

# Formulas of Multilevel Reliabilities for Tests with Ordered Categorical Responses



Zhenqiu (Laura) Lu, Minju Hong, and Seohyun Kim

## 1 Introduction

Reliability is a measure of overall internal consistency of a test. It has been widely used in statistical, psychological, educational, social and behavioral research when data are collected by responding to items in a test or a questionnaire (Bollen, 1989; Finney & DiStefano, 2006). A high value of reliability indicates the measure provides similar, reliable, and stable results under consistent conditions. There are many approaches that have been proposed to estimate reliabilities. Among them, the classical test theory (CTT) approach has been widely used, and the Cronbach alpha is the most popular reliability. But it only measures the lower bound on the consistency of a test (Green et al., 1977; Novick & Lewis, 1967; Sijtsma, 2009). Another approach using structural equation modeling (SEM) has been proposed to obtain more accurate reliabilities (Bentler, 2009; Bollen, 1989; Green & Yang, 2009; Miller, 1995; Raykov, 1997; Raykov & Shrout, 2002). But these methods focus on addressing continuous outcomes.

For tests with ordered categorical responses, Green and Yang (2009) proposed a nonlinear reliability coefficient within an SEM framework, and the nonlinear reliability has been found to be more accurate than the linear reliability that treats categorical scores as continuous. But they only considered the items with the same number of categories. Kim, Lu and Cohen (2020) extended their research to broader situations and proposed a general formula for reliabilities. But these formulas are only for single level data structure. And their research did not consider the composite

---

Z. Lu (✉) · M. Hong  
University of Georgia, Athens, GA, USA  
e-mail: [zlu@uga.edu](mailto:zlu@uga.edu)

S. Kim  
University of Virginia, Charlottesville, VA, USA

reliability or the coefficient Omega (McDonald, 1985) and the maximal reliability H of weighted sum (Bentler, 2007) for tests with categorical responses. So far, there has been no research on this topic.

In order to fill the gap, the current study reviewed various approaches to reliabilities, extended single level reliabilities to multilevel reliabilities, and provided closed-form formulas for multilevel nonlinear SEM reliabilities for tests with ordered categorical responses via a multilevel confirmatory factor analysis (MCFA) approach. Multilevel alpha was also considered.

## 2 Reliabilities

### 2.1 CTT Approach

Suppose there are  $J$  items in a test. In classical test theory (CTT), an observed score  $X_j$  on item  $j$  ( $j = 1, \dots, J$ ) is composed of two uncorrelated components, a latent true score or trait,  $T_j$ , and an error score,  $\epsilon_j$ , with mean of 0:

$$X_j = T_j + \epsilon_j$$

Let  $X$ ,  $T$  and  $\epsilon$  be the sum of observed scores, of true scores, and of error scores, respectively, across  $J$  items. Then

$$X = X_1 + X_2 + \dots + X_J = \sum_{j=1}^J X_j,$$

and  $T = \sum_{j=1}^J T_j$ ,  $\epsilon = \sum_{j=1}^J \epsilon_j$ , so we have  $X = T + \epsilon$ .

We want to make sure how much of variance of observed score is due to the latent true score versus the error. One measure is to use reliability. The reliability coefficient of a test is defined as the ratio of the true variance to the total variance, which is the sum of the true variance and the error variance. Mathematically, it is

$$\rho = \frac{\sigma_T^2}{\sigma_x^2},$$

where  $\sigma_T^2$  is the variance of  $T$ , and  $\sigma_x^2$  is the variance of  $X$  (Lord & Novick, 1968). The reliability quantifies the proportion or ratio. It is an estimation of how much random error might be in the scores around the true score.

**Cronbach Alpha** Under CTT, there are many ways to estimate reliability: test-retest, alternative forms, split-half, Spearman-Brown prophecy formula, and the Cronbach alpha (or coefficient Alpha). Among them, the Cronbach alpha (Cronbach, 1951) is the most commonly used. The Cronbach alpha is defined as

$$\alpha = \frac{J^2 \bar{\sigma}_{xx'}}{\sigma_x^2}$$

where  $J$  is the total number of items,  $\sigma_x^2$  is the variance of observed scores of the test  $X$ , and  $\bar{\sigma}_{xx'}$  is the mean of off-diagonal covariance between two parallel tests  $X$  and  $X'$ . Specifically, the alpha can be calculated as

$$\alpha = \frac{J^2 * \frac{\sum_i \sum_{j,i < j} \sigma_{x_i, x_j}}{\frac{J(J-1)}{2}}}{\sigma_x^2}$$

where  $\sum_i \sum_{j,i < j} \sigma_{\varepsilon_i, \varepsilon_j}$  is the sum of lower (or upper) off-diagonal covariance between items  $i$  and  $j$ .

### 2.2 CFA Approach

However, the Cronbach alpha is only a lower bound on the internal consistency of the test (Green et al., 1977; Novick & Lewis, 1967; Sijtsma, 2009). To get a better estimate, another approach to reliability is the structural equation modeling (SEM) approach. Specifically, this approach uses confirmatory factor analysis (CFA) to estimate reliability. Tests are assumed to have underlying factorial structure (factors, e.g., reading ability, math ability, or personality).

$$X_j^* = \lambda_{1j}\eta_1 + \lambda_{2j}\eta_2 + \dots + \lambda_{Mj}\eta_M + e_j,$$

where  $X_j^*$  is a continuous score for item  $j$ ,  $M$  is the number of latent factors,  $\eta_m$  ( $1 \leq m \leq M$ ) are latent factors weighted by corresponding factor loadings  $\lambda_m$ , and  $e_j$  is a measurement error term. We assume errors are independent and have variance  $\sigma_{\varepsilon_j}^2$  for item  $j$ .

Suppose there are  $J$  items in the test and  $X^*$  and  $T$  are the sum of observed scores and of true scores, respectively. Then  $X^* = \sum_{j=1}^J X_j^*$  and  $T = \sum_{j=1}^J \sum_{m=1}^M \lambda_{mj}\eta_m$ . Because item scores are presented by a confirmatory factor analysis model, the linear reliability  $\rho_{lin}$  is calculated as the ratio of true sum score variance to observed sum score variance where the true score variance is estimated using the CFA model above.

$$\rho_{lin} = \frac{\sigma_T^2}{\sigma_{X^*}^2} = \frac{Var\left(\sum_{j=1}^J \sum_{m=1}^M \lambda_{mj}\eta_m\right)}{Var\left(\sum_{j=1}^J X_j^*\right)}.$$

Here  $\rho_{lin}$  measures the linear proportion of observed sum score variance that is attributed to the latent factors,  $\eta_1, \eta_2, \dots, \eta_M$  (Bollen, 1989).

### Composite Reliability/Coefficient Omega $\omega$

If item scores have only one factor, then  $M = 1$  and the reliability above becomes

$$\rho_{lin} = \frac{\sigma_T^2}{\sigma_{X^*}^2} = \frac{Var\left(\sum_{j=1}^J \lambda_j \eta\right)}{Var\left(\sum_{j=1}^J X_j^*\right)}.$$

And if we assume that latent factor have unit variance  $Var(\eta) = 1$ , then the reliability  $\rho_{lin}$  is referred to as composite reliability or coefficient omega  $\omega$  (McDonald, 1985)

$$\omega = \frac{(\sum \lambda_{x_j})^2}{(\sum \lambda_{x_j})^2 + \sum \sigma_{\varepsilon_j}^2}$$

where  $\sum \lambda_{x_j}$  is a sum of the factor loading of item  $j$ , and  $\sum \sigma_{\varepsilon_j}^2$  is a sum of all error variances.

### Maximal Reliability $H$ for Weighted Sum

Composite reliability represents the relation between a scale's underlying latent factor and its unit-weighted composite, but a scale's unit-weighted composite may not optimally reflect its underlying latent construct. The true score variance estimated in factor analysis allows for heterogeneous indicator weights, so it is reasonable to allow heterogeneous weights when creating a scale's composite score.

$$X = w_1 X_1 + w_2 X_2 + \dots + w_J X_J = \sum_{j=1}^J w_j X_j$$

One approach to comparing true score variance for one common factor to the variance of a unit-weighted scale is presented as maximal reliability  $H$  (e.g., Bentler, 2007). When the weight vector

$$W = \left(\lambda' \psi^{-1} \lambda\right)^{-1/2} \psi^{-1} \lambda,$$

where  $\lambda$  is the factor loading matrix and  $\psi$  is the residue variance matrix, then the maximal reliability is given by

$$\rho_{w/x(\max)} = \frac{\lambda' \psi^{-1} \lambda}{\lambda' \psi^{-1} \lambda + 1}$$

By assuming that the variance of each item is 1, the standardized version of maximal reliability for a single common factor model can be expressed as follows (e.g., Hancock & Mueller, 2001; Geldhof et al., 2014).

$$H = \frac{\sum \frac{\lambda_{x_j}^2}{\sigma_{\varepsilon_j}^2}}{1 + \sum \frac{\lambda_{x_j}^2}{\sigma_{\varepsilon_j}^2}} = \frac{\sum \frac{\lambda_{x_j}^2}{1 - \lambda_{x_j}^2}}{1 + \sum \frac{\lambda_{x_j}^2}{1 - \lambda_{x_j}^2}} = \frac{1}{1 + \frac{1}{\sum \frac{\lambda_{x_j}^2}{1 - \lambda_{x_j}^2}}}$$

where  $\lambda_{x_j}^2$  is the squared standardized factor loading of item  $j$ , and  $\sigma_{\varepsilon_j}^2$  is the error variance of item  $j$ .

### 3 Multilevel Reliabilities

The data collected from social and educational areas often have multilevel structure. In these cases, multilevel reliabilities for multilevel structure were proposed.

#### 3.1 Multilevel Confirmatory Factor Analysis (MCFA)

Muthén (2011) defined a multilevel confirmatory factor analysis (MCFA) by assuming a one-factor model holds for both the between and the within components. The observed value of the  $p$ -dimensional variable  $y_{gi}$  is partitioned into three components:

$$y_{gi} = V + y_{Bg} + y_{wgi}$$

where  $y_{gi}$  is the observed value of individual  $i$  in group  $g$ ,  $V$  is a grand mean,  $y_{Bg}$  is the between-group part of the observed value, and  $y_{wgi}$  is the within-group part of the observed value. The multilevel CFA specifies a model at between-group level and within-group level separately. Suppose there are  $h$  factors between groups and  $m$  factors within groups, the between-group level CFA model is

$$y_{Bg} = \Lambda_{Bg} \eta_{Bg} + \varepsilon_{Bg}$$

where  $\Lambda_{Bg}$  is a  $(p \times h)$  matrix of factor loadings with elements  $\lambda_{Bg}$ 's,  $\eta_{Bg}$  is a  $h$ -dimensional vector of factor scores with the assumption of  $\eta_{Bg} \sim MN(0_h, \Psi_{h \times h})$ , and  $\varepsilon_{Bg}$  is a  $p$ -dimensional vector of errors with the assumption of  $\varepsilon_i \sim MN(0_p, \Phi_{p \times p})$ . And the within-group CFA model level is defined as

$$y_{wgi} = \Lambda_{wg} \eta_{wgi} + \varepsilon_{wgi}$$

where  $\Lambda_{wg}$  is a  $(p \times m)$  matrix of factor loadings with elements  $\lambda_{wg}$ 's,  $\eta_{wgi}$  is a  $m$ -dimensional vector of factor scores with the assumption of

$\eta_{wgi} \sim MN(0_m, \Psi_{m \times m})$ , and  $\varepsilon_{wgi}$  is a  $p$ -dimensional vector of errors with the assumption of  $\varepsilon_i \sim MN(0_p, \Phi_{p \times p})$ .

If there is only one factor either between-group or within-group, then the variance of the observed variable  $y$ ,  $\sigma_{y_{gi}}^2$ , is decomposed (Muthén, 2011) as

$$\begin{aligned} \sigma_{y_{gi}}^2 &= \lambda_{Bg}^2 \sigma_{\eta_{Bg}}^2 + \sigma_{\varepsilon_{Bg}}^2 + \lambda_{wg}^2 \sigma_{\eta_{wgi}}^2 + \sigma_{\varepsilon_{wgi}}^2 \\ &= \sigma_{BF}^2 + \sigma_{BE}^2 + \sigma_{WF}^2 + \sigma_{WE}^2 \end{aligned}$$

where  $\lambda_{Bg}$  is between-group factor loadings,  $\sigma_{\eta_{Bg}}^2$  is the variances of between-group factor scores,  $\sigma_{\varepsilon_{Bg}}^2$  is the variances of between-group errors,  $\lambda_{wg}$  is within-group factor loadings,  $\sigma_{\eta_{wgi}}^2$  is the variances of within-group factor scores, and  $\sigma_{\varepsilon_{wgi}}^2$  is the variances of within-group errors,  $\sigma_{BF}^2$  is a between-level factor score variance,  $\sigma_{BE}^2$  is a between-level error variance,  $\sigma_{WF}^2$  is a within-level factor score variance, and  $\sigma_{WE}^2$  is an within-level error variance.

### 3.2 Multilevel Reliabilities

Applying MCFA to reliability calculation,  $p$  dimensions become  $J$  items, and we assume there is only one factor either between-group or within-group.

**Multilevel Alpha** The multilevel alpha is calculated as

$$\begin{aligned} \text{within - group level } \alpha &= \frac{J^2 * \bar{\sigma}_{wgi, wgj}}{\sigma_{wg}^2} \\ \text{between - group level } \alpha &= \frac{J^2 * \bar{\sigma}_{Bgi, Bgj}}{\sigma_{Bg}^2} \end{aligned}$$

where  $J$  is the number of items,  $\bar{\sigma}_{wgi, wgj}$  is the average of the within-group covariance between items  $i$  and  $j$ ,  $\sigma_{wg}^2$  is the variance of the within-group part of observed value,  $\bar{\sigma}_{Bgi, Bgj}$  is the mean of the between-group covariance between items  $i$  and  $j$ , and  $\sigma_{Bg}^2$  is the variance of the between-group part of observed value.

**Multilevel Omega** And multilevel omega is obtained as

$$\text{within - group level } \omega = \frac{(\sum \lambda_{wgj})^2}{(\sum \lambda_{wgj})^2 + \sum \sigma_{\varepsilon_{wgi}}^2}$$

$$\text{between - group level } \omega = \frac{(\sum \lambda_{Bgj})^2}{(\sum \lambda_{Bgj})^2 + \sum \sigma_{\varepsilon Bgj}^2}$$

where  $\sum \lambda_{wgj}$  is a sum of within-group level squared factor loading of item  $j$ , and  $\sum \sigma_{\varepsilon wgi}^2$  is a sum of within-group level error variances,  $\sum \lambda_{Bgj}$  is a sum of within-group level squared factor loading of item  $i$ , and  $\sum \sigma_{\varepsilon Bgj}^2$  is a sum of within-group level error variances.

**Multilevel  $H$**  Maximal  $H$  is calculated by

$$\text{within - group level } H = \frac{1}{1 + \frac{1}{\sum \frac{\lambda_{wgj}^2}{1 - \lambda_{wgj}^2}}}$$

$$\text{between - group level } H = \frac{1}{1 + \frac{1}{\sum \frac{\lambda_{Bgj}^2}{1 - \lambda_{Bgj}^2}}}$$

where  $\lambda_{wgj}^2$  is the squared standardized within-group factor loading of item  $j$ ,  $\lambda_{Bgj}^2$  is the squared standardized between-group factor loading of item  $j$ .

### 4 Multilevel Reliability for Categories Responses

It is very common in social and behavioral sciences that items have ordered categories. When the observed data are ordinal categorical, fitting linear SEM models using the linear estimation method is not desirable because it violates the assumption and provides inflated chi-square estimates and attenuated factor loadings (Bollen, 1989). To address this problem, we consider the observed categorical scores ( $X_j$ ) are from underlying continuous variables ( $X_j^*$ ) and the nonlinear relationship between  $X_j$  and  $X_j^*$  is

$$X_j = \begin{cases} C_j - 1, & \text{if } X_j^* \geq v_{C_j-1} \\ \vdots & \vdots \\ 1, & \text{if } v_1 \leq X_j^* < v_2 \\ 0, & \text{if } X_j^* < v_1. \end{cases}$$

where  $C_j$  is the number of categories for  $X_j$ , and the  $v_i$  ( $i = 1, 2, \dots, C_j - 1$ ) are the category thresholds. If  $X_j^*$  is less than  $v_1$ ,  $X_j$  is equal to 0, for  $v_1 \leq X_j^* < v_2$ ,  $X_j$  is equal to 1, and if  $X_j^*$  is above  $v_{C_j-1}$ ,  $X_j$  is equal to  $C_j - 1$ . If the structure of the test is well-specified, this approach can estimate reliability more accurately than the linear SEM approach.

**Multilevel Alpha for Tests with Categorical Responses** Single level alpha for categorical responses can be calculated by using polychoric correlations if we assume the variance of each item is 1. Reliability calculated from parallel measures. For multilevel alpha, both within- and between-group levels reliabilities are calculated.

**Multilevel Composite Reliability for Tests with Categorical Responses** Single level composite reliabilities have been proposed by Kim, Lu and Cohen (2020) to investigate reliability with items having the same or different numbers of ordered categories. For multilevel composite reliabilities, the same formula can be applied at both within- and between-group levels. Multilevel Omega for tests with categorical responses is a simplified version for one factor models.

**Multilevel Maximal Reliability for Tests with Categorical Responses** There has been no research on this topic done before. In this article, we derived the numerical formula as follows. Suppose  $X$  and  $\tilde{X}$  are two parallel tests, which are two weighted sums,  $X = \sum_{j=1}^J w_j X_j$  and  $\tilde{X} = \sum_{j'=1}^J w_{j'} \tilde{X}_{j'}$ . To estimate the reliability for the nonlinear measurement model, the correlation between  $X$  and  $\tilde{X}$  is used, which is

$$\rho_{X\tilde{X}} = \frac{Cov(X, \tilde{X})}{\sqrt{\text{var}(X)\text{var}(\tilde{X})}}$$

The numerator for weighted sums is

$$Cov(X, \tilde{X}) = Cov\left(\sum_{j=1}^J w_j X_j, \sum_{j'=1}^J w_{j'} \tilde{X}_{j'}\right) = \sum_{j=1}^J \sum_{j'=1}^J w_j w_{j'} Cov(X_j, \tilde{X}_{j'}),$$

in which

$$Cov(X_j, \tilde{X}_{j'}) = E(X_j \tilde{X}_{j'}) - E(X_j) E(\tilde{X}_{j'})$$



$$\begin{aligned}
 &= \left( \sum_{k=1}^{C_j-1} \sum_{l=1}^{C_{j'}-1} \Phi_2(v_{jk}, h_{j'_l}; \rho_M) - (C_{j'} - 1) \sum_{k=1}^{C_j-1} \Phi_1(v_{jk}) - (C_j - 1) \right. \\
 &\quad \times \left. \sum_{l=1}^{C_{j'}-1} \Phi_1(h_{j'_l}) + (C_j - 1)(C_{j'} - 1) \right) \\
 &\quad - \left( - \sum_{k=1}^{C_j-1} \Phi_1(v_{jk}) + (C_j - 1) \right) \left( - \sum_{k=1}^{C_{j'}-1} \Phi_1(h_{j'_l}) + (C_{j'} - 1) \right) \\
 &= \sum_{k=1}^{C_j-1} \sum_{l=1}^{C_{j'}-1} \Phi_2(v_{jk}, h_{j'_l}; \rho_M) - \sum_{k=1}^{C_j-1} \Phi_1(v_{jk}) \sum_{l=1}^{C_{j'}-1} \Phi_1(h_{j'_l})
 \end{aligned}$$

and

$$\rho_M = \sum_{m=1}^M \sum_{m'=1}^M \lambda_{mj} \lambda_{m'j'} \rho_{\eta_m \eta_{m'}}$$

The denominator is

$$\begin{aligned}
 Var(X) &= Var \left( \sum_{j=1}^J w_j X_j \right) = \sum_{j=1}^J \sum_{j'=1}^J w_j w_{j'} Cov(X_j, X_{j'}) \\
 &= \sum_{j=1}^J \sum_{j'=1}^J w_j w_{j'} \left( \sum_{k=1}^{C_j-1} \sum_{l=1}^{C_{j'}-1} \Phi_2(v_{jk}, h_{j'_l}; \rho_{X_j^* X_{j'}^*}) \right. \\
 &\quad \left. - \sum_{k=1}^{C_j-1} \Phi_1(v_{jk}) \sum_{l=1}^{C_{j'}-1} \Phi_1(h_{j'_l}) \right)
 \end{aligned}$$

Therefore the reliability is calculated from parallel measures of ordered categories responses

$$\rho_{Cat} = \frac{\sum_{j=1}^J \sum_{j'=1}^J w_j w_{j'} \left[ \sum_{k=1}^{C_j-1} \sum_{l=1}^{C_{j'}-1} \Phi_2(v_{jk}, h_{j'_l}; \rho_M) - \sum_{k=1}^{C_j-1} \Phi_1(v_{jk}) \sum_{l=1}^{C_{j'}-1} \Phi_1(h_{j'_l}) \right]}{\sum_{j=1}^J \sum_{j'=1}^J w_j w_{j'} \left[ \sum_{k=1}^{C_j-1} \sum_{l=1}^{C_{j'}-1} \Phi_2(v_{jk}, h_{j'_l}; \rho_{X_j^* X_{j'}^*}) - \sum_{k=1}^{C_j-1} \Phi_1(v_{jk}) \sum_{l=1}^{C_{j'}-1} \Phi_1(h_{j'_l}) \right]}$$

For one factor models, the formula above can be greatly simplified. The multilevel maximal reliability of weighted sum will apply the formula at both within-group and between-group levels.

## 5 Conclusions

This study proposed a confirmatory factor analysis approach to multilevel reliability for tests with ordered categories item responses. It extended single level reliabilities to multilevel reliabilities, and provided closed-form formulas for calculating various types of multilevel nonlinear reliabilities, including the composite reliability, the coefficient Omega, and the maximal reliability.

## References

- Bentler, P. M. (2007). Covariance structure models for maximal reliability of unit-weighted composites. In *Handbook of latent variable and related models* (pp. 1–19). North-Holland.
- Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, *74*(1), 137–143.
- Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297–334.
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Information Age.
- Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91.
- Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement*, *37*(4), 827–838.
- Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, *74*(1), 155–167.
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future – A festschrift in honor of Karl Jöreskog* (pp. 195–216). Scientific Software International.
- Kim, S., Lu, Z., & Cohen, A. S. (2020). Reliability for tests with items having different numbers of ordered categories. *Applied psychological measurement*, *44*(2), 137–149.
- Lord, F. M., & Novick, R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley.
- McDonald, R. P. (1985). *Factor analysis and related methods*. Erlbaum.
- Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling*, *2*(3), 255–273.
- Muthén, B. O. (2011). Mean and covariance structure analysis of hierarchical data. UCLA: Department of Statistics, UCLA. Retrieved from <https://escholarship.org/uc/item/1vp6w4sr>.
- Novick, M., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, *32*(1), 1–13.
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*(2), 173–184.
- Raykov, T., & Shrout, P. E. (2002). Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling*, *9*(2), 195–212.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107–120.