Multiple-Component Measurement Instruments in Heterogeneous Populations: Is There a Single Coefficient Alpha?

Educational and Psychological
Measurement

2019, Vol. 79(2) 399–412

© The Author(s) 2017

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0013164417733305

journals.sagepub.com/home/epm

Tenko Raykov¹, George A. Marcoulides²,
Michael Harrison³, and Natalja Menold⁴

Abstract

This note confronts the common use of a single coefficient alpha as an index informing about reliability of a multicomponent measurement instrument in a heterogeneous population. Two or more alpha coefficients could instead be meaningfully associated with a given instrument in finite mixture settings, and this may be increasingly more likely the case in empirical educational and psychological research. It is argued that in such situations explicit examination of class-invariance in the alpha coefficient must precede any statements about its possible value in the studied population. The approach permits also the evaluation of between-class alpha differences as well as point and interval estimation of the within-class alpha coefficients. The method can similarly be used in situations with (a) known class membership when distinct (sub)populations are investigated while their number is known beforehand and membership in them is observed for studied persons, as well as (b) in settings where only the number of latent classes is known for a population under investigation. The outlined procedure is illustrated with numerical data.

Keywords

coefficient alpha, latent class, mixture, measuring instrument, reliability, unobserved heterogeneity

¹Michigan State University, East Lansing, MI, USA

²University of California, Santa Barbara, CA, USA

³University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

⁴GESIS - Leibniz Institute for the Social Sciences, Mannheim, Germany

Corresponding Author:

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

Email: raykov@msu.edu

Reliability of multicomponent measurement instruments has attracted enormous interest over the past several decades by methodologists and substantive scholars in the educational and behavioral sciences. These instruments, such as tests, scales, self-reports, inventories, subscales, test batteries, testlets, questionnaires, or surveys (or parts thereof), are widely used in these and cognate disciplines owing to their highly useful property of providing multiple converging pieces of information about underlying constructs of theoretical and empirical interest (e.g., McDonald, 1999). A main index of measurement instrument reliability, which has been to date most frequently used in behavioral and social studies, is Cronbach's coefficient alpha (α ; Cronbach, 1951). This index has attracted a great deal of generally valid criticism since the 1960s (e.g., Sijtsma, 2009), yet at the same time there is a significant body of research that points also to its utility and applicability under specific conditions (e.g., Raykov, 1997, 2012; Raykov & Marcoulides, 2015; Raykov & Marcoulides, 2017, and references therein).

Essentially all research using coefficient alpha has traditionally been concerned with what may be referred to as standard analytical methods. A main feature of these methods is their tacit assumption of single class populations, such as those that are associated with limited unobserved heterogeneity. As a consequence, the available methods for evaluating coefficient alpha do not attend to the possibility that measurement instruments are commonly used with populations characterized by increasing diversity of their members and substantial unobserved heterogeneity due to being mixtures of multiple subpopulations or classes (latent classes, or unobserved groupings). These latent class mixtures present serious challenges to standard analytical approaches, in particular those for point and interval estimation of coefficient alpha for multi-item instruments under consideration. As elaborated for instance in Raykov and Marcoulides (2015) as well as in Raykov, Marcoulides, and Li (2016), when mixtures are not adequately accounted for, the application of standard or single-class methods can produce seriously biased parameter estimates, incorrect standard errors, and misleading hypothesis test results and substantive conclusions, in particular with respect to reliability of measurement instruments.

Attending to these issues, recently Raykov and Marcoulides (2015) outlined a latent variable modeling (LVM) procedure for point and interval estimation of the popular coefficient omega in mixture settings (e.g., McDonald, 1999), which is a widely used index of scale, test, or composite reliability (referred to henceforth as "instrument reliability"). The aim of the present note is to extend their generic approach to the case of evaluation of class-specific alpha coefficients as indices informing about reliability, including their between-class identity examination and evaluation of their differences, as well as point and interval estimation of the within-class alpha coefficients.

Background, Notation, and Assumptions

To attain the goals of this article, we assume that a set of (approximately) continuous measures are given, designated by $Y_1, \dots, Y_p$, which are the components of a

measuring instrument under consideration, such as a psychometric scale or test (referred to as “scale” or “instrument” below; $p > 1$). We will not assume, however, that these observed measures are homogeneous. That is, no stipulation will be made in the sequel that any subset of $Y_1, \dots, Y_p$, or these p measures collectively, evaluate a single underlying latent variable (such as an ability, trait, aptitude, attitude, or proficiency of interest). Furthermore, we assume that the instrument has been administered to a sample of independent subjects from a studied population (see the “Conclusion” section for a possible extension). Thereby, we presume that the population is associated with substantial unobserved heterogeneity, that is, represents a mixture of an unknown number k of latent classes (subpopulations; $k > 1$), with the individual person membership in these classes being unknown or unobserved. Similarly, the routine assumption in contemporary finite mixture settings is also made below that each person belongs to one and only one class.

As mentioned earlier, a main index aimed at informing about instrument reliability is the popular coefficient alpha that is defined in a studied population as follows:

$$\alpha = \frac{p}{p-1} \left[1 - \sum_{j=1}^p \text{Var}(Y_j) / \text{Var}(Y) \right], \quad (1)$$

where $Y = Y_1 + \dots + Y_p$ is the overall sum score (total scale score) of the manifest measures and $\text{Var}(\cdot)$ symbolizes variance. (Throughout this article, $\text{Var}(Y) > 0$ is assumed, i.e., the existence of observed individual differences on the scale score Y is postulated.) In the following discussion, we will be concerned primarily with class-specific alpha coefficients and their examination for identity, as well as with their point and interval estimation and that of their cross-class differences.

Coefficient Alpha for Measurement Instruments in Heterogeneous Populations

Class-Specific Alpha Coefficients

When a studied population consists of k latent classes, following Equation (1) one can define the class-specific coefficient alpha for a measurement instrument under consideration as

$$\alpha_{(c)} = \frac{p}{p-1} \left[1 - \sum_{j=1}^p \text{Var}_{(c)}(Y_j) / \text{Var}_{(c)}(Y) \right], \quad (2)$$

where the subindex “(c)” refers to the c th class ($c = 1, \dots, k$). That is, coefficient alpha in the c th class is a nonlinear function of the variances of the individual components, $Y_1, \dots, Y_p$, and of their overall sum score in that class.

We notice from Equation (2) that in general there are k different alpha coefficients, even for a single measurement instrument and a single population being

studied with it. Hence, when one is concerned with a (sufficiently) heterogeneous population, the concept of “coefficient alpha” may actually lose its direct meaning, just like the concepts of scale reliability and validity can (e.g., Raykov et al., 2016; Raykov & Marcoulides, 2015). We attend to this theoretically and empirically important matter in more detail next.

When Can One Speak of a Coefficient Alpha for a Measurement Instrument and Population?

From Equation (2) it is seen that one can speak of a (single) coefficient alpha associated with a given measurement instrument and studied population if only if

$$\alpha_{(1)} = \alpha_{(2)} = \dots = \alpha_{(k)} \quad (3)$$

holds. Otherwise, that is, with at least two of the k alpha coefficients in Equation (3) being unequal, there is in fact no single coefficient alpha that is relevant for the population at large. Hence, a simple reference to “coefficient alpha” for that instrument and population is then misleading.

An equivalent representation of Equation (3) that will be helpful below is obviously as follows:

$$\begin{aligned} \alpha_{(1)} &= \alpha_{(2)}, \\ \alpha_{(2)} &= \alpha_{(3)}, \\ &\dots \\ \alpha_{(k-1)} &= \alpha_{(k)}, \end{aligned} \quad (4)$$

that is (see Equation 1, after some straight-forward algebra),

$$\begin{aligned} \sum_{j=1}^p \text{Var}_{(1)}(Y_j) / \text{Var}_{(1)}(Y) &= \sum_{j=1}^p \text{Var}_{(2)}(Y_j) / \text{Var}_{(2)}(Y), \\ \sum_{j=1}^p \text{Var}_{(2)}(Y_j) / \text{Var}_{(2)}(Y) &= \sum_{j=1}^p \text{Var}_{(3)}(Y_j) / \text{Var}_{(3)}(Y), \\ &\dots \\ \sum_{j=1}^p \text{Var}_{(k-1)}(Y_j) / \text{Var}_{(k-1)}(Y) &= \sum_{j=1}^p \text{Var}_{(k)}(Y_j) / \text{Var}_{(k)}(Y). \end{aligned} \quad (5)$$

Hence, one can only then speak of a (single) coefficient alpha associated with a given measurement instrument and studied population, when the sequence (5) of $k - 1$ equalities holds. (We note in passing that for a population with only $k = 2$ classes, the sequence (4) reduces to the single equation $\alpha_{(1)} = \alpha_{(2)}$, which will be of relevance in the “Illustration on Data” section.)

Examining Class-Invariance in Alpha Coefficients

From the preceding discussion it is seen that class invariance in alpha coefficients occurs only when the equation sequence (5) holds. Hence, a test of class-invariance, or of the existence of a single (meaningfully referred to) coefficient alpha, amounts to a test of the $k - 1$ nonlinear constraints (5).

Given that we did not assume any particular latent structure underlying the observed measures $Y_1, \dots, Y_p$ under consideration (in particular, we did not assume the unidimensionality model to hold for them), to conduct this test use can be made of the approach discussed in Raykov and Marcoulides (2010). In that procedure, the following formal p -factor model is used that is based on p dummy zero-mean latent variables comprised in the $p \times 1$ vector $\underline{\eta}$, which are defined, respectively, as equal to the corresponding observed measures collected themselves in the $p \times 1$ vector $\underline{Y}$ (in the rest of this article, underlining denotes vector):

$$\underline{Y} = \underline{\mu} + \Lambda \underline{\eta} + \underline{\varepsilon}. \quad (6)$$

In Equation (6), $\underline{\mu}$ is the $p \times 1$ vector of observed means, Λ is the $p \times p$ identity matrix (e.g., Johnson & Wichern, 2002), and $\underline{\varepsilon}$ is the $p \times 1$ vector of error terms that are all set equal to 0. Hence, Equation (5) entails

$$\text{Cov}(\underline{Y}) = \text{Cov}(\underline{\eta}), \quad (7)$$

where $\text{Cov}(\cdot)$ denotes covariance matrix of the matrix within parentheses (e.g., Raykov & Marcoulides, 2008).

Therefore, from Equations (6) and (7), it follows that as a result of postulating model (6) all variances and covariances of the observed measures $Y_1, \dots, Y_p$ which quantities completely determine coefficient alpha (see Equation 1), are “re-parameterized” as the corresponding variances and covariances of their respective dummy latent variables in $\underline{\eta}$ (cf. Raykov & Marcoulides, 2010). Thus, based on the preceding discussion in this article, one could only then speak of a single coefficient alpha for a measurement instrument and population of concern, when for the pertinent parameters of model (6) the nonlinear constraints (5) are fulfilled; in that case, in all latent classes there is a single coefficient alpha, which is therefore meaningfully referred to as coefficient alpha for the studied instrument and population.

We next readily observe that using the LVM (B. O. Muthén, 2002) methodology we can fit model (6) in each latent class being considered. This can be achieved, for instance, by using the popular LVM software Mplus (L. K. Muthén & Muthén, 2017; further details are provided in Raykov & Marcoulides, 2010, and we refer for them to the latter source, as well as to the appendix of this article). More specifically, for the aims of this article, during the initial process of class enumeration, which is referred to below as Step 1 of the presently discussed LVM procedure, model (6) needs to be postulated in each of the multiple classes with all its parameters being unconstrained across classes. In other words, while examining the issue of number of latent classes in the population (based on the individual instrument components), any

used version of model (6) must postulate it as having all its free parameters unconstrained across the classes. The restrictions in Equation (5), which represent a series of $k - 1$ nonlinear parameter constraints, are subsequently readily introduced in the k -class model fitting process, referred to as Step 2, and lead to a correspondingly nested model ($k > 1$). Alternatively, in this Step 2, confidence intervals (CIs) at any prespecified confidence level can be constructed for each pair of classes, to assess the pertinent alpha coefficient differences across them (see below and the appendix). If this examination of alpha class differences is to be achieved by hypothesis testing in Step 2, within-class fitting of the model in Equation (6) and imposing constraints (5) is to be carried out, as is typically done in applications of the likelihood ratio theory and test in case of measure normality (e.g., Johnson & Wichern, 2002), or of the relevant corrected chi-square difference test with violations of the latter assumption (L. K. Muthén & Muthén, 2017). The alternative examination of class differences in coefficient alpha via CIs in this Step 2, as mentioned above, allows point and interval estimation of the cross-class alpha discrepancies as well as estimation of the within-class alpha coefficients (see the “Illustration on Data” section for an example).

The discussed set of LVM activities are illustrated in the following section demonstrating an application of the outlined procedure (see the appendix for the Mplus source code and an R-function for the interval estimation of alpha coefficients; cf. Raykov & Marcoulides, 2015).¹

Illustration on Data

For the purposes of this section, we use a simulated data set for $n = 1,000$ cases (grouped) in $k = 2$ classes of size 500 each, on a scale with $p = 4$ homogeneous components $Y_1, \dots, Y_4$, which is a setting that reflects a common situation in contemporary empirical educational and psychological research frequently concerned with unidimensional multicomponent measurement instruments (cf. Raykov & Marcoulides, 2015). In the first class, multinormal data were generated according to the following model:

$$\begin{aligned} Y_1 &= \xi + \delta_1, \\ Y_2 &= .75\xi + \delta_2, \\ Y_3 &= .75\xi + \delta_3, \\ Y_4 &= .75\xi + \delta_4, \end{aligned} \tag{8}$$

where ξ was a standard normal variate and the error terms $\delta_1, \dots, \delta_4$ were zero-mean independent normal variates with variances .5 each. The normal data in the second class were generated using the same model defined by the four equations in (8) but with η being normal with mean 2 and variance 2 and the error terms being independent standard normal variates each.

In Step 1 of using the coefficient alpha examination procedure outlined in this article, we successively fit model (6) for the above four observed measures, Y_1 through Y_4 , with $k = 1, 2$, and 3 classes; as indicated earlier, no cross-class constraints are

Table 1. Model Selection Indices for Fitted Mixture Models.

K	BIC	<i>p</i>
1	12558.942	n/a
2	12457.523	.000
3	12474.432	.101

Note. *k* = number of classes, *p* = *p* value for the Lo–Meng–Rubin–adjusted likelihood ratio test; n/a = not applicable; BIC = Bayesian information criterion. The four-class model was not associated with a replicated best likelihood even after 2,500 random starts with 250 second-stage iterations (L. K. Muthén & Muthén, 2017). We interpret this finding as consistent with the pertinent attempt for “over-extraction” of classes, which is not supported by the analyzed data that were generated based on a model with two classes (see Equations 8 and surrounding discussion). The BIC for the four-class model, which was associated with the “best” log-likelihood obtained thereby, was 12479.249, that is, more than 19 units higher/worse than the BIC of the preferred two-class model (see main text). Models with five or more classes were not subsequently pursued in this example (which activity would not necessarily be seen as warranted, given that the analyzed data were generated following a two-class model).

imposed thereby on any model parameter. (The pertinent Mplus command file for Step 1 and the subsequently selected two-class model, is the first source code presented in the appendix; see also the footnote to Table 1 regarding the four-class model.) The resulting Bayesian information criterion (BIC) indices and *p* values of the Lo–Meng–Rubin adjusted likelihood ratio test are presented in Table 1.

Table 1 suggests preferring the model with *k* = 2 classes over the other fitted models, based on its BIC and *p* value of the Lo–Meng–Rubin adjusted likelihood ratio test. The reason is that the BIC is the smallest for the model with *k* = 2 classes, and that its Lo–Meng–Rubin adjusted likelihood ratio test *p*-value is for the last time significant (as class number increases; see also footnote to Table 1; e.g., Wang & Wang, 2012). This is a correct decision, given that the data were generated using a two-class model (see Equations 8 and the discussion surrounding them).

In Step 2 of applying the procedure of this article, we use the selected two-class model to point and interval estimate the possible class difference in the pair of class-specific alpha coefficients for the instrument under consideration (cf. Raykov & Marcoulides, 2015, and see the Step 2 Mplus source code in the appendix). To this end, we introduce the alpha difference in Equation (5) (top equation) as an “external” parameter and request its interval estimation. This also yields standard errors for the two class-specific alpha coefficients. The resulting two within-class alpha estimates are as follows (with standard errors within parentheses):

$$\begin{aligned} \hat{\alpha}_{(1)} &= .829(.016), \\ \text{and} \\ \hat{\alpha}_{(2)} &= .807(.026), \end{aligned} \tag{9}$$

where $\alpha_{(1)}$ and $\alpha_{(2)}$ denote these two alpha coefficients and a hat symbolizes estimate of the parameter underneath.

Using these estimates, their standard errors, and the R-function “ci.alpha” in the appendix (which is identical to the R-function “ci.rel” in Raykov & Marcoulides, 2011, chap 7, and included in the appendix merely for the sake of completeness of this article), the 95% CI for coefficient alpha in Class 1, that is, for $\alpha_{(1)}$, results as (.795, .858). Similarly, the 95% CI for coefficient alpha in Class 2, $\alpha_{(2)}$, results as (.751, .853).² We note that these two CIs overlap to a marked degree, suggesting that both classes may be associated with the same population coefficient alpha for the four-component scale under consideration.

Independently of this observation, we examine next the point and interval estimates of the cross-class difference in alpha coefficients, defined as $\Delta\alpha = \alpha_{(1)} - \alpha_{(2)}$. This difference is also point and interval estimated in the same modeling session (see end of the Step 2 Mplus source code in the appendix), and is as follows:

$$\hat{\Delta\alpha} = .021(.026)(-.029, .072), \quad (10)$$

where the 95% CI provided by the software, based on the delta method (e.g., Raykov & Marcoulides, 2004), is stated within final parentheses.

To test now for no between-class differences in Cronbach's alpha coefficients, that is, the null hypothesis $H_0: \alpha_{(1)} = \alpha_{(2)}$ (in this example), we can make use of the so-called “duality principle” (Raykov & Marcoulides, 2008). Accordingly, testing a point hypothesis, like H_0 , against a bidirectional alternative (here $H_1: \alpha_{[1]} \neq \alpha_{[2]}$), is equivalent to checking whether the hypothetical value for their difference (i.e., 0 for $\Delta\alpha$ in the example) is contained in the corresponding CI for the respective parameter, $\Delta\alpha$. Because the 95% CI (10) of $\Delta\alpha$ obtained above covers 0, it is suggested that at the respective .05 significance level the null hypothesis H_0 of class identity in alpha coefficients, that is, the hypothesis stating $H_0: \Delta\alpha = 0$, can be retained. Thereby, according to the CI in (10), a range of plausible population values for the between-class difference $\Delta\alpha$ in alpha coefficients at the 95% confidence level used, stretches from $-.03$ through $.07$ (rounded off). Based on these analyses, we can conclude that one can speak in the present example of a single coefficient alpha (for the population from which the studied sample/data set could be formally considered as having been “drawn”; see also Note 2).

Since we know here the (true) parameters of the model having generated the analyzed data set, using the direct (true) variance and covariance implications from the corresponding Equations (8) and surrounding discussion (i.e., the so-called covariance algebra rules; Raykov & Marcoulides, 2006), we can now use Equation (1) to determine the true (population) alpha coefficients in each of the classes. Proceeding in this way, we obtain $\alpha_{(1)} = .835$ and $\alpha_{(2)} = .809$, respectively. These values are quite close to their estimates furnished with the LVM procedure of this note and reported in Equation (9), and are each positioned well within the respective 95% CIs rendered earlier with the same method. In addition, the true (population) difference in the two within-class alpha coefficients is $\Delta\alpha = .026$, which is similarly quite close to its

estimated counterpart with this method and reported in Equation (10). We also note that this (true) alpha difference $\Delta\alpha$ is well within its CI obtained above with the LVM approach of the present article.

Conclusion

This note was concerned with estimation of the popular Cronbach alpha coefficient in populations characterized by considerable unobserved heterogeneity. Populations that represent mixtures of unknown numbers of latent classes are of growing relevance in the behavioral, business, educational, medical, organizational, and social sciences, and present an increasing concern in them and cognate disciplines. The article outlined a widely applicable modeling approach allowing one (a) to examine for identity class-specific alpha coefficients, and (b) to point and interval estimate their difference as well as these within-class coefficients. Furthermore, the discussed method can also be directly used (c) to examine the alpha coefficient differences in prespecified distinct populations in case the subjects' class membership is known, and (d) to point and interval estimate these coefficients' difference (as well as to test various hypotheses about it; cf. Raykov & Marcoulides, 2008, 2015; see also Note 1). Moreover, the method is straightforwardly applicable in settings where only the number of latent classes is known but not subjects' membership in them. Last but not least, the procedure of this article is readily used in an empirical setting with the popular LVM program Mplus (L. K. Muthén & Muthén, 2017; see the appendix for needed source codes).

Following Raykov and Marcoulides (2015), who were concerned with the case of scale reliability in heterogeneous populations, we find it appropriate to caution researchers about blindly relying on reported values of alpha coefficients published in measurement instrument manuals. As elaborated in the last cited source, these manuals may well be based (a) on outdated population definitions (i.e., populations may have in fact changed qualitatively since the publication of the manuals), and no less importantly (b) on pilot studies that considered for analytic purposes their sampled populations as consisting of only a single class, which is a critical assumption that in our opinion is with time becoming in general increasingly less likely to be satisfied. Therefore, the alpha coefficients found in those instrument manuals are likely to be (a) potentially biased and (b) not applicable for one, several, or any of the latent classes (subpopulations) in a contemporary population of interest in the behavioral and social sciences and cognate disciplines, if not even misleading.

At this point it is also worthwhile indicating several limitations of the procedure discussed in this article. As described, it is currently applicable with (approximately) continuous scale components (cf. Raykov & Marcoulides, 2015). With up to mild deviations from normality, which do not result from piling at the end of any individual component(s), use of the robust maximum likelihood (ML) method is worth considering (MLR; L. K. Muthén & Muthén, 2017), perhaps even with components having as few as five to seven possible answer options and, in particular, relatively

symmetric distributions (cf. Raykov & Marcoulides, 2011). As mentioned earlier, we also assume that the studied subjects are independent of each other, that is, they are not clustered or nested within Level-2 or higher order units, such as schools, hospitals, cities, treatment centers, teams, companies, and so on. We similarly conjecture that the MLR estimation method may as well have some robustness to violations of this classical independence assumption if their extent is limited, like the degree of nonnormality of the scale components. Further research is clearly needed before one may place trust in such a potential recommendation. In addition, the procedure is best used with large samples, since it rests on ML or MLR estimation, which are the key methods in mixture analysis at present (e.g., B. O. Muthén, 2002; L. K. Muthén & Muthén, 2017). Future research is hopefully going to help develop possible guidelines for determining sample size when one could rely on the underlying large-sample theory as having obtained practical relevance in empirical studies.

In conclusion, this article offers to researchers a widely applicable means for evaluating the popular Cronbach's alpha coefficient in populations with unobserved heterogeneity that are mixtures of latent classes. Together with the scale reliability and criterion validity procedures in Raykov and Marcoulides (2015) and in Raykov et al. (2016), the approach outlined here allows one to make more informed conclusions about the psychometric quality of studied multicomponent measurement instruments in finite mixture settings.

Appendix

Mplus Source Codes for Studying Alpha Coefficients in Finite Mixture Settings, and R-Function for Interval Estimation of Alpha Coefficients

```
TITLE:      TWO-CLASS MODEL FOR STUDYING ALPHA COEFFICIENTS IN FINITE
            MIXTURE SETTINGS. (ILLUSTRATION SECTION EXAMPLE.)
            STEP 1 OF OUTLINED PROCEDURE (CLASS ENUMERATION). SOURCE
            CODE 1.
            (SEE ALSO NOTE AT END OF CODE.)
DATA:      FILE = <NAME OF RAW DATA FILE>;
VARIABLE:  NAMES = Y1-Y4;
            CLASSES = C(2);
ANALYSIS:  TYPE = MIXTURE;
            STARTS = 1500 150;
MODEL:     %OVERALL% ! SEE EQUATION (6).
            ETA1 BY Y1@1;
            ETA2 BY Y2@1;
            ETA3 BY Y3@1;
            ETA4 BY Y4@1;
            Y1-Y4@0;
            [ETA1-ETA4@0];
            [Y1-Y4@0];

            %C#2% ! NO CONSTRAINTS ACROSS CLASSES ARE NEEDED.
```

```

ETA1 BY Y1@1;
ETA2 BY Y2@1;
ETA3 BY Y3@1;
ETA4 BY Y4@1;
Y1-Y4@0;
[Y1-Y4];
ETA1-ETA4;

```

```
OUTPUT:      TECH11;
```

Note. When fitting the model with fewer/more classes, delete/add the pertinent section for the dropped/added classes (i.e., the section starting with “%C#%” and finishing with “ETA1-ETA4;” 7 lines later). The two-class model, with the above source code, is preferred in the “Illustration on Data” section example.

```

TITLE:      TWO-CLASS MODEL FOR STUDYING ALPHA COEFFICIENTS IN FINITE
             MIXTURE SETTINGS. (ILLUSTRATION SECTION EXAMPLE DATA.)
             STEP 2 OF OUTLINED PROCEDURE - POINT AND INTERVAL
             ESTIMATION OF WITHIN-CLASS ALPHA COEFFICIENTS AND THEIR
             CROSS-CLASS DIFFERENCE.
             SOURCE CODE 2. (SEE NOTE AT END OF CODE.)

DATA:       FILE = <NAME OF RAW DATA FILE>;

VARIABLE:   NAMES = Y1-Y4;
             CLASSES = C(2);

ANALYSIS:   TYPE = MIXTURE;
             STARTS = 1500 150;

MODEL:      %OVERALL%
             KSI1 BY Y1@1;
             KSI2 BY Y2@1;
             KSI3 BY Y3@1;
             KSI4 BY Y4@1;
             Y1-Y4@0;
             [KSI1-KSI4@0];
             [Y1-Y4@0];
             KSI1-KSI4 (S11-S14);
             KSI1 WITH KSI2-KSI4 (S112-S114);
             KSI2 WITH KSI3-KSI4 (S123-S124);
             KSI3 WITH KSI4 (S134);

             %C#2%
             KSI1 BY Y1@1;
             KSI2 BY Y2@1;
             KSI3 BY Y3@1;
             KSI4 BY Y4@1;
             Y1-Y4@0;
             [Y1-Y4];
             KSI1-KSI4;

```

```
KSI1-KSI4 (S21-S24) ;
KSI1 WITH KSI2-KSI4 (S212-S214) ;
KSI2 WITH KSI3-KSI4 (S223-S224) ;
KSI3 WITH KSI4 (S234) ;
```

MODEL CONSTRAINT:

```
NEW (ALPHA1, ALPHA2, DELTA) ;
ALPHA1 = 1.333*(1-(S11+S12+S13+S14) /
  (S11+S12+S13+S14+2*(S112+S113+S114
    +S123+S124+S134)) ) ;
ALPHA2 = 1.333*(1-(S21+S22+S23+S24) /
  (S21+S22+S23+S24+2*(S212+S213+S214
    +S223+S224+S234)) ) ;
! CONFIDENCE INTERVAL-BASED EXAMINATION OF ALPHA DIFFERENCES.
DELTA = ALPHA1-ALPHA2;
```

OUTPUT: CINTERVAL;

Note. In general, in lieu of the multiplier 1.333 for ALPHA1 and ALPHA2 use the pertinent ratio $p/(p - 1)$ ($p > 1$, as assumed throughout this article).

```
ci.alpha = function(r, se){ # R-function for interval estimation of
coefficient alpha
  l = log(r/(1-r))
  sel = se/(r*(1-r))
  ci_l_lo = l-1.96*sel
  ci_l_up = l+1.96*sel
  ci_lo = 1/(1+exp(-ci_l_lo))
  ci_up = 1/(1+exp(-ci_l_up))
  ci = c(ci_lo, ci_up)
  ci
}
```

Note 1. This R-function is identical to the R-function “ci.rel” in Raykov and Marcoulides (2011, chap 7), and is presented here merely for the sake of completeness of the present article. (This R-function needs to be correspondingly modified if a confidence interval at another level is to be obtained; for instance, for a 90% confidence interval, use 1.645 in lieu of 1.96 in the two middle commands of the function). Use this R-function only if the estimated coefficient alpha is positive.

Note 2. To use this R-function, at the R prompt type/submit “ci.alpha(v, se),” where for “v” the class-specific alpha estimate is entered and for “se” its associated standard error (see pertinent section of software output for Step 2 of the outlined procedure in this article). This R-function is also applicable for interval estimation of a common alpha coefficient, when the null hypothesis H_0 of no cross-class alpha differences in Equation (5) is retained/not rejected (see Note 1 for details pertaining to point and interval estimation of the class-specific alpha coefficients).

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.

Notes

1. The current paragraph in the main text in effect outlines an approach to point and interval estimation of the common alpha coefficient for a given multicomponent instrument under consideration, in case the hypothesis (3) stating lack of between-class differences in alpha coefficients is retained. This approach consists of fitting then the formal model (6) to the data from all classes, whereby the set of cross-class constraints (5) is implemented in the model that ensures class invariance in alpha. This constrained model yields a point estimate of the common alpha coefficient and a standard error for it. Using then the R-function “ci.alpha” in the appendix on that estimate and standard error, furnishes a CI for the common alpha coefficient in the studied population (for any prespecified confidence level). We stress that this case of between-class identity in alphas is the only circumstance when one can meaningfully speak of a coefficient alpha for a given measurement instrument and population of concern (cf. Raykov & Marcoulides, 2015; Raykov et al., 2016).
2. As implied from the introductory section and earlier discussion in this article, its purpose is not to address, relate to, or use the complex relationship that exists between Cronbach’s coefficient alpha and the scale reliability coefficient (for details on this relationship, see e.g., Raykov, 1997, 2012; Raykov & Marcoulides, 2017). In particular, because of alpha not being in general identical to reliability (Novick & Lewis, 1967), it is possible that one could speak of a single coefficient alpha but not of a single reliability coefficient for a given scale and heterogeneous population, or possibly conversely (cf. Raykov & Marcoulides, 2015).

References

- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis*. Upper Saddle River, NJ: Prentice Hall.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, 29, 81-117.
- Muthén, L. K., & Muthén, B. O. (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. (2012). Scale development using structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 472-492). New York, NY: Guilford Press.
- Raykov, T., & Marcoulides, G. A. (2004). Using the delta method for approximate interval estimation of parametric functions in covariance structure models. *Structural Equation Modeling, 11*, 659-675.
- Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum.
- Raykov, T., & Marcoulides, G. A. (2008). *An introduction to applied multivariate analysis*. New York, NY: Taylor & Francis.
- Raykov, T., & Marcoulides, G. A. (2010). Group comparisons in the presence of missing data using latent variable modeling techniques. *Structural Equation Modeling, 17*, 135-149.
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York, NY: Taylor & Francis.
- Raykov, T., & Marcoulides, G. A. (2015). Scale reliability evaluation in heterogeneous populations. *Educational and Psychological Measurement, 75*, 875-892.
- Raykov, T., & Marcoulides, G. A. (2017). Thanks coefficient alpha, we still need you! *Educational and Psychological Measurement, 79*, 200-210. doi:10.1177/0013164417725127
- Raykov, T., Marcoulides, G. A., & Li, T. (2016). Evaluation of criterion validity in heterogeneous populations. *Educational and Psychological Measurement, 76*, 1026-1044.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*, 107-120.
- Wang, J., & Wang, X. (2012). *Structural equation modeling. Applications using Mplus*. New York, NY: Wiley.