

# A Direct Latent Variable Modeling Based Method for Point and Interval Estimation of Coefficient Alpha

Educational and Psychological  
Measurement

2015, Vol. 75(1) 146–156

© The Author(s) 2014

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0013164414526039

epm.sagepub.com



Tenko Raykov<sup>1</sup> and George A. Marcoulides<sup>2</sup>

## Abstract

A direct approach to point and interval estimation of Cronbach's coefficient alpha for multiple component measuring instruments is outlined. The procedure is based on a latent variable modeling application with widely circulated software. As a by-product, using sample data the method permits ascertaining whether the population discrepancy between alpha and the composite reliability coefficient may be practically negligible for a given empirical setting. The outlined approach is illustrated with numerical data.

## Keywords

coefficient alpha, confidence interval, latent variable modeling, population discrepancy

Cronbach's coefficient alpha ( $\alpha$ ; Cronbach, 1951) is currently one of the most frequently used psychometric indexes in the educational, behavioral, and social sciences. Its history can be traced back almost a century to the work of Kuder and Richardson (1937) and their formulas KR20 and KR21, as well as to Guttman's subsequent research on scale reliability (Guttman, 1945; cf. McDonald, 1999). Over the past few decades,  $\alpha$  has attracted a great deal of attention and interest among

---

<sup>1</sup>Michigan State University, East Lansing, MI, USA

<sup>2</sup>University of California at Santa Barbara, CA, USA

## Corresponding Author:

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

Email: raykov@msu.edu

behavioral and social researchers and has been mostly used as an index possibly informing about the reliability of overall sum scores from multicomponent measuring instruments (Miller, 1995; Raykov & Marcoulides, 2011).

At the same time,  $\alpha$  has been the focus of numerous critical discussions in the psychometric literature, ranging from contained to outright heated criticism (see, e.g., Raykov, 2012, for a balanced treatment). In particular, as has been demonstrated more than four decades ago, with uncorrelated errors population  $\alpha$  does not exceed the population composite reliability unless the scale's components are essentially tau-equivalent, when the two coefficients are identical (Novick & Lewis, 1967). Similarly, population  $\alpha$  can be very close to the population scale reliability for unidimensional scales with uncorrelated errors and uniformly high loadings on the common construct (for details, see Raykov, 1997). Therefore, under the latter circumstances—which are empirically examinable/testable (see below)—point and interval estimation of the scale reliability coefficient can be accomplished in practical terms by point and interval estimation of coefficient alpha (see Raykov, West, & Traynor, 2013).

With this impressive interest in  $\alpha$  by substantive and methodologically oriented scholars, it seems desirable to have a routinely applicable means of its point and interval estimation that can be readily used by empirical researchers with widely circulated software. The purpose of the present note is to outline an estimation approach accomplishing this aim. The approach is based on the latent variable modeling (LVM) methodology and at the software level employs the increasingly popular LVM program Mplus (Muthén & Muthén, 2012). As a by-product of the approach, one can also ascertain for a given empirical setting if the discrepancy between  $\alpha$  and reliability at large is negligible.

## A Parameterization Framework for Point and Interval Estimation of Alpha With Popular Latent Variable Modeling Software

Denote by  $X_1, X_2, \dots, X_p$  ( $p > 2$ ) a set of observed measures, such as the components of a test, scale, testlet, subscale, inventory, self-report, survey, or questionnaire part (generically referred to as “scale” in this article).<sup>1</sup> These devices represent highly popular means of measuring underlying latent dimensions in contemporary educational, behavioral, and social research. Coefficient alpha is defined in a studied population as

$$\alpha = \frac{p}{p-1} \left[ 1 - \frac{\sum_{i=1}^r \text{Var}(X_i)}{\text{Var}(X)} \right] = \frac{p}{p-1} \cdot \frac{\sum_{i \neq j} \text{Cov}(X_i, X_j)}{\text{Var}(X)}, \quad (1)$$

where  $X = X_1 + X_2 + \dots + X_p$  is the overall scale sum (composite) score that is often of main interest in empirical work, whereas  $Var(\cdot)$  and  $Cov(\cdot, \cdot)$  denote variance and covariance of the variables involved (Cronbach, 1951).

The remainder of this article uses instrumentally the classical test theory framework (e.g., Zimmerman, 1975). Accordingly, each observed component (measure) score  $X_i$  is decomposable into the sum of true score  $T_i$  and error score  $E_i$ :

$$X_i = T_i + E_i \quad (i = 1, \dots, p).$$

In case of scale homogeneity (unidimensionality), the following congeneric test model holds (Jöreskog, 1971):

$$\underline{X} = \underline{a} + \underline{b}T + \underline{E}, \quad (2)$$

where  $\underline{X}$  is the  $p \times 1$  vector of observed measures (scale components),  $T$  denotes the assumed common latent construct evaluated by the  $p$  measures in question and with variance set at 1 for model identification,  $\underline{a}$  is a  $p \times 1$  vector of intercepts,  $\underline{b}$  is the  $p \times 1$  vector of manifest variable loadings on  $T$ , and  $\underline{E}$  is the  $p \times 1$  vector of error terms with zero means and uncorrelated with  $T$ . (For  $T$  one could take the true score of the first measure,  $X_1$ , without loss of generality in the following discussion; underlining is used to denote vector in this article.)

We note that in the setting of concern to this article, the congeneric Model (2) is not empirically distinguishable from the single-factor model that is frequently used in applications. With this in mind, to achieve its aims the remainder of the note considers a special case of the single-factor model, formally with  $m = p$  underlying factors that (a) have a unit loadings each on a corresponding observed component and (b) are associated with zero error variances (e.g., Raykov & Marcoulides, 2010):

$$\underline{Y} = \underline{a} + \Lambda \underline{\xi} + \underline{\delta}, \quad (3)$$

where  $\Lambda = I_p$  is the  $p \times p$  identity matrix,  $\underline{\xi}$  is a  $p \times 1$  vector of (dummy) latent variables, and  $\underline{\delta}$  is a  $p \times 1$  vector of error terms with zero variance each.<sup>2</sup> From Equation (3) follows directly

$$Cov(\underline{Y}) = Cov(\underline{\xi}), \quad (4)$$

with  $Cov(\cdot)$  denoting the covariance matrix for the vector in parentheses, as well as that the model in Equation (4) is saturated and with perfect fit to a given data set (observed means, variances, and covariances). Therefore,  $\alpha$  can be represented as

$$\alpha = \frac{p}{p-1} \left[ 1 - \frac{\sum_{i=1}^p Var(\xi_i)}{Var(X)} \right] = \frac{p}{p-1} \left[ 1 - \frac{\sum_{i=1}^p Var(\xi_i)}{\sum_{i=1}^p Var(\xi_i) + 2 \sum_{1 \leq i < j \leq p} Cov(\xi_i, \xi_j)} \right] \quad (5)$$

(Raykov et al., 2013).

## Evaluation of Coefficient Alpha Using Latent Variable Modeling With Popular Software

The implications of the preceding discussion for the aims of this article are as follows. Within the framework of the model defined in Equation (3), which is readily accessed using LVM,  $\alpha$  is a function only of the variances and covariances of the dummy latent variables  $\xi$ . These variances and covariances are however the parameters of Model (3), which are estimated when fitting it to data. A point estimate of  $\alpha$  results then as follows (with hat denoting parameter estimate):

$$\hat{\alpha} = \frac{p}{p-1} \left[ 1 - \frac{\sum_{i=1}^p V\hat{ar}(\xi_i)}{\sum_{i=1}^p V\hat{ar}(\xi_i) + 2 \sum_{1 \leq i < j \leq p} C\hat{ov}(\xi_i, \xi_j)} \right]. \quad (6)$$

This estimate of coefficient alpha is readily obtained with Mplus, and the needed source code is provided in Appendix A (as applicable to the example in the illustration section). We stress that while point estimates of  $\alpha$  are easily furnished with other popular software (such as SPSS, Stata, or SAS, to name a few), there are two advantages associated with the presently described approach. One, a confidence interval for coefficient alpha is readily obtainable within the same LVM framework. Indeed, this is straightforwardly achieved by employing the popular bootstrap approach (Efron & Tibshiriani, 1993), and the Mplus source code in Appendix A includes also the pertinent request for such an interval estimate. As a second advantage, based on sample data, one can ascertain as a by-product of this approach whether the population discrepancy between  $\alpha$  and the scale reliability coefficient is negligible for a given empirical setting (see Note 2).

As discussed in detail in Raykov (1997), the population slippage or discrepancy of coefficient alpha from the population scale reliability coefficient can be substantial depending on model parameters and their relative magnitude, but need not always be large, as indicated earlier in this note. Therefore, it is important for a researcher considering use of  $\alpha$  as an index of reliability, to be in a position to ascertain if in an empirical setting he or she is in a situation where this slippage is minimal and for practical reasons ignorable. A procedure examining empirically this slippage and responding to the question whether the alpha-reliability discrepancy is negligible is provided in Raykov et al. (2013), where it was developed in the context of complex sample designs. A readily used version of its associated Mplus command file for the single-level case with no design variables, which is of concern in this note, is given in Appendix B. This procedure is directly used on a given data set from a unidimensional scale with uncorrelated errors, when one is interested in responding to the query whether  $\alpha$  can be practically used as a substitute of the scale reliability coefficient (see also Note 2).

The discussed approach to point and interval estimation of coefficient alpha and examining if the population alpha-reliability discrepancy is practically ignorable for an empirical setting is demonstrated next on numerical data.

### Illustration on Data

For the demonstration aims of this section, we use a simulated data set from a scale with  $p = 5$  components, which was generated according to the following model (cf. Equation 2):

$$\begin{aligned} X_1 &= 3 + .75 T + \varepsilon_1, \\ X_2 &= 3 + .75 T + \varepsilon_2, \\ X_3 &= 3 + .80 T + \varepsilon_3, \\ X_4 &= 3 + .85 T + \varepsilon_4, \\ X_5 &= 3 + .85 T + \varepsilon_5, \end{aligned} \tag{7}$$

where  $\eta$  was standard normal and  $\varepsilon_1$  through  $\varepsilon_7$  were independent normal zero-mean variates with variance .5 each.

Since the average loading used in this data generation process is above .7 and the discrepancies of the individual component loadings from it are all well below .2 in absolute value (as is readily seen in Equations 7), from Raykov (1997) it follows that an upper bound of the population difference between  $\alpha$  and the reliability coefficient of the sum  $X = X_1 + \dots + X_5$  is .02. Hence, the population discrepancy between  $\alpha$  and the composite reliability coefficient is practically negligible, that is,  $\alpha$  effectively equals the scale reliability at large. Given that we know the population parameters underlying the data simulation process, we can also work out here the population reliability coefficient,  $\rho$ , as follows (e.g., Bollen, 1989; an asterisk is used to denote multiplication next):

$$\rho = (.75 + .75 + .8 + .85 + .85)^2 / [(.75 + .75 + .8 + .85 + .85)^2 + 5 * .5] = .865. \tag{8}$$

The covariance matrix associated with the simulated data set on the 5 scale components under consideration (Equations 7) is presented in Table 1.

To apply the outlined procedure for point and interval estimation of coefficient alpha, we fit Model (3) to these data (see Mplus source code in Appendix A). As mentioned earlier, this model is saturated and hence associated with perfect overall fit. The model yields for  $\alpha$  the estimate  $\hat{\alpha} = .863$ , with a 95% confidence interval (CI) of (.849, .874). We note in passing that, as expected (see above in this section), the estimate of coefficient alpha is quite close to the population reliability coefficient  $\rho$  and in fact both are practically identical (see also Equation 8). In addition, we observe that the 95% CI for alpha obtained with this procedure also covers that population reliability coefficient.

**Table 1.** Covariance Matrix of Analyzed Data Set.

Variable	$Y_1$	$Y_2$	$Y_3$	$Y_4$	$Y_5$	$Y_6$	$Y_7$
$Y_1$	1.019						
$Y_2$	0.575	1.077					
$Y_3$	0.564	0.602	1.136				
$Y_4$	0.628	0.635	0.683	1.231			
$Y_5$	0.625	0.651	0.660	0.741	1.248		

Note. Since reliability and coefficient alpha do not depend on observed variable intercepts (denoted  $a$  in Equations 2 and 3), observed variable means are inconsequential for the discussion in this note and thus not presented.

In empirical research, one typically does not know the population reliability coefficient. Hence, if he or she is considering use of the alpha estimate as such of composite reliability, it would be necessary that the researcher ascertains first if he/she is in a situation where based on the sample data it can be suggested that the population discrepancy between alpha and reliability is practically ignorable (see Note 2). As discussed in detail in Raykov (1997), for unidimensional scales with uncorrelated errors this will be the case if in the studied population the average loading is above .7 and the largest (in absolute value) component loading difference from that average is below .2, after setting at 1 the underlying true score variance (for model identification purposes).

To ascertain if this is also the case for the presently considered data set, we apply the above-mentioned single-level, no-design-variable version of the Raykov et al. (2013) procedure to examine whether the population alpha-reliability slippage is practically ignorable here. To this end, we first fit the unidimensional model (single-factor model) to this data set, and find that it is tenable: chi-square = 4.259, degrees of freedom = 5,  $p = .513$ , root mean square error of approximation (RMSEA) = 0, with a 90% CI (0, .041; see Mplus source code in Appendix B). The resulting parameter estimates, with standard errors, are presented in Table 2. In this model, as mentioned earlier, of particular interest are the loading estimates. Their average and component differences from it, along with their standard errors and 95% CIs, are presented in Table 3, with these standard errors and CIs being obtained with the Raykov et al. (2013) test version that is of relevance here (see also Appendix B and notes to it).

As seen from Table 3, (a) the CI for the average loading is entirely above .7 and (b) the CI of each difference between component loading and that average is entirely well within the interval  $(-.2, .2)$ . Hence, the presently considered data set is in fact an example where (with a high degree of confidence one may suggest that) the difference between coefficient alpha and scale reliability is practically negligible at large. Therefore,  $\alpha$  can be used in the role of a scale reliability coefficient in this example.

With this in mind, we deduce that the scale score  $X$  defined as the sum of the 5 components  $X_1$  through  $X_5$  under consideration in this section (see Equations 7) represents a composite with satisfactory reliability that is estimated essentially as equal to .863 and with a 95% CI of (.849, .874).

**Table 2.** Parameter Estimates, Standard Errors, Their Ratios, and  $p$  Values for the Fitted Unidimensional Model.

Parameter	Estimate	SE	Estimate/SE	$p$ value
Loadings				
Y1	0.734	0.027	27.416	.000
Y2	0.762	0.027	27.850	.000
Y3	0.782	0.031	24.916	.000
Y4	0.858	0.032	26.506	.000
Y5	0.854	0.032	26.724	.000
Intercepts				
Y1	2.944	0.031	94.635	.000
Y2	3.042	0.033	91.245	.000
Y3	3.041	0.034	88.310	.000
Y4	2.974	0.035	84.043	.000
Y5	3.013	0.035	85.747	.000
Error variances				
Y1	0.480	0.025	19.023	.000
Y2	0.497	0.029	17.221	.000
Y3	0.524	0.029	17.844	.000
Y4	0.495	0.031	15.977	.000
Y5	0.519	0.030	17.304	.000
Latent variance:				
F	1.000 <sup>a</sup>			

Note. Software parameter estimate presentation used.

a. Parameter fixed for model identification (cf. Raykov, 1997).

## Conclusion

This note was concerned with the highly popular Cronbach coefficient alpha in the educational, behavioral, and social sciences. A directly applicable LVM procedure was outlined that can be readily used with widely circulated LVM software to point and interval estimate coefficient alpha. As a by-product, the approach can also be used to ascertain if one is in an empirical situation where alpha and the reliability of a considered scale can be treated as practically identical at large. While in the general case coefficient alpha is not a consistent estimator of composite reliability and has a number of downsides (discussed in detail, e.g., in Raykov, 2012), under unidimensionality and uncorrelated errors alpha can be very close to reliability at large if the average construct loading—given unitary latent variance—is in excess of .7 and the component loading deviations from it are within the interval  $(-.2, .2)$ . This set of conditions can be examined/tested using LVM and the procedure outlined in this note (see also Note 2).

A limitation of the modeling approach underlying this note is the requirement for large samples. The reason is that it is grounded in the LVM methodology that itself relies critically on asymptotic theory (Muthén, 2002). Similarly, the bootstrap is a large-sample method for sampling distribution approximation that is essential for the present approach when it comes to interval estimation of coefficient alpha. Therefore,

**Table 3.** Average Loading and Its Difference From Individual Component Loadings in Fitted Unidimensional Model—Point and Interval Estimates (see also Table 2; cf. Raykov, West, & Traynor, 2013).

Parameter	Estimate	SE	95% Confidence interval
Average loading			
AVE_LAM	0.798	0.020	(0.739, 0.838)
Individual loading differences from average loading			
L1A	-0.064	0.022	(-0.107, -0.023)
L2A	-0.036	0.022	(-0.084, 0.004)
L3A	-0.016	0.023	(-0.062, 0.028)
L4A	0.060	0.024	(0.016, 0.106)
L5A	0.056	0.023	(0.012, 0.098)

**Note.**

1. AVE\_LAM = average loading; LA1-LA5 = differences between component loading (first through fifth, respectively) and average loading.
2. Check if (i) the confidence interval of the average loading is entirely above .7 and (ii) the confidence intervals of its differences from the individual loadings are entirely within (-.2, .2), to suggest population near identity of alpha and the scale reliability coefficient for a given empirical setting (Raykov, 1997, pp. 342-344). When both (i) and (ii) hold, the point and interval estimates of alpha and reliability can be treated as essentially interchangeable empirically, and the Mplus code in Appendix A can be used to point and interval estimate scale reliability by point and interval estimating coefficient alpha instead.

in the absence currently of sufficiently precise guidelines for determining desirable sample size, caution is advised when applying the approach outlined in this article with samples having less than several hundred observations.

A related limitation of the procedure in this note is the following. While it can be used for point and interval estimation of coefficient alpha with components having any scale (i.e., binary, ordinal, or interval scaled; see Mplus command file in Appendix A), its application for ascertaining whether alpha's slippage from reliability at large is negligible assumes that the components  $X_1, \dots, X_p$  are (approximately) continuous (see Mplus command file in Appendix B). With components having at least 5-7 possible values, the application of that part of the outlined procedure for point and interval estimation of alpha's slippage from reliability in a studied population (Appendix B), is possible by using robust maximum likelihood (cf. Raykov & Marcoulides, 2011).

In conclusion, this note outlined a readily and widely applicable LVM-based procedure for (a) point and interval estimation of the popular coefficient alpha using a widely circulated LVM software as well as (b) ascertaining, in a given empirical setting, if the alpha-scale reliability population discrepancy could be considered practically ignorable for a multicomponent measuring instrument under consideration.

## Appendix A

### *Mplus Source Code for Point and Interval Estimation of Coefficient Alpha*

TITLE: LVM POINT AND INTEVAL ESTIMATION OF COEFFICIENT ALPHA.



```

DATA:      FILE = <NAME OF RAW DATA FILE>;
VARIABLE:  NAMES = Y1-Y5;
ANALYSIS:  BOOTSTRAP = 2000;
MODEL:     KSI1 BY Y1@1; ! SEE EQUATIONS (3) AND THEIR IMMEDIATELY
           KSI2 BY Y2@1; ! FOLLOWING DISCUSSION.
           KSI3 BY Y3@1;
           KSI4 BY Y4@1;
           KSI5 BY Y5@1;
           Y1-Y5@0;
           KSI1-KSI5 (S1-S5);
           KSI1 WITH KSI2-KSI5 (S12-S15);
           KSI2 WITH KSI3-KSI5 (S23-S25);
           KSI3 WITH KSI4-KSI5 (S34-S35);
           KSI4 WITH KSI5 (S45);
MODEL CONSTRAINT:
           NEW(ALPHA, P, SC, SV);
           P = 5; ! enter here number of components in scale
           SC = 2*(S12+S13+S14+S15+S23+S24+S25+
           S34+S35+S45); ! modify correspondingly for  $p \neq 5$ ,
           SV = S1+S2+S3+S4+S5; ! as well as here.
           ALPHA = P/(P-1)*SC/(SC+SV); ! SEE EQUATION (1).
OUTPUT:    CINTERVAL(BCBOOTSTRAP);

```

*Note.* After the title for the analysis and naming the raw data file, names are assigned in the VARIABLE section. The ANALYSIS and OUTPUT sections request the bias-corrected bootstrap confidence intervals (at 90%, 95%, and 99% levels). The MODEL section defines Model 3 (see discussion immediately after Equation 3). The MODEL CONSTRAINT section introduces first place-holders for coefficient alpha, the sum of component variances and of their covariances, and then defines alpha following Equation (1). (Modify the expressions for SV and SC by adding/deleting applicable component variances and covariances, in case  $p \neq 5$ .)

## Appendix B

```

TITLE:     VERSION FOR THE SINGLE-LEVEL, NO-DESIGN-VARIABLE CASE OF
           THE RAYKOV, WEST, & TRAYNOR PROCEDURE FOR ASCERTAINING IF
           ONE COULD USE ALPHA FOR RELIABILITY IN A GIVEN EMPIRICAL
           SETTING (ESTIMATES CORRESPONDING POPULATION DISCREPANCY
           BETWEEN ALPHA AND SCALE RELIABILITY, AND EXAMINES IF IT IS
           IGNOREABLE; cf. Raykov, 1997, pp. 342-344; Raykov et al.,
           2013) .
DATA:      FILE = <NAME OF RAW DATA FILE>;
VARIABLE:  NAMES = Y1-Y5;
ANALYSIS:  BOOTSTRAP = 2000;

```

```

MODEL:      F BY Y1*(L1) ! defines single-factor (cong. test) model (2)
            Y2-Y5 (L2-L5) ;
            F@1; ! cf. Raykov (1997)- suitable model identif. constr.
MODEL CONSTRAINT:
            NEW(AVE_LAM, L1A, L2A, L3A, L4A, L5A) ;
            AVE_LAM = (L1+L2+L3+L4+L5)/5; ! average loading
            L1A = L1-AVE_LAM; ! difference of 1st to average loading
            L2A = L2-AVE_LAM; ! 2nd
            L3A = L3-AVE_LAM; ! 3rd
            L4A = L4-AVE_LAM; ! 4th
            L5A = L5-AVE_LAM; ! 5th
OUTPUT:     CINTERVAL(BCBOOTSTRAP) ;

```

*Note.* The MODEL CONSTRAINT section is used to examine if alpha and reliability are nearly identical at large (Raykov, 1997). For this to be the case, check in the output if (a) the confidence interval (at a prespecified confidence level, such as say 95%) for the average loading is entirely above .7, and (b) the confidence intervals for the component loading-to-average discrepancies are each entirely within (−.2, .2). (The metric of these numerical comparisons is set by fixing at 1 the latent variance—see last line of the MODEL section.)

## Acknowledgment

We are grateful to B. T. West for valuable discussions on estimation of coefficient alpha.

## 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. If  $p = 2$ , additional identifying restrictions will be needed, for instance, indicator loading equality (true score-equivalent/essentially tau-equivalent measures) and/or error variance equality (e.g., parallel measures; Raykov & Marcoulides, 2011). As implied by a scale's consideration, it is further assumed in this article that not all elements of  $\underline{b}$  in Equation (2) in the main text are zero, a condition easily fulfilled in empirical behavioral and social research.
2. As elaborated below, based on sample data one can examine whether the corresponding population slippage of coefficient alpha from the scale reliability coefficient is ignorable for a given empirical setting. (The conclusion about the pertinent population slippage of  $\alpha$

from reliability is drawn using sample data, as is typical in applications of statistics, and in this sense does not require availability of population data on the scale under consideration.)

## References

- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Efron, B., & Tibshiriani, R. J. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255-282.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36, 109-133.
- Kuder, G. F., & Richardson, M. W. (1937). The theory and estimation of test reliability. *Psychometrika*, 2, 151-160.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling*, 3, 255-273.
- Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, 29, 81-117.
- Muthén, L. K., & Muthén, B. (2012). *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 with fixed congeneric components. *Multivariate Behavioral Research*, 32, 329-353.
- Raykov, T. (2012). Scale construction and development using structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 472-492). New York, NY: Guilford.
- 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., West, B. T., & Traynor, A. (2013). *Evaluation of coefficient alpha in complex sample design studies*. Manuscript submitted for publication.
- Zimmerman, D. W. (1975). Probability measures, Hilbert spaces, and the axioms of classical test theory. *Psychometrika*, 30, 221-232.