the number of elements of $C_{\text{Negative-wording}}$. But there is no influence of the size of c on model fit if $0 < c \leq 1$.

The item-level method effect. This effect is defined in the same way as the wording effect except of the specification of the subsets of manifest variables. $C_{\text{Item-level effect}}$ includes the manifest variables with the overly large correlation while $C_{\text{Remainder}}$ comprises all the other manifest variables. If X_i is $\in C_{\text{Item-level effect}}$ then β_i of $\boldsymbol{\beta}_{\text{Item-level effect}}$ loading on the item-level latent variable, $\xi_{\text{Item-level effect}}$, is defined as

$$\beta_{\text{Item-level effect}}\,(i) = c \tag{9}$$

with $0 < c \leq 1$. Otherwise c is zero. We usually set $c = 0.5$.

Scaling for obtaining estimates of the variance parameter valid for serving as factor variances usually requires the modification the βs. We recommend the investigation of model fit before scaling for estimating factor variances.

Objective

The objective of this empirical study was threefold. First, it aimed to demonstrate that the control of method effects by including additional latent variables into the CFA measurement model improves model fit if there is corresponding common systematic variation. Second, it sought to reveal the gain in information on the relative amount of common systematic variation attributable to method effects as compared to the relative amount of common systematic variation attributable to the attribute. Third, the suitability of CFA measurement models for simultaneously investigating several method effects had to be explored. The considered models were the congeneric and bifactor models, as well as the one-factor and multi-factor versions of the original and basically tau-equivalent models.

Note. Further on we address CFA measurement models simply as models.

Materials and Methods

The correlation matrix of the study serving as example was created based on information recovered from a previously published study (Schweizer, Rauch, & Gold, 2011) focusing the assessment of personal optimism. The scale consisted of eight items with four graded response options. Four items were positively worded and the remaining items negatively. The sample included 308 university students (57 males and 245 females with the mean age of 23.76 years and the SD of 6.28).

A scale characteristic and a recent research result required the amendment of the formal descriptions of method effects presented in the previous section instead of their straight-forward realization. Recent research called into question the robustness of the item-position effect, suggesting that this effect could be disrupted by a change of the processing rule characterizing the items (von Gugelberg et al., 2025). Since 50 percent of the items of the optimism scale were worded negatively, such a disruption could be expected. Therefore, two correlated position-effect latent variables were inserted in the measurement models for representing the item-position effect instead of one.

Further an item-level method effect could be expected for the relationship of the first and second items of the scale. The first item taped the belief in the own (future) success and the second item the positivity of expectations regarding future outcomes. Although the wordings of the items differed, the contents were very similar. This degree of similarity could be the source for a high correlation.

Given these amendments, the empirical part of the study comprised the investigation of the structure of the optimism data by one-factor, three-factor and five-factor models. The one-factor models included the attribute latent variable with free factor loadings as congeneric model, with equal-sized numbers (0.125) (Equation 3) as factor loading constants of the original tau-equivalent model and with slope-corrected numbers (Equation 6) as factor loading constants of the basically tau-equivalent model. The multi-factor models had to be realized differently for the congeneric model and the other models.

Since the bifactor model as multi-factor version of the congeneric model only allowed for cross-loadings, it was not feasible to include all method-effect latent variables in one multi-factor model. Especially, it was not possible to include the second position-effect latent variable and the wording-effect latent variable into one model since they extended to the exactly same subset of manifest variables. Therefore, two three-factor models were created: the first included the attribute and the two item-position effect latent variables, and the second included the attribute, wording and item-level latent variables. They were realized by free parameters serving as factor loadings except in the case of the item-level method effect. Since this effect extended to two manifest variables only, it was realized by fixing the factor loadings to 0.5 and estimating the general factor loading indirectly (see Equation 5).

Using the other models, the multi-factor versions included all latent variables of interest (one attribute, two position-effect, one wording effect, and one item-level latent variables; see Figure 1). In these models, the general factor loadings were indirectly estimated by means of variance parameters (see Equation 5).

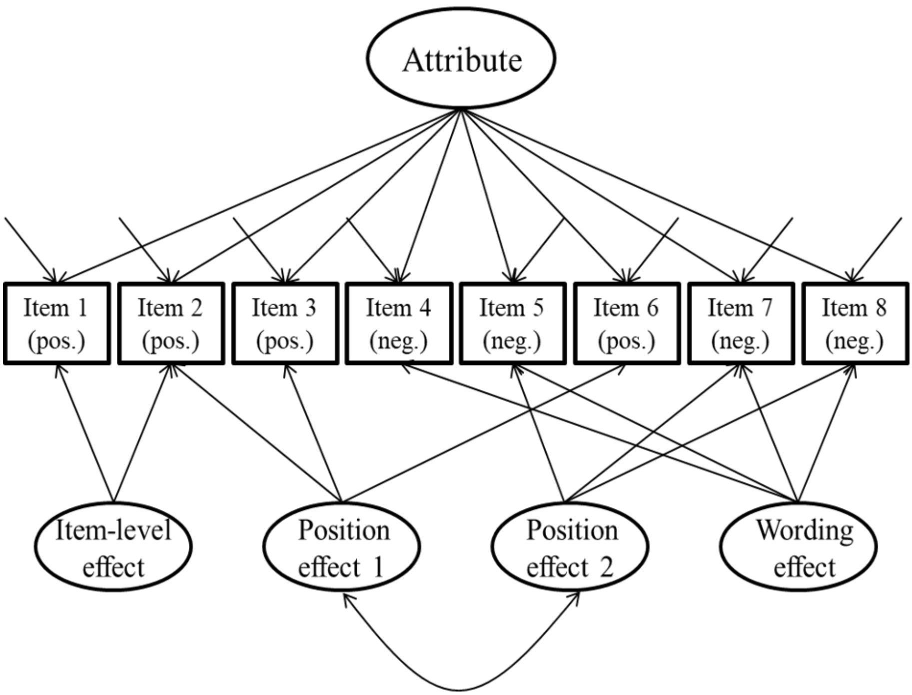

Figure 1
Five-factor CFA measurement model including the attribute latent variable (attribute), the item-level latent variable (item-level effect), the two position-effect latent variables (position effect 1 and position effect 2) and the wording latent variable (wording effect). Ellipses characterize latent variables and rectangles manifest variables. Arrows with long shafts depict assumed influences. Arrows with small shaft represent the residuals. The double-headed arrow represents the correlation of the position effects.

The statistical investigation included maximum likelihood estimation using LISREL software (Advice regarding LISREL notation is provided in Appendix B). Polychoric correlations estimated by PRELIS served as input. Considered fit indices and fit cutoffs (in parentheses) were χ^2, normed χ^2 as χ^2/df, RMSEA (≤ 0.06), SRMR (≤ 0.08), TLI (≥ 0.95), CFI (≥ 0.95), and AIC (DiStefano, 2016; Hu & Bentler, 1999). CFI differences (0.01) and RMSEA differences (0.015) served the comparison of models (Cheung & Rensvold, 2001).

Results

The investigation of the internal consistency of the available correlation matrix according to McDonald's (1999) Omega revealed a consistency estimate of 0.896 indicating a high degree of consistency. Overestimation due to a centroid-like factor solution could be ruled out as reason for the high consistency (Schweizer, Ren, & Wang, 2025).

Essential parts of the various measurement models are the attribute latent variables that differ from each other according to factor loadings. In the congeneric and bifactor models they are estimated individually. Figure 2A that is restricted to the upper and middle parts of Figure 1 illustrates this by using arrows with solid shafts linking manifest and latent variables.

In the basically tau-equivalent model (Figure 2B) fixed numbers originating from previous estimation are employed that is made apparent by dashed lines as arrow shafts linking manifest and latent variables while in the original tau-equivalent model (Figure 2C) dotted lines as arrow shafts signify the equivalence of factor loadings constants.

The estimated factor loadings and the constants of loading patterns of the attribute latent variable characterizing the various models are included in Table 1.

Table 1
Factor Loadings on the Attribute Latent Variable for the Congeneric, Bifactor, Original and Basic Tau-Equivalent CFA Measurement Models Using Polychoric Correlations as Input

Manifest variable	Congeneric factor loadings	Bifactor factor loadings 1	Bifactor factor loadings 2	Original tau-equivalent factor loadings	Basically tau-equivalent factor loadings
1	0.656	0.759	0.614	0.648	0.708
2	0.805	0.914	0.813	0.648	0.849
3	0.609	0.546	0.680	0.648	0.629
4	0.539	0.450	0.416	0.648	0.545
5	0.790	0.692	0.759	0.648	0.782
6	0.561	0.447	0.611	0.648	0.544
7	0.823	0.676	0.721	0.648	0.782
8	0.803	0.639	0.741	0.648	0.748
Mean	0.698	0.640	0.669	0.648	0.698
SD	0.120	0.158	0.124	0	0.114

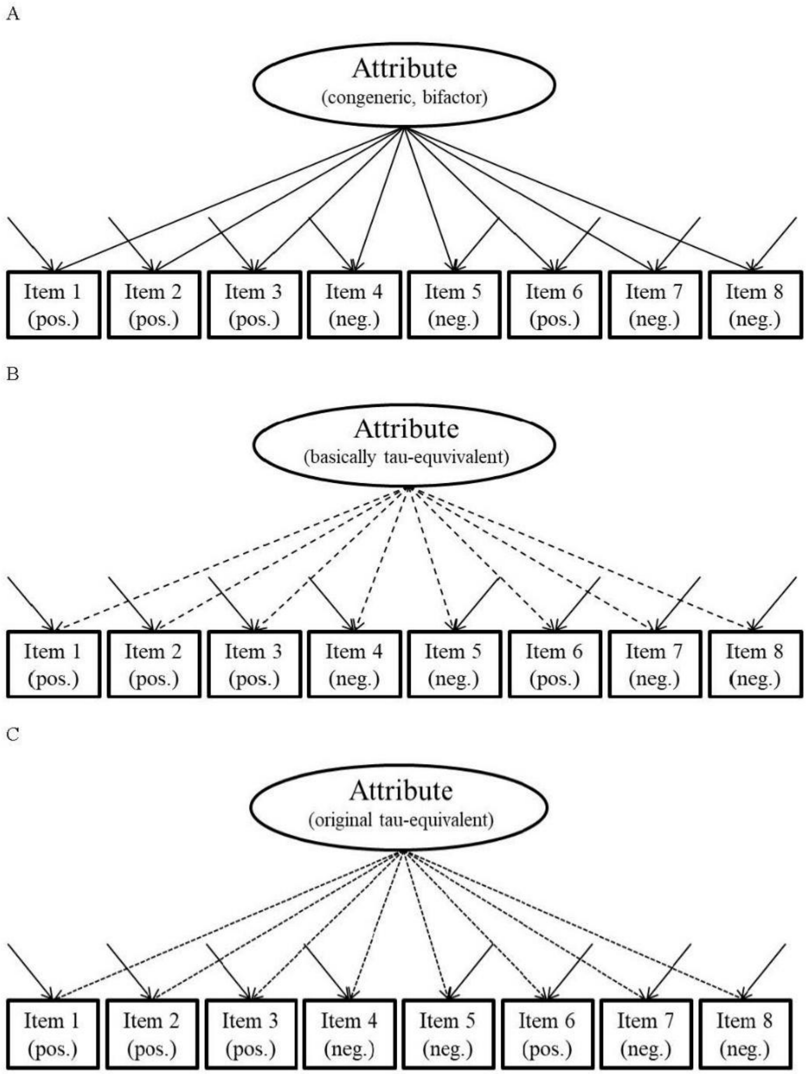

Figure 2
Illustration of measurement model parts regarding the attribute latent variable with solid lines representing estimated parameters, dashed line representing pre-estimated constants and dotted lines representing constants fixed to equal sizes.

While the first column of Table 1 identifies the manifest variables, the second column includes the factor loadings estimated by the one-factor congeneric model. There was variation between 0.539 and 0.823. The third column provides the factor loadings observed by the first bifactor model with additional position-effect latent variables. The estimates varied between 0.447 and 0.914. The factor loadings of the first and second manifest variables were especially large. In the fourth column, the factor loadings of the second bifactor model with additional wording and item-level latent variables are listed. The smallest factor loading was 0.416, and the largest one was 0.813. The loading pattern of the original tau-equivalent model is included in the fifth column. The numbers provided in this column were recovered from the estimate of the factor variance (Equation 5). They corresponded, as was required because of strict tau equivalence. The last column includes the loading pattern of the basically tau-equivalent model. The factor loadings varied between 0.545 and 0.849. The slope of the factor loading pattern was virtually zero (-0.008).

The results of investigating model fit

The fit results of investigating the structure of the example correlation matrix are reported in Table 2. The first to third rows include results for the congeneric and bifactor models. The one-factor congeneric model with the attribute latent variable was only good according to the SRMR result. Realized as model with position-effect latent variables, the CFI estimate additionally signified good model fit (second row). When realized as model with wording effect and item-level effect latent variables, SRMR, CFI and TLI indicated good model fit (third row). All RMSEA and CFI comparisons of one-factor and three-factor models were substantial.

The fourth and fifth rows of results were achieved by versions of the original tau-equivalent model. No one of the fit estimates for the one-factor model signified good model fit whereas the CFI estimate of the five-factor model surpassed the corresponding cutoff. The RMSEA and CFI comparisons of one-factor and multi-factor models were substantial.

The results of the sixth and seventh rows were obtained by the basically tau-equivalent model. As one-factor model, only the SRMR estimate reached the corresponding cutoff. In contrast, the multi-factor model signified good model fit, as indicated by SRMR, TLI and CFI estimates. Furthermore, in this model the RMSEA estimate could be considered marginally acceptable. All comparisons between one-factor and multi-factor models based on RMSEA and CFI showed substantial differences.

Note. Due to the fact that tau-equivalent models only required the estimation of the general factor loadings, larger degrees of freedom were reported for the original and basic tau-equivalent models than the congeneric and bifactor models. As a consequence, smaller AICs were observed that was especially obvious for the basically tau-equivalent model.

In sum, the worst model fit was observed for the one-factor models including the attribute latent variable only. Especially, the RMSEA estimates of all one-factor models were inacceptable (Brown, 2015, Hu & Bentler, 1999). The incorporation of additional latent variables to account for the common systematic variation of each method effect raised the fit results of the bifactor and basically tau-equivalent models to the required quality level in three out of four cases. But only the basically tau-equivalent model yielded a marginally acceptable RMSEA estimate (0.081) whereas the RMSEA of the second bifactor model was short of being inacceptable.

Table 2
Model Fit Results Observed in CFA According to the One-Factor (1F) and Multi-Factor (3F, 5F) Congeneric and Original as well as Basically Tau-equivalent Measurement Models Using Polychoric Correlations as Input

Measurement model	Latent variables	$\chi 2$	df	normed $\chi 2$	RMSEA	SRMR	TLI	CFI	AIC
Congeneric (1F)	AT	198.4	20	9.9	0.170	0.068	0.881	0.915	230.4
Bifactor 1 (3F)[1]	AT IPE1 IPE2	107.3	13	8.3	0.154	0.065	0.894	0.951	153.3
Bifactor 2 (3F)	AT WD ILE	60.8	15	4.1	0.100	0.039	0.958	0.977	102.8
Tau (1F)	AT	233.2	27	8.6	0.158	0.132	0.889	0.893	251.2
Tau (5F)	AT IPE1 IPE2 WD ILE	108.4	22	4.9	0.113	0.097	0.943	0.955	136.4
bTau (1F)	AT	207.1	27	7.6	0.147	0.075	0.913	0.916	225.1
bTau (5F)	AT IPE1 IPE2 WD ILE	66.7	22	3.0	0.081	0.041	0.971	0.977	94.7

Note. Tau abbreviates Tau-equivalent and bTau abbreviates Basically Tau-equivalent. Further AT abbreviates attribute, IPE1 first item-position effect, IPE2 second item-position effect, WD wording effect, and ILE item-level method effect.
[1] One factor loading on the first position-effect latent variable was *insignificant*.

The results of investigating the factor variances

Scaled factor variances were computed either according to the reference-group method (congeneric model and bifactor models) or the criterion-based method (original and basically tau-equivalent models) (Little, Slegers, & Card, 2006; Schweizer, Troche, & DiStefano, 2019). All estimates of the factor variances of the latent variables reached significance.

Table 3
Scaled Factor Variances and Percentages (in Italics) Observed Using the Bifactor and Original as well as Basically Tau-equivalent CFA Measurement Models

Latent variables in model	Attribute	IPE1	IPE2	WD	ILE	Sum
	Bifactor CFA measurement model					
AT IPE1 IPE2	3.45	0.61	0.73			4.79
	72.0	*12.7*	*15.2*			*100*
AT WD ILE	3.69			0.73	0.43	4.85
	76.0			*15.1*	*8.9*	*100*
	Tau-equivalent CFA measurement model					
AT IPE1 IPE2 WD ILE	3.36	0.42	0.33	0.47	0.24	4.82
	69.7	*8.7*	*6.8*	*9.8*	*5.0*	*100*
	Combination of CFA measurement models					
AT IPE1 IPE2 WD ILE	3.49	0.49	0.26	0.48	0.36	5.08
	68.7	*9.6*	*5.1*	*9.4*	*7.1*	*100*

Note. AT abbreviates attribute, IPE1 first item-position effect, IPE2 second item-position effect, WD wording effect, ILE item-level method effect.

The scaled factor variances of the latent variables of the multi-factor models are included in Table 3. The column titled *sum* provides the sums of the factor variances of all the latent variables included in each model. The sums observed for the bifactor models were 4.79 and 4.85, for the original tau-equivalent model it was 4.82, and for the basically tau-equivalent model it was 5.08, indicating a considerable degree of agreement.

The largest attribute factor variance was observed by a bifactor model (3.69). The largest factor variances of the method-effect latent variables were also observed for bifactor latent variables (item-position effect 1: 0.61, item-position effect 2: 0.73, wording effect: 0.73, item-level effect: 0.43).

As part of a bifactor model, the method-effect latent variables accounted for between 8.9 and 15.2 percent of the sums of factor variances. As part of the original tau-equivalent model, the additional latent variables accounted for between 5.0 and 9.8 percent, and as part of the basically tau-equivalent model for between 5.1 and 9.6 percent.

The estimate for the second item-position effect observed by the first bifactor model (0.73) was remarkably large as compared to what was observed for other models. The size was almost three times as large as what was observed for the corresponding latent variable of the tau-equivalent models. Also, the estimate for the wording effect of the second bifactor (0.73) was considerably larger than what was observed for the other models. Since the sums of factor variances of the bifactor models did not differ from the other sums, these results might be indications of a lack of discriminating between different method effects by the bifactor models.

In sum, all multi-factor CFA measurement models accounted for similar sums of factor variances. In all models, the attribute latent variables accounted for more than two thirds of the sum of factor variances. Each one of the considered method effects accounted for at least five percent of the sum of factor variances.

Discussion

This study reveals that some measurement characteristics linked to method effects impair model fit in confirmatory factor analysis unless there is appropriate control. The serial processing of items induces the item-position effect, and a change of the item wording leads to the wording effect. In both cases, common systematic variation is created that is unrelated to the attribute of interest (Maul, 2013; Sechrest et al. 2000) and needs to be captured by extra latent variables in order to prevent model misfit. The item-level method effect also consists in common systematic variation that can be modelled via a latent variable. This effect stems from an item-selection process that allows for a too large degree of similarity between a pair of items.

The scaled factor variances provide information about the impact of the method effects on measurement and enable the check of whether the attribute latent variable accounts for the superior amount of common systematic variation. Among these, the item-position effect emerged as the largest effect. In the bifactor model, as a whole it extends to 27.9 percent of the common systematic variation while according to the other models to between 14.7 and 15.5 percent. The wording effect comprises 15.1 percent of the common variation in the bifactor model, and between 9.4 and 9.8 according to other models. The percentages for the item-level method effect are between 5.0 and 8.9 percent.

The attribute latent variable accounts for between 72.0 and 76.0 percent of the common systematic variation when investigated by the bifactor model while between 68.7

and 69.7 by the other models. These percentages seem to suggest that the bifactor model is more efficient in accounting for attribute-related common systematic variation than the other models. However, the different percentages can simply be due to the different types of the latent variables since the free factor loadings of the bifactor model enable the better adaptability to the specificities of the data, as is apparent from the factor loadings of the first bifactor model (see Table 1). In this case, the factor loadings on the attribute latent variable signify an especially strong relationship of the first and second manifest variables with the attribute latent variable.

The sums of factor variances observed for the models are quite similar. They vary in the range between 4.79 and 5.08 although the models considerably differ according to the included sets of method effect latent variables. The bifactor models include only two method-effect latent variables that are either the two item-position effect latent variables or the combination of wording-effect and item-level effect latent variables while the other models include all four method-effect latent variables. One reason for this observation is that the bifactor models cannot discriminate between the second item-position effect and the wording effect because the corresponding latent variables are associated with the same set of manifest variables. Lack of discrimination in models including free factor loadings is also suggested by studies investigating the sensitivity of the congeneric model for influences of basic method effects regarding model fit (Schweizer, Ren, Wang, 2026).

Furthermore, it needs to be highlighted that reaching a sufficient degree of model fit turns out to be impossible without considering method effects. The estimates of model fit for each one-factor model do not reach the required fit standard. Even after including all method-effect latent variables, the original tau-equivalent model failed to meet this standard except of in one case. Given this predictable result (e.g., Graham, 2006), the basically tau-equivalent model was added to the set of considered models. It differs from the original tau-equivalent model by the relaxation of the equality requirement for factor loadings. This model can be realized by rotating the outcome of parameter estimation by the congeneric model as one of the two available methods so that the slope of the regression line describing the loading pattern becomes zero or close to zero. The attribute latent variable of the basically tau-equivalent model is designed to combine the representation of item specificity, as is characteristic of the congeneric model, with the equality requirement of the tau-equivalent model but in a relaxed version.

Another issue which we like to address is the gain for assessment when considering method effects. There is the separation of common systematic variation representing the attribute of interest from other common systematic variation so that a purified representation of the attribute becomes available leading to better results in investigations of external validity. Without taking method effects into consideration, other common systematic variation attenuates correlations between attribute latent variables on one hand and scale scores as well as other latent variables on the other hand so that smaller correlations are observed than otherwise. Such attenuation is avoided by considering method effects as part of CFA measurement models. Further representing method effects by latent variables implicitly means that to some degree method effect are differential effects. This implies that method common systematic variation contributes to the participants' test scores and is likely to lead to an increase in measurement error. Therefore, we suggest multi-factor measurement models as the outset for the development of scoring methods that decrease measurement error by eliminating influences of method effects on test results.

A limitation of the reported study is the lack of the generalizability of findings. The input to CFA is a correlation matrix obtained from data collected with optimism items. Therefore, it is an open question whether the performances of the models will be the same when the optimism items are replaced by other trait items. Also the specific arrangement of positively and negatively worded items can play a role. Further, the response format and the number of items can make a difference.

Finally, we like to highlight that the investigated method effects stem from the established way of collecting data in psychological assessment. Researchers should keep an eye on them to prevent that they get out of control. Specifically, identifying item-level method effects can provide clues for revisions. Information on the wording effect can help evaluating whether the change of wording means an advantage over uniformity of wording. Information on the item-position effect may be useful for controlling its impact on consistency and reliability estimates.

References

Arias, V. B., Ponce, F. P., & Martinez-Molina, A. (2022). How a few inconsistent respondents can confound the structure of personality survey data. *European Journal of Psychological Assessment, 39(6)*, 400–407. https://doi.org/10.1027/1015-5759/a000719

Bandalos, D. L. (2021). Item meaning and order as causes of correlated residuals in confirmatory factor analysis. *Structural Equation Modeling, 28(6)*, 903–913. https://doi.org/10.1080/10705511.2021.1916395

Baumgartner, H., & Steenkamp, J.-B. E. M. (2001). Response Styles in Marketing Research: A Cross-National Investigation. *Journal of Marketing Research*, *38(2)*, 143–156. http://www.jstor.org/stable/1558620

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

Campbell, D. T., & Mohr, P. J. (1950). The effect of ordinal position upon responses to items in a check list. *Journal of Applied Psychology, 34*, 62–67. https://doi.org/10.1037/h0061818

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9(2)*, 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Debeer, D., & Janssen, R. (2013). Modeling item-position effects within an IRT framework. *Journal of Educational Measurement*, *50(2)*, 164–185. https://doi.org/10.1111/jedm.12009

DiStefano, C., Morgan, G. B., & Motl, R. W. (2012). Detecting acquiescence underlying Rosenberg's self-steem scale using the Rasch partial credit model. In *Handbook on Psychology of Self-Esteem* (pp. 195-210). Nova Science Publishers, Inc.

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling, 13(3)*, 440–464. https://doi.org/10.1207/S15328007SEM1303_6

Embretson, S. E. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika*, *56(3)*, 495–515. https://doi.org/10.1007/BF02294487

Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement*, *66(6)*, 930–944. https://doi.org/10.1177/0013164406288165

Hartig, J., & Buchholz, J. (2012). A multilevel item response model for item position effects and individual persistence. *Psychological Test and Assessment Modelling, 54(4)*, 418–431.

Hohensinn, C., Kubinger, K.D., Reif, M., Holocher-Ertl, S., Khorramdel, L. & Frebort, M. (2008). Examining item-position effects in large-scale assessment using the Linear Logistic Test Model. *Psychology Science Quarterly, 50,* 391-402.

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)*, 1–55. https://doi.org/10.1080/10705519909540118

Knowles, E. S. (1988). Item context effects on personality scales: Measuring changes the measure. *Journal of Personality and Social Psychology*, *55(2)*, 312–320. https://doi.org/10.1037/0022-3514.55.2.312

Kubinger, K.D. (2009). Applications of the Linear Logistic Test Model in Psychometric Research. *Educational and Psychological Measurement, 69,* 232-244.

Landis, R. S., Edwards, B. D., Cortine, J. M. (2009). On the practice of allowing correlated residuals among indicators in structural equation models. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: verity and fable in the organizational and social sciences* (pp. 193-214). Routledge/Taylor & Francis Group.

Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling, 13(1)*, 59–72. https://doi.org/10.1207/s15328007sem1301_3

Lozano, J. H., & Revuelta, J. (2020). Investigating operation-specific learning effects in the Raven's Advanced Progressive Matrices: A linear logistic test modeling approach. *Intelligence*, *82.* https://doi.org/10.1016/j.intell.2020.101468

Maul, A. (2013). Method effects and the meaning of measurement. *Frontiers in Psychology, 4*:169. http://doi.org/10.3389/fpsyg.2013.00169

McDonald, R. P. (1999). *Test theory: a unified treatment*. Erlbaum.

Messick, S. (1966). The psychology of acquiescence: An interpretation of research evidence. *Research Bulletin (Educational Testing Service.)*, *1966(1)*, 1–44. https://doi.org/10.1002/j.2333-8504.1966.tb00357.x

Ren, X., Wang, T., Altmeyer, M., & Schweizer, K. (2014). A learning-based account of fluid intelligence from the perspective of the position effect. *Learning and Individual Differences, 31(1)*, 30–35. https://dx.doi.org/doi/10.1016/j.lindif.2014.01.002

Schweizer, K., Gold, A., & Krampen, D. (2022). A hierarchical model for data showing an item-position effect. *Psychological Test and Assessment Modeling, 64(3),* 253-271.

Schweizer, K., Gold, A., Krampen, D., & Troche, S. (2024). Conceptualizing correlated residuals as item-level method effects in confirmatory factor analysis. *Educational and Psychological Measurement, 84(5)*. http://doi.org/10.1177/00131644231218401

Schweizer, K., Rauch, W., & Gold, A. (2011). Bipolar items for the measurement of personal optimism instead of by unipolar items. *Psychological Test and Assessment Modeling, 53*, 399–413.

Schweizer, K., Ren, X., Wang, T. (2025). Overestimation of internal consistency by coefficient Omega in data giving rise to a centroid-like factor solution. *Educational and Psychological Measurement, 85*, 728–746. https://doi.org/10.1177/00131644241313447

Schweizer, K., Troche, S., & DiStefano, C. (2019). Scaling the variance of a latent variable while assuring constancy of the model. *Frontiers in Psychology (Section Quantitative Psychology and Measurement)*, 10, ArtID 887. https://www.doi.org/10.3389/fpsyg.2019.00887).

Schweizer, K., & Wang, T. (2026). Basically tau-equivalent confirmatory factor analysis considering item particularities. *Frontiers in Applied Mathematics and Statistics, 12*:1778863. https://doi.org/10.3389/fams.2026.1778863

Sechrest, L., Davis, M. F., Stickle, T. R., & McKnight, P. E. (2000). Understanding ‚method‘ variance. In L. Bickman (Ed.), *Research design: Donald Campbell's legacy* (pp. 63–87). Thousand Oaks, CA: Sage Publications.

Verguts, T., & De Boeck, P. (2000). A Rasch model for detecting learning while solving an intelligence test. *Applied Psychological Measurement*, *24(2)*, 151–162. https://doi.org/10.1177/01466210022031589

von Gugelberg, H. M., Schweizer, K., & Troche, S. J. (2025). Experimental evidence for rule learning as the underlying source of the item-position effect in reasoning ability measures. *Learning and Individual Differences, 118*. https://doi.org/10.1016/j.lindf.2024.102622

Weijters, B., Baumgartner, H., & Schillewaert, N. (2013). Reversed item bias: an integrative model. *Psychological Methods, 18(3)*, 320–334. https://doi.org/10.1037/a0032121

Appendix A.

Advice for performing basically tau-equivalent CFA

The steps necessary for conducting basically tau-equivalent CFA according to method 2 (of the two methods described in author), as employed in investigating the optimism data, are described.

Step 1: Perform one-factor CFA according to the congeneric measurement model for obtaining the estimated factor loadings (λ_t of Equation 6).

Step 2: Use the estimated factor loadings as criterion variable in ordinary linear regression analysis with linearly increasing numbers as predictor variable (we used 0.1, 0.2, 0.3, …) in order to obtain estimates of slope and intercept.

Step 3: Use slope and intercept for computing expected values (= ν_t of Equation 6).

Step 4: Compute the mean of the estimated factor loadings for achieving the elements of the goal vector (φ_t).

Step 5: Use Equation 6 for computing β_t^*.

Step 6: Repeat performing one-factor CFA with the fixed numbers of the loading pattern composed of the outcomes of step 5 (β_t^*) and with the variance parameter of the covariance matrix model (Equations 5) set free for estimation (ϕ*).

Step 7: Perform multi-factor CFA with additional components that are specified by fixed numbers and corresponding free variance parameters.

Appendix B:

Advice for one-factor CFA employing LISREL notation

See below the LISREL program code employed for investigating the matrix of polychoric correlations (sub-diagonal matrix including the main diagonal) according to the one-factor basically tau-equivalent measurement model (second method), congeneric measurement model, and original tau-equivalent model.

```
BASICALLY TAU-EQUIVALENT CFA MEASUREMENT MODEL
DA NI=8 NO=308 MA=KM

LA
p1 p2 p3 n1 n2 p4 n3 n4

KM

    1.000
    0.715   1.000
    0.377   0.549   1.000
    0.255   0.371   0.208   1.000
    0.482   0.638   0.500   0.499   1.000
    0.325   0.459   0.527   0.196   0.427   1.000
    0.485   0.618   0.432   0.595   0.627   0.459   1.000
    0.490   0.592   0.474   0.399   0.649   0.440   0.734   1.000

MO NX=8 NK=1 TD=DI,FR PH=FU,FI

LE
Optimism

VA      0.7078    LX      1      1
VA      0.8494    LX      2      1
VA      0.6287    LX      3      1
VA      0.5447    LX      4      1
VA      0.7818    LX      5      1
VA      0.5439    LX      6      1
VA      0.7821    LX      7      1
VA      0.7483    LX      8      1

FR PH 1 1

PD

OU ML MI RS SC ND=3 IT=1000 AD=OFF


CONGENERIC CFA MEASUREMENT MODEL
DA NI=8 NO=308 MA=KM

LA
p1 p2 p3 n1 n2 p4 n3 n4

KM

    1.000
    0.715   1.000
    0.377   0.549   1.000
    0.255   0.371   0.208   1.000
    0.482   0.638   0.500   0.499   1.000
```

```
     0.325    0.459    0.527    0.196    0.427    1.000
     0.485    0.618    0.432    0.595    0.627    0.459    1.000
     0.490    0.592    0.474    0.399    0.649    0.440    0.734    1.000

MO NX=8 NK=1 TD=DI,FR PH=FU,FI

LE
Optimism

FR      LX      1      1
FR      LX      2      1
FR      LX      3      1
FR      LX      4      1
FR      LX      5      1
FR      LX      6      1
FR      LX      7      1
FR      LX      8      1

VA 1 PH 1 1

PD

OU ML MI RS SC ND=3 IT=1000 AD=OFF
```

```
ORIGINAL TAU-EQUIVALENT CFA MEASUREMENT MODEL
DA NI=8 NO=308 MA=KM

LA
p1 p2 p3 n1 n2 p4 n3 n4

KM

     1.000
     0.715    1.000
     0.377    0.549    1.000
     0.255    0.371    0.208    1.000
     0.482    0.638    0.500    0.499    1.000
     0.325    0.459    0.527    0.196    0.427    1.000
     0.485    0.618    0.432    0.595    0.627    0.459    1.000
     0.490    0.592    0.474    0.399    0.649    0.440    0.734    1.000

MO NX=8 NK=1 TD=DI,FR PH=FU,FI

LE
Optimism

VA      0.125      LX      1      1
VA      0.125      LX      2      1
VA      0.125      LX      3      1
VA      0.125      LX      4      1
VA      0.125      LX      5      1
VA      0.125      LX      6      1
VA      0.125      LX      7      1
VA      0.125      LX      8      1

FR PH 1 1

PD

OU ML MI RS SC ND=3 IT=1000 AD=OFF
```