Reliability of Scales With Second-Order Structure: Evaluation of Coefficient Alpha's Population Slippage Using Latent Variable Modeling

Educational and Psychological
Measurement

2018, Vol. 78(6) 1123–1135

© The Author(s) 2017

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0013164417719948

journals.sagepub.com/home/epm

Tenko Raykov¹, Philippe Goldammer²,
George A. Marcoulides³, Tatyana Li¹, and
Natalja Menold⁴

Abstract

A readily applicable procedure is discussed that allows evaluation of the discrepancy between the popular coefficient alpha and the reliability coefficient of a scale with second-order factorial structure that is frequently of relevance in empirical educational and psychological research. The approach is developed within the framework of the widely used latent variable modeling methodology and permits point and interval estimation of the slippage of alpha from scale reliability in a population under investigation. The method is useful when examining the consistency of complex structure measuring instruments assessing higher order latent constructs and, under its assumptions, represents a generally recommendable alternative to coefficient alpha. The outlined procedure is illustrated using data from an authoritarianism study.

Keywords

coefficient alpha, factor analysis, latent structure, latent variable modeling, second-order factor model, reliability

¹Michigan State University, East Lansing, MI, USA

²Military Academy at ETH Zurich, Zurich, Switzerland

³University of California, Santa Barbara, CA, USA

⁴Leibniz Institute for the Social Sciences, Mannheim, Germany

Corresponding Author:

Tenko Raykov, Measurement and Quantitative Methods, Michigan State University, 443a Erickson Hall, East Lansing, MI 48824, USA.

Email: raykov@msu.edu

Reliability is a major psychometric quality index commonly used in the educational and behavioral sciences. Coefficient alpha (e.g., Cronbach, 1951) is frequently used to appraise the reliability of multicomponent measuring instruments that are highly popular in these and cognate disciplines. Meanwhile, a large body of research on the properties of alpha has accumulated that explicates the conditions under which this coefficient can be considered sufficiently informative about scale reliability in studied populations. In particular, Raykov (1997) showed that under restrictive yet testable conditions, coefficient alpha has minimal population slippage from scale reliability, which is negligible for most empirical purposes (see also Raykov & Marcoulides, 2015; Raykov, West, & Traynor, 2015). When these conditions do not hold, however, alpha can be seriously misleading (e.g., Sijtsma, 2009), as can also be the closely related index “alpha if item deleted” that is currently often used by researchers involved in scale construction and development.

Despite these limitations, the continued popularity of coefficient alpha and the last-mentioned index across various disciplines remains extensive. There seems to be also a tendency among some applied scientists to use alpha as a coefficient presumably informing about reliability with measuring instruments possessing complex structure that cannot be even approximately considered unidimensional. A particular type of such measuring instruments taps into several interrelated constructs that load on a second-order factor common to them. These instruments are likely to be of relevance in empirical research when validity considerations and especially construct underrepresentation concerns may lead scholars to include components related to more than one latent variable in an instrument being developed. For example, a mathematics ability test could consist of one part evaluating algebra ability, another assessing geometry ability, a third measuring trigonometry ability, and a fourth comprising problems assessing abstract thinking ability; thereby, these four abilities may load on a second-order factor representing the targeted mathematics ability (e.g., Raykov, Marcoulides, & Menold, 2018). In situations like this, the reliability of the overall scale score is oftentimes of particular interest. However, as may be implied from Novick and Lewis (1967), it would in general be incorrect to use coefficient alpha then despite its still high popularity as an index of overall scale consistency. The reason is that with uncorrelated measurement errors alpha is a lower bound of reliability, with the population discrepancy of the former from the latter possibly being of marked magnitude (e.g., Raykov, 1997; for the correlated error case, when that slippage of alpha can be considerable, see for instance Bollen, 1980, and Raykov & Marcoulides, 2011).

The present note aims to contribute to the study of the population discrepancy between coefficient alpha and scale reliability, focusing specifically on measuring instruments possessing second-order factorial structure. To this end, we discuss a latent variable modeling (LVM; e.g., B. O. Muthén, 2002) procedure, which allows point and interval estimation of the population slippage of alpha from the reliability coefficient pertaining to the overall sum score (or mean) of the components of such instruments. The goal of the article is also to raise awareness about what may be considered at times nearly indiscriminate use of coefficient alpha in some areas of

empirical educational and psychological research, particularly with tests, scales, or inventories possessing a second-order latent structure, where alpha can be seriously misleading as a lower bound of reliability. The discussed method is illustrated on data from a study involving measurement of the multidimensional concept of authoritarianism (e.g., Beierlein, Asbrock, Kauff, & Schmidt, 2014, and references therein).

Background, Notation, and Assumptions

To achieve the aims of this note, we assume that a set of (approximately) continuous measures are given (cf. Raykov et al., 2018). We will denote them as $y_1, y_2, \dots, y_p$ ($p > 1$) and presume that they represent the components of a measuring instrument whose sum score reliability is of interest to evaluate. Given the concerns of the article, we posit the following second-order factorial structure for these p components:

$$\underline{y} = \underline{\mu} + \Lambda \underline{\eta} + \underline{\varepsilon}, \text{ and} \quad (1)$$

$$\underline{\eta} = \Gamma \xi + \underline{\delta}. \quad (2)$$

In Equations (1) and (2), which represent the general model underlying this article, $\underline{y} = (y_1, y_2, \dots, y_p)'$ is the $p \times 1$ vector of test components with a positive definite covariance matrix (as assumed for any covariance matrix in this note), $\underline{\mu}$ is the $p \times 1$ vector of associated intercepts, $\underline{\eta}$ is the $q \times 1$ vector of first-order common factors assumed with zero mean and unit variance ($1 < q < p$), and ξ is a second-order factor with zero mean and unit variance (underlining denotes vector in this article). Furthermore, in this pair of model equations Λ is a $p \times q$ matrix of first-order factor loadings (with the components in $\underline{y}$ assumed without any cross-factor loadings; see the “Conclusion” section for extension), Γ is the $q \times 1$ second-order factor loading matrix, $\underline{\varepsilon}$ is the $p \times 1$ vector of unique factors with zero means and a diagonal covariance matrix that are uncorrelated with ξ and $\underline{\eta}$, and $\underline{\delta}$ is the $q \times 1$ vector of residual terms that are uncorrelated with ξ , $\underline{\eta}$, and $\underline{\varepsilon}$ (e.g., Harman, 1976). Moreover, we assume that the number of first-order constructs (i.e., the η s) and that of their corresponding indicators are such that the overall model in Equations (1) and (2) is identified (and that if necessary, suitable parameter constraints are introduced to achieve its identification). We similarly posit that whenever used, this model is plausible for a studied population. Last but not least, throughout the article we stipulate that the population under investigation consists of independent cases (persons; see the “Conclusion” section for a possible extension).

Evaluating the Population Slippage of Coefficient Alpha From Reliability of Measuring Instruments With Second-Order Factorial Structure

As discussed in detail in Raykov et al. (2018), Equations (1) and (2) imply that it would not be appropriate, strictly speaking, to view as unidimensional the instrument

comprising the components y , despite its feature of evaluating a single second-order factor ξ . It may thereby be useful, however, to consider for certain theoretically and empirically relevant aims—in addition to the sum scores associated with individual first-order factors—the overall sum score

$$X = y_1 + y_2 + \cdots + y_p, \quad (3)$$

or the weighted sum

$$W = w_1 y_1 + w_2 y_2 + \cdots + w_p y_p \quad (4)$$

using weights w_j that are known beforehand or estimated.

The well-known coefficient alpha (α) is defined as (e.g., Crocker & Algina, 2006):

$$\alpha = \frac{p}{p-1} \left[1 - \sum_{i=1}^p \text{Var}(y_i) / \text{Var}(X) \right], \quad (5)$$

where $\text{Var}(\cdot)$ denotes variance of the random variable within parentheses. As can be readily seen from Equation (5), alpha is “agnostic” of a second-order factor structure that could be associated with the components y_i ($i = 1, \dots, p$). In fact, even though each of the variances participating in the right-hand side could be expressed then in terms of the parameters of the second-order factor model (Equations 1 and 2), those expressions like the fact that underlying is a second-order structure are irrelevant for the value of alpha in a given sample or population of concern.

Measuring instruments possessing second-order factorial structure have attracted considerable interest in recent decades by methodologists and substantive researchers. In particular, as discussed in Raykov and Marcoulides (2012) the reliability of the overall sum score X is then representable as (see Equations 1 and 2, and recall the uncorrelated error term assumption for the first-order factor indicators)

$$\rho_X = \frac{\sum_{j=1}^q L_j^2 \gamma_j^2 + \sum_{j=1}^q L_j^2 \psi_j}{\sum_{j=1}^q L_j^2 \gamma_j^2 + \sum_{j=1}^q L_j^2 \psi_j + \sum_{j=1}^q \theta_j}, \quad (6)$$

where $\theta_j = \text{Var}(\varepsilon_j)$, L_j are the sums of the first-order factor loadings of the components pertaining to the j th factor η_j , $\psi_j = \text{Var}(\delta_j)$ are the associated latent disturbance variances, γ_j are the loadings of the j th first-order factor η_j on the second order factor ξ (recall also the earlier made assumption of all factor variances being 1; $j = 1, \dots, q$; see below and the appendix).

From Equations (5) and (6), it follows with some straightforward algebra that the population discrepancy Δ between coefficient alpha and the scale reliability coefficient for the setting of relevance in this article is

$$\begin{aligned} \Delta = \alpha - \rho_X &= \frac{p}{p-1} \left[1 - \sum_{i=1}^p \text{Var}(y_i) / \text{Var}(X) \right] - \frac{\sum_{j=1}^q L_j^2 \gamma_j^2 + \sum_{j=1}^q L_j^2 \psi_j}{\sum_{j=1}^q L_j^2 \gamma_j^2 + \sum_{j=1}^q L_j^2 \psi_j + \sum_{i=1}^p \theta_i} \\ &= \frac{p}{p-1} \left[1 - \sum_{i=1}^p (\lambda_i^2 + \theta_i) / \left(\sum_{j=1}^q L_j^2 \gamma_j^2 + \sum_{j=1}^q L_j^2 \psi_j + \sum_{i=1}^p \theta_i \right) \right] \\ &\quad - \frac{\sum_{j=1}^q L_j^2 \gamma_j^2 + \sum_{j=1}^q L_j^2 \psi_j}{\sum_{j=1}^q L_j^2 \gamma_j^2 + \sum_{j=1}^q L_j^2 \psi_j + \sum_{i=1}^p \theta_i}, \end{aligned} \quad (7)$$

where for notational simplicity λ_i denote the loadings of the components y_i on their pertinent factors (and the unit latent variance assumption is made use of; $i = 1, \dots, p$). We note that in the present setting the discrepancy Δ is in general positive. This follows from a main result in Novick and Lewis (1967), which states that with uncorrelated errors population alpha is a lower bound of scale reliability unless the scale components are essentially tau-equivalent (i.e., evaluating the same common true score with the same units of measurement) when it equals reliability (see also Raykov & Marcoulides, 2011).

From the right-hand side of Equation (7), we realize that the population slippage of alpha from scale reliability for instruments with second-order factorial structure is a nonlinear function of parameters of the model defined in Equations (1) and (2). Thus, when that model is plausible for a studied population, after fitting it to sample data from the latter using LVM a researcher can point and interval estimate this slippage. Specifically, substituting the estimates of the model parameters into the right-hand side of Equation (7) we obtain the estimate of the population slippage of alpha as

$$\begin{aligned} \hat{\Delta} &= \frac{p}{p-1} \left[1 - \sum_{i=1}^p (\hat{\lambda}_i^2 + \hat{\theta}_i) / \left(\sum_{j=1}^q \hat{L}_j^2 \hat{\gamma}_j^2 + \sum_{j=1}^q \hat{L}_j^2 \hat{\psi}_j + \sum_{i=1}^p \hat{\theta}_i \right) \right] \\ &\quad - \frac{\sum_{j=1}^q \hat{L}_j^2 \hat{\gamma}_j^2 + \sum_{j=1}^q \hat{L}_j^2 \hat{\psi}_j}{\sum_{j=1}^q \hat{L}_j^2 \hat{\gamma}_j^2 + \sum_{j=1}^q \hat{L}_j^2 \hat{\psi}_j + \sum_{i=1}^p \hat{\theta}_i}, \end{aligned} \quad (8)$$

where a hat denotes estimate of the parameter underneath (see Equation 7). In particular, when the popular maximum likelihood (ML) method of parameter estimation is used (with observed variable normality; e.g., Bollen, 1989), Equation (8) renders the ML estimator of the population slippage of coefficient alpha from the reliability

coefficient of the overall scale score X , due to the invariance property of ML estimators (e.g., Casella & Berger, 2002).

In an empirical study the point and interval estimates of the population discrepancy Δ between alpha and scale reliability, which are obtained with the discussed procedure, will be very informative when a researcher is interested in evaluating the extent to which use of coefficient alpha can be misleading and underestimating reliability of scales with second-order factorial structure (with uncorrelated error terms). In the next section, we illustrate this LVM-based method using empirical data from an authoritarianism study. We exemplify there the degree to which being concerned with coefficient alpha rather than scale reliability itself can misinform a researcher using measuring instruments with a higher order latent structure.

Illustration on Empirical Data

For the aims of this section, we use data from a study of $n = 238$ members of an online panel representing a sample of German adults (Internet users), which was concerned with examining right-wing authoritarianism. For further details on this study, we refer to Raykov et al. (2018), who used another portion of that data set for the purpose of illustrating a method for evaluating latent criterion validity. As discussed in the last source, authoritarianism as a second-order factor is measured in this empirical study by the so-called Short Scale of Authoritarianism (SSA) consisting of nine indicators (items) that cover three latent dimensions referred to as Aggression, Submission, and Conventionalism (see Table 1 in Raykov et al., 2018, for specifics regarding these items and dimensions).

To illustrate the LVM procedure for evaluation of the population discrepancy between coefficient alpha and scale reliability discussed earlier in this article, we fit to the used data set the model defined by Equations (1) and (2) with $p = 9$ and $q = 3$ (see the appendix for the needed Mplus source code). This model includes a total of four latent constructs—the above three first-order factors of Aggression, Submission, and Conventionalism with three indicators each, and their second-order factor Authoritarianism. Since the nine observed variables (SSA components) were evaluated each using a 5-point numeric fully verbalized rating scale with no extreme skewness, they are considered for the illustration purposes of this section as approximately continuous measures on which the robust ML method of model testing and parameter estimation is applied (cf. DiStefano, 2002; see also Raykov & Marcoulides, 2011). To deal with a notable proportion of missing data in the manifest measures and counteract possible violations of the missing at random (MAR) assumption underlying this method, we also include as an auxiliary variable the score from the so-called Left-Right Self-Placement (LRSP) scale (e.g., Enders, 2010; see also Raykov et al., 2018, for rationale behind this auxiliary variable choice).

The described model, when fitted to the data from these nine SSA measures, was found to be associated with the following tenable fit indices: chi-square (χ^2) = 28.773, degrees of freedom (df) = 24, p value (p) = .229, and root mean square

Table 1. Parameter Estimates, Standard Errors, *t* Values, and Two-Tailed *p* Values Associated With Fitted Model (Software Output Format).

Parameter		Estimate	SE	<i>t</i> Value	<i>p</i> Value
ETA1	BY				
Y1		1.009	0.068	14.737	.000
Y2		0.982	0.067	14.692	.000
Y3		0.895	0.069	12.916	.000
ETA2	BY				
Y4		0.695	0.068	10.169	.000
Y5		0.666	0.062	10.685	.000
Y6		0.560	0.058	9.641	.000
ETA3	BY				
Y7		0.619	0.068	9.169	.000
Y8		0.907	0.063	14.328	.000
Y9		0.678	0.056	12.027	.000
KSI	BY				
ETA1		0.631	0.079	7.982	.000
ETA2		0.830	0.095	8.752	.000
ETA3		0.556	0.079	7.072	.000
Intercepts					
Y1		2.761	0.077	35.655	.000
Y2		3.132	0.076	41.406	.000
Y3		2.680	0.076	35.112	.000
Y4		2.943	0.064	46.111	.000
Y5		2.016	0.059	34.145	.000
Y6		1.807	0.055	32.719	.000
Y7		3.332	0.068	48.815	.000
Y8		2.460	0.066	37.217	.000
Y9		2.085	0.058	36.005	.000
Variances					
KSI		1.000	0.000	999.000	999.000
Residual variances					
Y1		0.440	0.069	6.379	.000
Y2		0.425	0.066	6.460	.000
Y3		0.608	0.072	8.502	.000
Y4		0.504	0.072	6.979	.000
Y5		0.404	0.061	6.674	.000
Y6		0.423	0.052	8.120	.000
Y7		0.741	0.075	9.864	.000
Y8		0.235	0.070	3.367	.001
Y9		0.347	0.049	7.095	.000
ETA1		0.602	0.100	6.046	.000
ETA2		0.311	0.158	1.972	.049
ETA3		0.691	0.087	7.922	.000
New/additional parameters					
S1		2.886	0.155	18.627	.000
S2		1.921	0.121	15.888	.000

(continued)

Table 1. Continued

Parameter	Estimate	SE	t Value	p Value
S3	2.205	0.128	17.214	.000
ALPHA	0.599	0.011	56.339	.000
SR_2NDO	0.804	0.014	57.825	.000
DELTA	−0.204	0.004	−55.031	.000

Note. In addition to the self-explanatory variable reference (see Equations 1 and 2, as well as the Mplus source code in the appendix), SE = standard error; SR_2NDO = reliability coefficient of the overall scale score X (see Equation 3); $\#j = L_j$ from main text ($j = 1, 2, 3$).

error of approximation (RMSEA) = .043 with a 90% confidence interval (0, .073). The resulting parameter estimates in it are presented in Table 1.

In this plausible model, of particular interest are the estimates of coefficient alpha as well as scale reliability, and especially their difference. As seen from the last three rows of Table 1 the estimate of coefficient alpha, .599, is markedly lower than that of the authoritarianism scale reliability, .804. Their discrepancy is thereby estimated as $\hat{\Delta} = .599 - .804 = -.204$, which is of considerable magnitude. Moreover, the used software (see the appendix) also yields the 95% confidence interval for the population slippage of alpha as $(-.212, -.197)$. The latter interval represents with high confidence all plausible population values for this discrepancy between alpha and the scale reliability coefficient, and suggests substantial misestimation of reliability by alpha. This interpretation is further strengthened by examining the same-level confidence intervals of alpha and of scale reliability. Using the R-function “ci.rel” in Raykov and Marcoulides (2011, Chap. 7), the confidence interval for coefficient alpha results as $(.573, .624)$, and that for scale reliability as $(.771, .833)$. These two intervals do not overlap, with the one for reliability being positioned markedly above that for coefficient alpha, thus further contributing to the impression of considerable underestimation of reliability by alpha.¹ Last but not least, the reliability of the used authoritarianism scale seems to be of essentially “acceptable” level, being estimated above .80 with the method used in this article (and with a 95% confidence interval stretching from the high .70 through the low to mid .80s). We stress that this scale reliability related conclusion cannot be reached if one were to be using instead coefficient alpha, whose estimate possesses here a much lower magnitude.

Conclusion

This note was concerned with a LVM procedure for point and interval estimation of the population slippage of the popular coefficient alpha from the reliability coefficient of a scale with a second-order factorial structure. The article also aimed at raising awareness about use of coefficient alpha in lieu of scale reliability itself in empirical settings with this type of complex latent structure (and uncorrelated

measurement errors; see below). Specifically, under the assumptions made in the article (see Equations 1 and 2 and surrounding discussion), use of coefficient alpha cannot be generally recommended for the purpose of evaluating scale reliability. Rather, an application of Equation (6) is recommended then for point and interval estimation of reliability of measuring instruments with second-order factor structure (see next and Raykov & Marcoulides, 2012).

Several limitations of the discussed approach are worthwhile pointing out here (cf. Raykov et al., 2018). One, the procedure assumes (approximately) continuous individual scale components (first-order factor indicators). In case of indicator normality, as mentioned above use of ML estimation is appropriate and yields ML estimates of the alpha to reliability discrepancy, as well as of the overall scale's reliability and coefficient alpha if of interest (see the second term in the right-hand side of Equation 7, and Raykov & Marcoulides, 2012). With up to mild deviations from normality, which do not result from piling at scale end for an individual component(s), it may well be recommendable to use the robust ML method (MLR; L. K. Muthén & Muthén, 2017), possibly also with components having as few as 5 to 7 response options (e.g., DiStefano, 2002; see also the appendix). We encourage however further research on the robustness of the MLR method in such situations. With fairly large samples, weighted least squares (WLS) estimation is also available with nonnormal continuous instrument components (e.g., Bollen, 1989). Also, the outlined procedure is best used with large samples, owing to the fact that its application rests on ML, robust ML, or WLS estimation, with all of them grounded in asymptotic statistical theory (e.g., B. O. Muthén, 2002). Future research will hopefully contribute to the development of possible guidelines for determining sample sizes at which one could rely on that large-sample theory.

Moreover, we assumed that observations (studied persons) were independent, that is, not clustered or nested within (higher order) Level-2 units, such as schools, clinicians, interviewers, physicians, neighborhoods, cities, and so on. One may conjecture that the robust ML estimation method may also have some robustness to limited violations of this classical independence assumption, especially when the degree of non-normality is not pronounced. We are not aware, however, of sufficient research in this area that could help determine the extent and conditions under which one may trust such a potential recommendation. Alternatively, one may consider standard error and overall goodness-of-fit corrections in the presence of clustering effect, which would not affect the point estimates but will affect the interval estimates of reliability, alpha, and their discrepancy (e.g., L. K. Muthén & Muthén, 2017). Furthermore, the present article implicitly assumes relatively limited population heterogeneity which permits one to consider single-class (as opposed to mixture) modeling as used throughout the article. Similarly, the developments in this note assumed no cross-factor loadings for the first-order factor indicators, but its method is also readily applicable in the latter case after a minor modification to appropriately include cross-loadings in the second term of Equation (7).

We reiterate that the plausibility and identification of the model defined by Equations (1) and (2), which underlies this article, is essential when its procedure is used in applications (cf. Raykov et al., 2018). Where these conditions are not satisfied, the method cannot be generally recommended as it may yield misleading parameter estimates, standard errors, and statistical test results with respect to scale reliability, coefficient alpha, and their discrepancy. Lack of identification of the overall model may be expected with an insufficient number of indicators for one or more of the first-order factors (and/or a sufficiently small number of the latter), and may be resolved by adding appropriate parameter constraints that reflect substantively plausible parameter relationships in studied populations (e.g., Raykov & Marcoulides, 2006). Last but not least, as indicated on several occasions in the note, it rests on the assumption of uncorrelated error terms associated with the indicators of the first-order factors. Hence, the outlined procedure cannot be recommended when two or more such errors correlate, but can be readily extended to accommodate correlated errors by including their covariances in the denominator of the second term of the right-hand side of Equation (7) (see also Bollen, 1980).

In conclusion, this article offers to empirical educational, behavioral, and social scientists a readily applicable means for point and interval estimation of reliability of scales that possess second-order factorial structure as well as its difference from the popular coefficient alpha. The note also contributes to raising the awareness particularly among applied scientists of the fact that coefficient alpha can be seriously misleading if used for the purpose of reliability estimation for complex structure multiple-component measuring instruments, in particular with second-order factor structure.

Appendix

Mplus Source Code for Evaluating the Population Slippage of Coefficient Alpha From Reliability of Measuring Instruments With Second-Order Factorial Structure

```
TITLE:          LVM POINT AND INTERVAL ESTIMATION OF THE DIFFERENCE BETWEEN
                COEFFICIENT ALPHA AND RELIABILITY OF A SCALE WITH SECOND-ORDER
                FACTOR STRUCTURE (AND UNCORRELATED MEASUREMENT ERRORS
                ASSOCIATED WITH THE FIRST-ORDER FACTORS).

DATA:          FILE = <raw data file name>;

VARIABLE:      NAMES = Y1-Y9 Z1; ! 9 ITEMS OF AUTHORITARIANISM SCALE AND LRSP.
                MISSING = ALL(-999);
                USEV = Y1-Y9;
                AUXILIARY = (M) Z1;

ANALYSIS:      ESTIMATOR = MLR;

MODEL:         ETA1 BY Y1*(L1) ! L#J = LAMBDA#J IN EQUATION (7).
                Y2-Y3(L2-L3); ! 1ST FIRST-ORDER FACTOR.
                ETA2 BY Y4*(L4)
                Y5-Y6(L5-L6); ! 2ND FIRST-ORDER FACTOR.
```

ETA3 BY Y7*(L7)
 Y8-Y9(L8-L9); ! 3RD FIRST-ORDER FACTOR.
 Y1-Y9(TH1-TH9); ! ERROR VARIANCES (FOR 1ST-ORDER FACTOR INDICATORS).
 KSI BY ETA1*(G1) ! SECOND-ORDER FACTOR
 ETA2-ETA3(G2-G3);
 KSI@1; ! CONTRIBUTES TO MODEL IDENTIFICATION (SEE ALSO BELOW).
 ETA1-ETA3(PSI1-PSI3); ! LATENT DISTURBANCE VAR'S (1ST-ORDER FACTORS).

MODEL CONSTRAINT:

NEW(ALPHA, SR_2NDO, S1, S2, S3, DELTA); ! CREATING 'PLACE-HOLDERS' FOR PARAMETERS.

S1 = L1+L2+L3; ! S#J = L#J IN EQUATION (7).

S2 = L4+L5+L6;

S3 = L7+L8+L9;

ALPHA = 9/8*(1-(L1^2+TH1+L2^2+TH2+L3^2+TH3+L4^2+TH4+L5^2+TH5+L6^2+TH6
 + L7^2+TH7+L8^2+TH8+L9^2+TH9)/(S1^2*G1^2+S2^2*G2^2
 + S3^2*G3^2+S1^2*PSI1 + S2^2*PSI2+S3^2*PSI3
 +TH1+TH2+TH3+TH4+TH5+TH6+TH7+TH8+TH9)); ! SEE EQUATION (5).

SR_2NDO = (S1^2*G1^2+S2^2*G2^2+S3^2*G3^2+S1^2*PSI1
 + S2^2*PSI2+S3^2*PSI3)/ (S1^2*G1^2+S2^2*G2^2+S3^2*G3^2+S1^2*PSI1
 +S2^2*PSI2+S3^2*PSI3+TH1+TH2+TH3+TH4+TH5+TH6+TH7+TH8+TH9); !

SEE EQUATION (6).

DELTA = ALPHA-SR_2NDO; ! SEE EQUATION (7).

PSI1 = 1 - G1^2; ! CONTRIBUTES TO MODEL IDENTIFICATION (FIXING LAT. VAR. AT 1).

PSI2 = 1 - G2^2;

PSI3 = 1 - G3^2;

OUTPUT: CINTERVAL;

Note. See annotating comments and Table 1 for variable and auxiliary quantities notation. (For an introduction to the Mplus syntax, see e.g., Raykov & Marcoulides, 2006.)

Authors' Note

This research was in part conducted while T. Raykov was visiting the Leibniz Institute for the Social Sciences, Mannheim, Germany.

Acknowledgment

Thanks are due to B. Rammstedt for helpful support.

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.

Note

1. In the present authoritarianism empirical study the population value of the scale reliability coefficient of actual relevance is unknown, as is that of coefficient alpha. Hence, the alpha and reliability estimate and confidence interval interpretations offered in the present paragraph of the main text are based only on their sample estimates and their numerical comparison, and thus do not imply corresponding identical statements about their population values.

References

- Beierlein, C., Asbrock, F., Kauff, M., & Schmidt, P. (2014). Die Kurzsкала Autoritarismus (KSA-3): Ein ökonomisches Messinstrument zur Erfassung dreier Subdimensionen autoritärer Einstellungen [The Kurzsкала Authoritarianism (KSA-3): An economic measuring instrument for the acquisition of three sub-dimensions of authoritarian attitudes]. In D. Danner & A. Glöckner-Rist (Eds.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen* [Compilation of social science items and scales]. doi:10.6102/zis228
- Bollen, K. A. (1980). Issues in the comparative measurement of political democracy. *American Sociological Review*, 45, 370-390.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.
- Casella, G., & Berger, J. (2002). *Statistical inference*. Monterey, CA: Wadsworth.
- Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt College.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling*, 9, 327-346.
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.
- Harman, H. H. (1976). *Modern factor analysis*. Chicago, IL: University of Chicago Press.
- Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, 29, 81-117.
- Muthén, L. K., & Muthén, B. (2017). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.
- Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurement. *Psychometrika*, 32, 1-13.
- Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence for fixed congeneric components. *Multivariate Behavioral Research*, 32, 329-354.
- Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum.
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York, NY: Taylor & Francis.
- Raykov, T., & Marcoulides, G. A. (2012). Evaluation of validity and reliability of hierarchical scales. *Structural Equation Modeling*, 19, 495-508.
- Raykov, T., & Marcoulides, G. A. (2015). A direct latent variable modeling based procedure for evaluation of coefficient alpha. *Educational and Psychological Measurement*, 75, 146-156.

- Raykov, T., Menold, N., & Marcoulides, G. A. (2018). Studying latent criterion validity for complex structure measuring instruments using latent variable modeling. *Educational and Psychological Measurement*, 78, 905-917. doi:10.1177/0013164417698017
- Raykov, T., West, B. T., & Traynor, A. (2015). Evaluation of coefficient alpha for multiple component measuring instruments in complex sample designs. *Structural Equation Modeling*, 22, 429-438.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74, 107-120.