Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 18, Number 1, February 2013 ISSN 1531-7714

Unit-weighted scales imply models that should be tested!

André Beauducel & Anja Leue, *University of Bonn*

In several studies unit-weighted sum scales based on the unweighted sum of items are derived from the pattern of salient loadings in confirmatory factor analysis. The problem of this procedure is that the unit-weighted sum scales imply a model other than the initially tested confirmatory factor model. In consequence, it remains generally unknown how well the model implied by the unit-weighted sum scales fits the data. Nevertheless, the derived unit-weighted sum scales are often used in applied settings. The paper demonstrates how model parameters for the unit-weighted sum scales can be computed and tested by means of structural equation modeling. An empirical example based on a personality questionnaire and subsequent unit-weighted scale analyses are presented in order to demonstrate the procedure.

The investigation of model fit is common in structural equation modeling (SEM) and confirmatory factor analysis (CFA), as well as in the context of item response theory. Several different fit indexes and cut-off values for fit indexes have been proposed, evaluated, and discussed for SEM and CFA (Barrett, 2007; Beauducel & Wittmann, 2005; Hu & Bentler, 1999; Tanaka, 1993). In several areas of psychological assessment, scales are formed as unit-weighted aggregates (sums) of item responses. The unit-weighted sum of items usually represents the raw score of a test and is sometimes called unweighted score (Lord & Novick, 1968). Examples for the calculation of unit-weighted sum scales based on a conventional CFA model can be found in several different areas of psychological assessment (e.g., Baloğlu & Zelhart, 2007; Brown & Krishnakumar, 2007; Lee & Chokkanathan, 2008; Moraitou & Efklides, 2009; Norton, 2007; Prinzie, Onghena & Hellinckx, 2007; Van der Linden, d'Acremont, Zermatten, Jermann, Larøi, Willems, Juillerat & Bechara, 2006). Usually researchers are satisfied by performing some model investigation by means of CFA and compose unit-weighted sum scales based on the variables with salient CFA loadings. The models that are implied by the unit-weighted scales are usually not tested, although structural equation modeling could be used in order to test these models.

The confirmatory model that corresponds to a unit-weighted scale can be directly calculated when the scale comprises only one factor or component. When a unit-weighted scale is computed as a sum of items, the correlations between the items and the scale can be calculated subsequently. In a model with a single factor or component as well as in a model with orthogonal factors or components, the loadings are the correlations between the items and the factor or component (Gorsuch, 1983; Harman, 1976). The loadings representing the correlations between the items and the factors are usually called structure coefficients and in a CFA with correlated factors the interpretation is usually based on the pattern coefficients. However, in a model with only one factor or in an orthogonal factor model the structure coefficients and the pattern coefficients are identical (Gorsuch, 1983; Harman, 1976). Therefore, in a single-factor or in a single-component model the correlations of the variables with the scale can be regarded as the loadings of the factor or component. Accordingly, these correlations can be fixed as the loadings of the model and the model fit can be assessed in SEM. Thus, the item-scale correlations (i.e., the not part-whole corrected item-total correlations) are the model parameters (loadings) of the unit-weighted scale model. Formally, the SEM obtained by this procedure is a confirmatory component model (CCM), since the model can be generated from the unit-weighted scores of the observed variables. It could also be regarded as a special form of the regression component model (Schönemann & Steiger, 1978). If several intercorrelated unit-weighted scales are analyzed, the intercorrelations of the scales must be taken into account when the loadings are calculated for a corresponding CCM. It is possible to calculate the loadings of the unit-weighted scale model from the correlations between the items and the scales and from the intercorrelations of the scales. However, in order to assess the model fit of several unit-weighted scales, it would be sufficient to calculate a single component model for each scale. This would allow for a separate evaluation of the model fit of each scale.

Possibly, researchers are not aware of the fact that unit-weighted scales imply specific models that can be tested with SEM and that these models are usually different from those that are tested in an initial CFA. Although the strategy of performing CFA as a basis for unit-weighted sum scales is rather common, this state of affairs leaves an unknown gap of misfit between the originally tested CFA models and the models implied by the unit-weighted scales that are generally used in applied settings. Therefore, the present paper presents a method that allows for testing the fit of the models implied by unit-weighted scales. Based on this method, the fit obtained for conventional CFA models can be compared with the fit that occurs for the models implied by unit-weighted scales. If the model implied by the unit-weighted scales does not fit to the data, there are different ways to deal with the problem: Some items with insufficient loadings might be deleted or factor score predictors (Harman, 1976; Grice, 2001b; Beauducel & Rabe, 2009) might be computed.

It should be noted that we do not argue against the use of unit-weighted scales. This misunderstanding might occur, because we will demonstrate that in some cases, when the conventional CFA model fits the data, the model implied by the unit-weighted scale does not fit the data. However, if the models implied by unit-weighted scales fit well to the data, there is nothing wrong with the unit-weighted scales so that they might be used in applied settings if their reliability is sufficient. Therefore, the present paper does not contradict the results provided by Grice (2001a), who found that unit-weighted scales based on factor score coefficients compared favorably with factor score predictors.

Another possible misunderstanding might be that CFA models with equal salient loadings should be tested as a basis for computing unit-weighted scales. Testing for equal salient loadings might be of interest, but the CFA loadings are the weights of the common factors when the observed variables (items) are predicted from the common factors. In contrast, when unit-weighted scales are computed, the unit-weights are the weights for the prediction of the scales from the observed variables. Even when all items have a unit-weight as a predictor of the scale, there might be a variability of the item-total correlations. This variability should also show up in the loadings of the CCM corresponding to the scale. Therefore, a model with equal loadings does not correspond to the model implied by unit-weighted scales.

The evaluation of the CCM implied by unit-weighted scales comprises the following steps: First, a conventional CFA model is calculated. Second, unit-weighted scales are computed for the variables with salient loadings. Third, the correlations of the items with the corresponding unit-weighted scale are computed. Finally, the loadings are fixed to the item-total correlations obtained in step three and the model fit is calculated by means of SEM. In the following, these steps are demonstrated by means of an empirical example.

Methods

Sample and instrumentation

The investigation is based on a sample of 446 German participants (240 females; age in years: M= 34.57; SD=12.86) that were recruited through newspaper advertisement. The participants filled in a German paper-pencil short version of the Zuckerman-Kuhlman-Personality-Questionnaire (ZKPQ; Zuckerman, Kuhlman, & Camac, 1988; Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993). The short version of the ZKPQ measures the five personality dimensions proposed in the alternative five-factor model of Zuckerman et al. (1988) by means of 60 dichotomous items. The alternative five-factor model comprises

Impulsive Sensation Seeking (ISS), Anxiety (ANX), Aggression-Hostility (AGH), Sociability (SOC), and Activity (ACT). The short version of the ZKPQ comprises 13 items for ISS, 17 items for ANX, 9 items for AGH, 14 items for SOC, and 7 items for ACT. All participants gave written informed consent and participated voluntarily in the study.

Statistical analysis

CFA with maximum-likelihood estimation of the short version of the ZKPQ was calculated by means of LISREL 8.8 (Jöreskog & Sörbom, 2006), because maximum-likelihood estimation is the most common estimation method. For the CFA model, the salient loadings on each factor were freely estimated, the remaining loadings were fixed to zero and the factor variances were fixed to one. The inter-factor correlations were freely estimated. For each factor, a separate CCM representing the model of the corresponding unit-weighted scale was also computed. In the CCM all loadings were fixed to the value of the correlation between the item and the corresponding unit-weighted scale. The error terms of the items were fixed to one minus the squared correlation between the item and the unit-weighted scale. In order to specify the correlations between the items and the components as loadings, the CCM analyses were based on the intercorrelations of the items and the variances of the components were fixed to one. An example for a LISREL syntax that illustrates the calculation of the CCM for the ACT scale is given in the Appendix. Model fit was evaluated according to some conventional fit indices, which are the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR). The fit indices were evaluated according to conventional cut-off values (Hu & Bentler, 1999; Beauducel & Wittmann, 2005). Unit-weighted scales and the correlations between the scales and the items (item-total correlations) were computed by means of SPSS 20.

Results

The CFA results for the short version of the ZKPQ are presented in Table 1. The model fit was acceptable so that researchers could aim at composing a unit-weighted scale for each of the five factors. Cronbach's Alpha for the unit-weighted scales was .83 for the IMP scale, .87 for the ANX scale, .74 for the AGH scale, .87 for the SOC scale, and .77 for the ACT scale. It should be noted that the CCM loadings presented in Table 2 are the correlations between the items and the respective unit-weighted scale. The fit of the CCM corresponding to the unit-weighted scales was acceptable for the IMP and ANX scale, very good for the AGH scale, and insufficient for the SOC scale, because the CFI was smaller than .90 and the RMSEA was larger than .10. The fit of the ACT scale should also be regarded as insufficient, because the CFI was below .90 and the RMSEA as well as the SRMR were larger than .10.

Discussion

A method to compute the fit of the models implied by unit-weighted item sum scales was proposed. Unit-weighted sum scales imply component models in which the component loadings are the correlations between the unit-weighted scale and the corresponding items. These models can be evaluated by means of SEM. If the model implied by the unit-weighted scale fits the data, the unit-weighted scale can be regarded as acceptable. Otherwise, it could be conceived to exclude items with moderate or small salient loadings in order to calculate a shortened unit-weighted scale. Additionally, the decision on the variables to be included into the unit-weighted scales could be based on the salient factor score coefficients (Grice & Harris, 1998; Grice, 2001a) and not on the salient factor loadings. This might lead to modified unit-weighted scales whose fit should, however, also be evaluated. Another possibility would be to compute factor score predictors instead of unit-weighted scales (e.g., DiStefano, Zhu, & Mindrila, 2009), even when this leads to the problem to choose an optimal factor score predictor (Grice, 2001b; Krijnen, 2006). However, all these possibilities of scale improvement should encourage researchers to evaluate by means of the method proposed in this paper whether it can be justified to compose unit-weighted sum scales.

The method to calculate the model fit of the unit-weighted sum scales was illustrated by means of an empirical example based on a short form of a questionnaire for the alternative five-factor model of

Table 1: CFA model for the short version of the ZKPQ (completely standardized solution)

item	IMP	ANX	AGH	SOC	ACT	item	IMP	ANX	AGH	SOC	ACT
1	.46	-	-	-	-	31	-	-	.57	-	-
2	.36	-	-	-	-	32	-	-	.56	-	-
3	.60	-	-	-	-	33	-	-	.38	-	-
4	.42	-	-	-	-	34	-	-	.52	-	-
5	.63	-	-	-	-	35	-	-	.52	-	-
6	.62	-	-	-	-	36	-	-	.49	-	-
7	.54	-	-	-	-	37	-	-	.43	-	-
8	.42	-	-	-	-	38	-	-	.42	-	-
9	.50	-	-	-	-	39	-	-	.54	-	-
10	.64	-	-	-	-	40	-	-	-	.40	-
11	.51	-	-	-	-	41	-	-	-	.44	-
12	.58	-	-	-	-	42	-	-	-	.35	-
13	.50	-	-	-	-	43	-	-	-	.62	-
14	-	.50	-	-	-	44	-	-	-	.51	-
15	-	.45	-	-	-	45	-	-	-	.66	-
16	-	.58	-	-	-	46	-	-	-	.59	-
17	-	.47	-	-	-	47	-	-	-	.62	-
18	-	.70	-	-	-	48	-	-	-	.67	-
19	-	.51	-	-	-	49	-	-	-	.51	-
20	-	.51	-	-	-	50	-	-	-	.67	-
21	-	.41	-	-	-	51	-	-	-	.57	-
22	-	.51	-	-	-	52	-	-	-	.51	-
23	-	.54	-	-	-	53	-	-	-	.72	-
24	-	.42	-	-	-	54	-	-	-	-	.67
25	-	.68	-	-	-	55	-	-	-	-	.56
26	-	.47	-	-	-	56	-	-	-	-	.54
27	-	.49	-	-	-	57	-	-	-	-	.63
28	-	.51	-	-	-	58	-	-	-	-	.41
29	-	.77	-	-	-	59	-	-	-	-	.49
30	-	.45	-	-	-	60	-	-	-	-	.72

Inter-factor correlations:

	IMP	ANX	AGH	SOC
ANX	-.06			
AGH	.22	.27		
SOC	.43	-.18	.10	
ACT	.15	.00	-.04	.21

Model fit:

$\chi^2(1700) = 3419.24$; $p < .001$; CFI = .91;
RMSEA= .050; SRMR = .065

Note. ISS= Impulsive Sensation Seeking; ANX= Anxiety; AGH= Aggression-Hostility; SOC= Sociability; ACT= Activity.

personality (Zuckerman et al., 1993). Although the model fit of the overall CFA was acceptable, only three of the models corresponding to the unit-weighted scales had an acceptable or excellent fit (IMP, ANX, AGH), whereas the model fit of two other unit-weighted scale models was insufficient (SOC, ACT). Moreover , it was interesting that model fit of the SOC scale was insufficient although its Cronbach's Alpha coefficient was .87. Thus, the fit of the model corresponding to the unit-weighted scale and Cronbach's Alpha reveal different aspects of psychometric quality of a scale. The fit of the model implied by a unit-weighted scale is related both to reliability and validity. Since the reliability of a unit-

Table 2: CCM for the unit-weighted scales of the short version of the ZKPQ (completely standardized solution)

IMP		ANX		AGH		SOC		ACT	
item	loading	item	loading	item	loading	item	loading	item	loading
1	.52	14	.54	31	.63	40	.44	54	.73
2	.43	15	.52	32	.62	41	.53	55	.62
3	.64	16	.61	33	.47	42	.46	56	.63
4	.50	17	.52	34	.60	43	.66	57	.70
5	.67	18	.69	35	.58	44	.60	58	.54
6	.64	19	.55	36	.58	45	.64	59	.60
7	.59	20	.56	37	.56	46	.60	60	.74
8	.49	21	.47	38	.48	47	.64		
9	.56	22	.56	39	.62	48	.68		
10	.65	23	.58			49	.59		
11	.57	24	.51			50	.69		
12	.63	25	.69			51	.65		
13	.53	26	.53			52	.56		
		27	.52			53	.70		
		28	.54						
		29	.74						
		30	.51						
Model fit:									
$\chi^2(91)$ =163.68; p < .001; CFI = .97; RMSEA= .047; SRMR = .066		$\chi^2(153)$=376.79; p < .001; CFI = .96; RMSEA= .067; SRMR = .067		$\chi^2(45)$ =77.20; p = .002; CFI = .97; RMSEA= .047; SRMR = .082		$\chi^2(105)$=645.95; p < .001; CFI = .88; RMSEA= .113; SRMR = .088		$\chi^2(28)$ = 148.04; p < .001; CFI = .89; RMSEA= .119; SRMR = .105	

Note. In the CCM the loadings are the correlations between the items and the unit-weighted scale. ISS= Impulsive Sensation Seeking; ANX= Anxiety; AGH= Aggression-Hostility; SOC= Sociability; ACT= Activity.

weighted scale might be sufficient even when its validity is low, the fit of the model implied by the unit-weighted scale provides information beyond reliability. Accordingly, the present results indicate that it might be insufficient only to report an overall CFA and the Cronbach's Alpha for the unit-weighted scales composed from the items with salient loadings.

To sum up, the method proposed here allows for an evaluation of the model fit of the model implied by unit-weighted sum scales. Since most raw scores are unweighted and therefore unit-weighted sums of items, the proposed method applies to most raw scores. Moreover, the model implied by unit-weighted scales can easily be specified, because only the correlations between the unit-weighted scales have to be entered for the specification of the loadings and the error terms fixed according to the loadings (see Appendix). Researchers might therefore be encouraged to test the fit of the model implied by unit-weighted sum scales whenever they want to use those scales in order to provide a psychometric justification of their scales.

References

Baloğlu, M. & Zelhart, P.F. (2007). Psychometric properties of the Revised Mathematics Anxiety Rating Scale. *The Psychological Record, 57,* 593-611.

Barrett, P. (2007). Structural equation modeling: Adjudging model fit. *Personality and Individual Differences, 42,* 815-824.

Beauducel, A. & Rabe, S. (2009). Model-related factor score predictors for confirmatory factor analysis. *British Journal of Mathematical and Statistical Psychology, 62,* 489-506.

Beauducel, A. & Wittmann, W.W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling, 12,* 41-75.

Brown, T.L. & Krishnakumar, A. (2007). Development and validation of the Adolescent Racial and Ethnic Socialization Scale (ARESS) in African American Families. *Journal of Youth Adolescence, 36*, 1072-1085.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations fort he applied researcher. *Practical Assessment, Research, and Evaluation, 14*. Available online: http://pareonline.net/getvn.asp?v=14&n=20.

Gorsuch, R.L. (1983). *Factor Analysis*. Hillsdale, NJ: Lawrence Erlbaum.

Grice, J.W. (2001a). A comparison of factor scores under conditions of factor obliquity. *Psychological Methods, 6,* 67-83.

Grice, J.W. (2001b). Computing and evaluation factor scores. *Psychological Methods, 6,* 430-450.

Grice, J.W. & Harris, R.J. (1998). A comparison of regression and loading weights for the computation of factor scores. *Multivariate Behavioral Research, 33,* 221-247.

Harman, H.H. (1976). *Modern factor analysis* (3rd ed.). Chicago: The University of Chicago Press.

Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Krijnen, W.P. (2006). Some results on mean square error for factor score prediction. *Psychometrika, 71,* 395-409.

Lord, F.M. & Novick, M.R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Jöreskog, K. G. & Sörbom, D. (2006). *LISREL 8.8*. Chicago: Scientific Software International.

Lee, A.E.Y. & Chokkanathan, S. (2008). Factor structure of the 10-item CES-D scale among community dwelling older adults in Singapore. *International Journal of Deriatric Psychiatry, 23,* 592-597.

Moraitou, D. & Efklides, A. (2009). The Blank in the Mind Questionnaire (BIMQ). *European Journal of Psychological Assessment, 25,* 115-122.

Norton, P.J. (2007). Depression Anxiety and Stress Scales (DASS-21): Psychometric analysis across racial groups. *Anxiety, Stress, & Coping, 20,* 253-265.

Prinzie, P., Onghena P., & Hellinckx, W. (2007). Reexamining the Parenting Scale. *European Journal of Psychological Assessment, 23,* 24-31.

Schönemann, P. H. & Steiger, J. H. (1976). Regression component analysis. *British Journal of Mathematical and Statistical Psychology, 29,* 175-189.

SPSS (2011). *SPSS for Windows, Release 20.0.0.* Chicago: Author.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10–39). Newbury Park, CA: Sage.

Van der Linden, M., d'Acremont, M., Zermatten, A., Jermann, F., Larøi, F., Willems, S., Juillerat, A.-C., & Bechara, A. (2006). A french adaptation of the UPPS Impulsive Behavior Scale. *European Journal of Psychological Assessment, 22,* 38-42.

Zuckerman, M., Kuhlman, D.M. & Camac, C. (1988). What lies beyond E and N? Factor analyses of scales believed to measure basic dimensions of personality. *Journal of Personality and Social Psychology*, *54*, 96-107.

Zuckerman, M., Kuhlman, D.M., Joireman, J., Teta, P. & Kraft. (1993). A comparison of three structural models for personality: The big three, the big five, and the alternative five. *Journal of Personality and Social Psychology*, *65*, 757-768.

Appendix

```
TI 'Confirmatory component model (CCM) for the ACT scale'
DA NI=7 NO=446 MA=KM
RA FI='C:\DATA.psf'
MO NX=7 NK=1 TD=DI,FI LX=FI PH=FI
LK
SCALE

VA 0.730 LX 1 1
VA 0.622 LX 2 1
```

```
VA 0.626 LX 3 1
VA 0.697 LX 4 1
VA 0.539 LX 5 1
VA 0.599 LX 6 1
VA 0.743 LX 7 1


CO TD 1 1 = 1 - LX 1 1**2
CO TD 2 2 = 1 - LX 2 1**2
CO TD 3 3 = 1 - LX 3 1**2
CO TD 4 4 = 1 - LX 4 1**2
CO TD 5 5 = 1 - LX 5 1**2
CO TD 6 6 = 1 - LX 6 1**2
CO TD 7 7 = 1 - LX 7 1**2

VA 1 PH 1 1

OU it=10 SC
```

Citation:

Beauducel, André & Leue, Anja (2013). Unit-weighted scales imply models that should be tested! *Practical Assessment, Research & Evaluation*, 18(1). Available online: http://pareonline.net/getvn.asp?v=18&n=1

Authors:

André Beauducel,
University of Bonn, Faculty of Arts,
Institute of Psychology,
Kaiser-Karl-Ring 9, 53111 Bonn,
Germany
email: beauducel [at] uni-bonn.de

Anja Leue,
University of Bonn, Clinic of Epileptology
and Institute of Psychology,
Kaiser-Karl-Ring 9, 53111 Bonn,
Germany,
email: anja.leue [at] uni-bonn.de