International Journal of Assessment Tools in Education
2022, Vol. 9, No. 3, 583–592
https://doi.org/10.21449/ijate.946231
Published at https://ijate.net/ https://dergipark.org.tr/en/pub/ijate *Research Article*

Which scale short form development method is better? A Comparison of ACO, TS, and SCOFA

Hakan Kogar [1,*]

[1]Akdeniz University, Faculty of Education, Department of Educational Sciences, Antalya

ARTICLE HISTORY

Received: June 01, 2021
Revised: Mar. 13, 2022
Accepted: July 12, 2022

Keywords:
Short form,
Scale development,
Ant colony optimization,
Tabu search,
Stepwise confirmatory factor analysis.

Abstract: The purpose of this study is to identify which scale short-form development method produces better findings in different factor structures. A simulation study was designed based on this purpose. Three different factor structures and three simulation conditions were selected. As the findings of this simulation study, the model-data fit and reliability coefficients were reported for each factor structure in each simulation condition. All analyses were conducted under the R environment. According to the findings of this study, the increase in the level of misspecification and the decrease in the sample size can significantly affect the model-data fit. In a situation where the factor structure of the scale is getting more and more complex, model-data fit and Omega coefficients decrease. For scales with a unidimensional factor structure, all of the scale short-form development methods are recommended. For scales with multidimensional factor structure, Ant Colony Optimization, and Stepwise Confirmatory Factor Analysis algorithms and for scales with bifactor factor structure, the ACO algorithm is recommended. When viewed from the framework of metaheuristic algorithms, it has been identified that ACO produces better findings than Tabu Search.

1. INTRODUCTION

The use of short forms of psychological measurement tools has become widespread, especially in the last 20 years. Is it more important than the scale has a well-prepared factor structure but contains many items, or is it more important to obtain sufficient proof of validity by using time correctly with fewer items? This question has accelerated the work on short-form development. The main reason for this situation is to reduce the time and cost required for the application of the test, and the effort and length of the test that the participant would spend on the test items are appropriate (Kleka & Soroko, 2018). Due to these important reasons, academic studies to shorten the long forms of the scales have started to gain an important place in the social science literature. It should be noted that this situation has theoretical reasons as well as practical reasons. According to the Classical Test Theory (CTT), the high number of items makes significant contributions to the reliability of test scores, construct, and content validity. Many items are needed based on CTT to make valid and reliable measurements (Anastasi, 1982; Nunnally, 1978).

*CONTACT: Hakan Koğar hkogar@gmail.com Akdeniz University, Faculty of Education, Department of Educational Sciences, Antalya, Türkiye

e-ISSN: 2148-7456 /© **IJATE** 2022

Nowadays, many new methods have been developed for scale short-form development (Leite et al., 2008; Olaru et al., 2015). It is possible to talk about dozens of methods within the framework of Metaheuristic Algorithms, Factor Analysis, Item Response Theory, and Rasch analysis. However, it is still known in the literature that classical approaches are frequently used in short-form development. These approaches are mostly based on item statistics or factor analysis approach. Item-total correlation and factor loadings or removing items with a low contribution to internal consistency constitute the basic implementation form of these approaches. This means determining the short form according to the item characteristics (Janssen et al., 2017). However, developing a short form of a scale is a much more complex and comprehensive process.

Classical short-form development approaches potentially have a significant disadvantage when it is desired to reduce the number of items of a previously validated measurement instrument. In these techniques, psychometrically poor items can be detected through item reliability or item-total correlation. However, in this case, depending on the test item excluded, the statistical findings of the remaining items and the findings of the general test will vary. Therefore, a stepwise item selection for the development of a short form will result in different item groups depending on the order of items eliminated (Janssen et al., 2017).

One of the methods developed to overcome many of the disadvantages of these classical approaches is Stepwise Confirmatory Factor Analysis (SCOFA). With this method, latent variable models are used, which provide a comprehensive framework for testing measurement models. This model overcomes one of the disadvantages of classical approaches by focusing directly on dimensionality and factor structure. Reducing the item pool in a measurement process can affect the factor structure of the instrument (Schroeders et al., 2016).

Obtaining a short form with both a strong factor structure and a high validity is quite difficult with traditional short-form development methods. Metaheuristic optimization algorithms have the potential to solve these difficulties because they can optimize and test with multiple validity criteria simultaneously for the developed short forms. Ant Colony Optimization (ACO), the first of the metaheuristic algorithms discussed within the scope of this research, was firstly developed by Colorni et al. (1991) as a metaheuristic to solve a wide range of combination-based problems and can be applied to situations where there are many possible solutions for a problem to be graded based on various criteria. This approach does not require the best solution to existing scale but instead focuses on finding a solution within the set of possible solutions that best meet certain criteria. One of the potential problems that this approach can solve is the development of the short form of the scales, undoubtedly. In this context, any combination of selected items is a possible solution, and these possible solutions will vary according to the previously established degree of competence (Janssen et al., 2017).

Tabu Search (TS), another metaheuristic algorithm, was designed by Marcoulides & Falk (2018) for short-form development.

> “TS examines each set of short forms created by changing one item at a time. The main idea behind the TS method is to consistently identify the best short-form currently selected by examining other short forms neighbor the best available short form. If a neighbor examined short form fits better than the existing short form, it is selected as the most suitable new short form. If not, the neighbor short form under study is marked as "taboo." In other words, it is placed in a separate list so that it will not be re-evaluated until certain criteria are met." (Raborn et al., 2020, p. 5).

Within the scope of this research, ACO and TS algorithms were selected from metaheuristic algorithms with the simulation study conducted by Raborn et al. (2020), taking into account other application-oriented studies and prevalence. In addition, due to the frequent use of classical approaches in the literature, the SCOFA technique, which is thought to represent these approaches and is based on iterations such as metaheuristic algorithms, was chosen. A simulation study was designed based on the technique of developing the short forms of these three scales. The research questions of this study are as follows:

1. How are the model-data fit and reliability coefficients obtained from different scale short-form development methods according to different factor structures?
2. How are the model-data fits and reliability coefficients obtained from different scale short-form development methods according to different sample sizes, the correlation between factors, and model misspecification?
3. Which scale short-form development method performs best under various conditions?

2. METHOD

2.1. Simulation Conditions

In this study, various factor structures were primarily defined. Unidimensional, multidimensional, and bifactor structures were selected for this study. For each factor structure, research findings from real data sets were used. Simulation-based data sets were produced on various other features, especially the factor loadings of these measurement tools. Instructor Self-Disclosure Scale with 18 items for unidimensional structure (Cayanus & Martin, 2004), the Multidimensional Health Locus of Control Scale (LaNoue et al., 2015) with 3 factors and 18 items for the multidimensional structure, and The Anxiety Sensitivity Index-3 (Ebesutani et al., 2014) with 3 factors and 18 items for the bifactor structure were selected. Factor loading values of the scale vary between 0.28 - 0.70. The correlation coefficients between the factors vary between 0.41 - 0.45. Especially attention has been paid to ensure that each of these studies has an equal number of items. It is aimed to reduce the number of items in each factor structure by half. In this way, it is planned to develop the short form of the Instructor Self-Disclosure Scale with 9 items and a unidimensional scale, and the other scales with a total of 9 items, 3 items in each factor. Confirmatory Factor Analysis (CFA) was repeated on 9 items determined by each scale short form method.

Three different conditions have been manipulated for the purpose of the simulation study based on Raborn et al., (2020). These features are sample size, model misspecification, and correlation between factors.

- *Sample Size:* In CFA, it is stated that a sample size of at least 200 is required for accurate model estimates (Gatignon, 2010; Singh et al., 2016). Considering other similar simulation studies (French & Finch, 2011; Yang & Liang, 2013), two different sample sizes were determined as 200 and 500. Only findings based on sample size are included in the unidimensional structure.

- *Correlation between factors:* This condition were identified to be 0.00, 0.25 and 0.50 in a study by Batley & Boss (1993), 0.20, 0.50 and 0.70 in a study by Jiang et al. (2016), 0.10, 0.40 and 0.70 in a study by Van Abswoude et al. (2004a) and 0.00, 0.20, 0.40, 0.60, 0.80 and 1.00 in a study by Van Abswoude et al. (2004b). In this study, based on the correlations determined by these studies, two different correlation values were selected: 0.30 and 0.70 only for multidimensional factor structure. There is no correlation between factors in a unidimensional structure. In addition, in bifactor models, correlations between dimensions were not included as a simulation condition because "The bifactor model incorporates a general factor, onto which all items load directly, plus a series of orthogonal (i.e., specified as uncorrelated) factors each loading on a sub-set of items." (Reise, 2012, p. 682).

- *Model Misspecification:* Model Misspecification was applied by ensuring that some of the observed variables were included in the factor that was not loaded. In the multidimensional and bifactor models, 6 items were selected, and it was ensured that these items were loaded in the factors where these items were not loaded in pairs. Three different models of misspecification have been selected: No Misspecification (0.00), 0.30, and 0.60 (Raborn et al., 2020). The misspecified loadings (loading on the incorrect factor) were not the same as the loadings simulated to be real datasets. Misspecification was not applied to the model in unidimensional structure.

2.2. Data Simulation and Analysis

All data sets were simulated in R v4.0.4 (R Core Team, 2018) using the simulateData function on lavaan 0.6-8 package (Rosseel, 2012). All factor structures fitted on lavaan 0.6-8 package (Rosseel, 2012). First, population models were created with the findings of real data in the production of data belonging to unidimensional, multidimensional, and bifactor structures. These models are then calibrated to the null model. This process is repeated for each simulation condition. 100 iterations have been used per each simulation condition.

The scale short-form selection with ACO and TS was implemented with the ShortForm 0.4.6 package (Raborn & Leite, 2018). This package uses the lavaan package (Rosseel et al., 2012) to fit unidimensional, multidimensional, and bifactor CFA analysis. Based on previous research (Marcoulides & Falk, 2018; Raborn et al., 2020), some tuning parameters were used for each meta-heuristic algorithm. For ACO, 20 consecutive steps for convergence, 0.9 evaporation, 20 ants, and 50 maximum steps for no improvement were tuned. For TS, 5 tabu sizes for each condition and 50 iterations were specified. Since the iterations made with the ShortForm package were very slow, the number of iterations was limited (Raborn et al., 2020).

SCOFA analysis was implemented with lavaan 0.6-8 package (Rosseel, 2012). SCOFA algorithm, which iteratively deletes the item with the lowest factor loading from the item pool, is a standard scale short-form development procedure (Kruyen et al., 2013). "After estimating a CFA for the original factor structure, the item with the lowest factor loading is removed. The model is then re-estimated with the reduced item set and again the item with the lowest factor loading is removed. This procedure is repeated until the predetermined number of items for the short version is reached." (Schroeders et al., 2016, p. 8). Weighted Least Squares Mean and Variance adjusted (WLSMV) estimator are used for parameter estimation.

Model-fit was checked using several fit indices, including the Comparative Fit Index (CFI), the Tucker–Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA). CFI and TLI values above 0.90 and 0.95 reflect acceptable and excellent fit, respectively, while RMSEA below or near 0.05 indicates an acceptable fit of data to a model (Hu & Bentler 1999, pp. 24-26).

Omega coefficients were computed from semTools package 0.5-4 (Jorgensen et al., 2014). For all factor structures, omega coefficients as composite reliability is computed. Omega hierarchical (ωH) and omega hierarchical subscale (ωHS) are computed for bifactor structures. Omega hierarchical subscales are computed for multidimensional and omega total coefficients are computed for unidimensional structures.

As the findings of this simulation study, model-data fit and reliability coefficients were reported for each factor structure in each simulation condition.

3. RESULT

3.1. Findings from Unidimensional Factor Structure

In the unidimensional factor structure, findings were reported only according to the changes in the sample size according to the simulation conditions.

Table 1. *Model-data fits and omega coefficients from the unidimensional factor structure.*

SS	ACO				TS				SCOFA			
	CFI	TLI	RMSEA	RF-g	CFI	TLI	RMSEA	RF-g	CFI	TLI	RMSEA	RF-g
200	.980	.985	.034	.830	.985	.989	.024	.778	.980	.985	.034	.830
500	.998	.998	.011	.812	.972	.979	.036	.794	.999	.999	.009	.812

SS: Sample Size, RF-g: Reliability of general factor

In Table 1, it has been observed that all model-data fits obtained according to the unidimensional factor structure indicate a good fit. When the sample size is 200 TS algorithm, it is seen that the SCOFA method produces the best results when it is 500. When the sample size is 200, the model-data fit and Omega coefficients obtained from ACO and SCOFA techniques are the same. In case the sample size is 500, the model-data fit and Omega coefficients obtained from ACO and SCOFA techniques are very close to each other. The Omega coefficients produced by ACO and SCOFA methods for both sample sizes are the same and higher than the coefficient produced by the short form obtained with TS. As the sample size increases, the model data fit generally increases, while the Omega coefficients decreases except for TS.

3.2. Findings from Multidimensional Factor Structure

In the multidimensional factor structure, findings were reported for all simulation conditions.

Table 2. *Model-data fits from the multidimensional factor structure.*

MS	SS	CBF	ACO			TS			SCOFA		
			CFI	TLI	RMSEA	CFI	TLI	RMSEA	CFI	TLI	RMSEA
None (0.0)	200	0.3	.998	.998	.006	.994	.996	.010	.991	.994	.013
		0.7	.994	.996	.012	.994	.996	.012	.974	.983	.027
	500	0.3	.984	.989	.017	.983	.989	.015	.972	.981	.022
		0.7	.990	.993	.015	.984	.989	.017	.994	.996	.012
Minor (0.3)	200	0.3	.995	.997	.009	.998	.998	.007	.924	.950	.045
		0.7	.915	.943	.053	.922	.948	.049	.856	.904	.078
	500	0.3	.913	.942	.040	.895	.930	.042	.902	.935	.049
		0.7	.979	.986	.026	.963	.975	.032	.984	.990	.024
Major (0.6)	200	0.3	.936	.957	.054	.892	.928	.057	.934	.956	.054
		0.7	.943	.962	.060	.789	.859	.079	.949	.966	.063
	500	0.3	.937	.958	.058	.890	.926	.051	.937	.958	.058
		0.7	.974	.982	.037	.952	.968	.039	.990	.994	.026

MS: Misspecification, SS: Sample Size, CBF: Correlation Between Factors

According to the findings obtained from the multidimensional factor structure, in Table 2, all model-data fit values of the situation where there was no misspecification showed a good fit. When the sample size was 200 and the correlation between factors was 0.3, the model-data fit values were the highest. Although all three scale short-form development methods produced similar findings, it can be said that the findings of ACO and TS were better. In the case of minor misspecification, when the sample size was 200 and the correlation between factors was 0.3, the sample size was 500 and the correlation between factors was 0.7, with high model-data fits. Some minor and major misspecification conditions findings were obtained with TS and SCOFA, it is observed that sufficient model-data fit was not achieved. In case the major misspecification, when only the sample size was 500 and the correlation between factors was 0.7, all scale short-form development methods showed sufficient model-data fit. Findings obtained from ACO and SCOFA showed sufficient model-data fit and were similar, specifically in no misspecification conditions; it is seen that the TS algorithm can generally obtain values far from adequate model-data fit.

Table 3. *Omega coefficients from the multidimensional factor structure.*

MS	SS	CBF	ACO			TS			SCOFA		
			RF-1	RF-2	RF-3	RF-1	RF-2	RF-3	RF-1	RF-2	RF-3
None (0.0)	200	0.3	.583	.606	.664	.437	.575	.585	.586	.595	.635
		0.7	.679	.577	.698	.575	.431	.421	.681	.735	.696
	500	0.3	.639	.648	.665	.552	.492	.498	.737	.758	.458
		0.7	.662	.577	.604	.576	.507	.501	.665	.716	.697
Minor (0.3)	200	0.3	.483	.563	.466	.536	.561	.410	.536	.563	.529
		0.7	.487	.490	.563	.487	.490	.563	.526	.456	.563
	500	0.3	.485	.541	.545	.464	.542	.422	.485	.507	.544
		0.7	.462	.561	.511	.389	.457	.487	.473	.562	.512
Major (0.6)	200	0.3	.426	.529	.392	.550	.440	.600	.559	.528	.615
		0.7	.399	.653	.336	.425	.532	589	.449	.643	.494
	500	0.3	.508	.438	.525	.511	.409	.519	.504	.592	.525
		0.7	.475	.612	.527	.472	.476	.548	.474	.640	.546

MS: Misspecification, SS: Sample Size, CBF: Correlation Between Factors, RF-1: Reliability of First Factor, RF-2: Reliability of Second Factor, RF-3: Reliability of Third Factor

According to the findings obtained from the multidimensional factor structure, in Table 3, there was a decrease in the Omega coefficient of none to major misspecification, generally. The better Omega coefficient values were the case where the sample size was 500 and the correlation between factors was 0.7 in the no misspecification compared with other short-form techniques. It can be said that the Omega coefficients of ACO and SCOFA are similar. Especially in the no misspecification, the Omega coefficients obtained by the TS algorithm are lower than the other algorithms.

3.3. Findings from Bifactor Factor Structure

In the bifactor factor structure, findings were reported according to the changes in the sample size and the correlation between factors according to the simulation conditions.

Table 4. *Model-data fits from the bifactor factor structure.*

MS	SS	ACO			TS			SCOFA		
		CFI	TLI	RMSEA	CFI	TLI	RMSEA	CFI	TLI	RMSEA
None (0.0)	200	.983	.992	.028	.981	.990	.027	.981	.990	.030
	500	.994	.997	.016	.991	.996	.020	.976	.988	.033
Minor (0.3)	200	.991	.996	.021	.969	.987	.039	.969	.987	.039
	500	.994	.998	.018	.897	.945	.068	.828	.908	.093
Major (0.6)	200	.881	.941	.072	.869	.934	.077	.892	.946	.076
	500	.914	.957	.068	.918	.959	.057	.955	.977	.045

MS: Misspecification, SS: Sample Size

In Table 4, according to the findings obtained under the bifactor structure, all model-data fits obtained in the no misspecification showed a good fit. However, in the case of minor misspecification, the model-data fit in the case where the sample size obtained from TS and SCOFA algorithms is 500 was not sufficient. Findings obtained with only ACO showed a good fit. In case of major misspecification and the sample size of 200, no scale short-form development method could show enough model-data fit. In the case of the sample size of 500, although the whole scale short-form development method showed sufficient model-data fit, the better findings were obtained with SCOFA compared with other short-form techniques.

Table 5. *Omega coefficients from the bifactor factor structure.*

MS	SS	ACO				TS				SCOFA			
		RF-1	RF-2	RF-3	RF-g	RF-1	RF-2	RF-3	RF-g	RF-1	RF-2	RF-3	RF-g
None (0.0)	200	.528	.160	.368	.763	.426	.483	.410	.844	.502	.176	.383	.760
	500	.327	.152	.186	.772	.396	.155	.176	.774	.428	.104	.229	.770
Minor (0.3)	200	.444	.507	.437	.403	.485	.284	.378	.713	.424	.456	.537	.344
	500	.365	.447	.543	.413	.684	.202	.518	.675	.336	.419	.308	.560
Major (0.6)	200	.213	.219	.351	.467	.211	.465	.464	.585	.183	.237	.217	.506
	500	.019	.170	.215	.417	.214	.384	.153	.534	.285	.174	.273	.490

MS: Misspecification, SS: Sample Size, RF-1: Reliability of First Factor, RF-2: Reliability of Second Factor, RF-3: Reliability of Third Factor, RF-g: Reliability of general factor

In Table 5, according to the findings obtained from the bifactor factor structure, there is a decrease in the Omega coefficient of none to major misspecification. The best Omega coefficients were obtained when there was no misspecification and the sample size was 200. Although the Omega values obtained from the scale short-form development methods are similar, it can be said that the Omega values obtained with the TS algorithm are higher.

4. DISCUSSION and CONCLUSION

It has been determined that all the methods of developing the short form of the scale in the unidimensional factor structure can select the short form with sufficient psychometric properties. Although it is not possible to mention a significant difference between the methods, the Omega coefficients produced by ACO and SCOFA methods are the same and higher than the coefficient produced by the short form obtained with TS. As the sample size increased, the model-data fit generally increased, while a slight decrease was observed in the Omega coefficients. In this case, for scales with unidimensional factor structure, all of the scale short-form development methods used in this study can be recommended.

In the multidimensional factor structure, all model-data fit values for the situation where there is no misspecification shows a good fit. In the case of minor misspecification, it was determined that some findings obtained by TS and SCOFA did not provide sufficient model-data fit. In the case of major misspecification, the findings obtained from ACO and SCOFA show sufficient model-data fit and are similar; it can be said that the TS algorithm can generally obtain values far from adequate model-data fit. It has been determined that the Omega coefficients of ACO and SCOFA are similar. Especially in the absence of misspecification, the Omega coefficients obtained by the TS algorithm are lower than the other algorithms. It is possible to say that the Omega coefficients increased with the increase in the sample size. In this case, for scales with multidimensional factor structure, ACO and SCOFA, which are among the scale short form development methods used in this study, can be recommended.

According to the findings obtained under the bifactor structure, all model-data fits obtained under no misspecification condition show a good fit. In the case of minor misspecification, TS and SCOFA algorithms could not show sufficient model-data fit, but in the case of major misspecification, no scale short form development technique could show sufficient model-data fit. In this case, it is recommended to use ACO, one of the short form development methods used in this study, for scales with a bifactor factor structure.

When viewed from the framework of metaheuristic algorithms, it has been determined that ACO produces better findings than TS. This finding is similar to the study of Raborn et al. (2020). In Raborn et al.'s (2020) study, the simulated annealing (SA) technique showed better performance in terms of fit indices and reliability indices. Next comes the ACO. Since the SA technique was not included in this study, it is possible to say that the findings of both studies are similar.

According to the findings of this study, the increase in the level of misspecification and the decrease in the sample size can significantly affect the model-data fit. In a situation where the factor structure of the scale is getting more and more complex, model-data fit and Omega coefficients decrease. Especially in cases where the factor structure is complex and the sample size is relatively low, it may be recommended to apply multiple scales short-form development methods and to continue studies on the methods that produce the best results.

This study was not carried out on two samples as suggested in the scale short form development studies. Using the required first sample as a "training sample" and choosing the item for the short form with this sample; the second sample should be used as a "testing sample" and the validity of the short form should be ensured with this sample. With such an application, it is ensured that the new short-form is validated in a new sample (Raborn et al., 2020). It is recommended to use these techniques and similar ones.

Declaration of Conflicting Interests and Ethics

The author declares no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the author.

Orcid

Hakan Koğar https://orcid.org/ 0000-0001-5749-9824

REFERENCES

Anastasi, A. (1982). *Psychological Testing* (5th ed.). Macmillan.

Batley, R.M., & Boss, M.W. (1993). The effects on parameter estimation of correlated dimensions and a distribution-restricted trait in a multidimensional item response model. *Applied Psychological Measurement*, *17*(2), 131-141. https://doi.org/10.1177/014662169301700203

Cayanus, J.L., & Martin, M.M. (2004). An instructor self-disclosure scale. *Communication Research Reports*, *21*(3), 252-263. https://doi.org/10.1080/08824090409359987

Colorni, A., Dorigo, M., & Maniezzo, V. (1991). *Distributed optimization by ant colonies*. In: Varela, F. and Bourgine, P., Eds., Proceedings of the European Conference on Artificial Life, ECAL'91, Paris, Elsevier Publishing, Amsterdam, 134-142.

Ebesutani, C., McLeish, A.C., Luberto, C.M., Young, J., & Maack, D.J. (2014). A bifactor model of anxiety sensitivity: Analysis of the Anxiety Sensitivity Index-3. *Journal of Psychopathology and Behavioral Assessment*, *36*(3), 452-464. https://doi.org/10.1007/s10862-013-9400-3

French, B.F., & Finch, W.H. (2011). Model misspecification and invariance testing using confirmatory factor analytic procedures. *The Journal of Experimental Education*, *79*(4), 404-428. https://doi.org/10.1080/00220973.2010.517811

Gatignon, H. (2010). *Confirmatory Factor Analysis. In Statistical Analysis of Management Data* (pp. 59-122). Springer. https://doi.org/10.1007/978-1-4419-1270-1_4

Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Janssen, A.B., Schultze, M., & Grötsch, A. (2017). Following the ants: Development of short scales for proactive personality and supervisor support by Ant Colony Optimization. *European Journal of Psychological Assessment*, *33*(6), 409. https://doi.org/10.1027/1015-5759/a000299

Jiang, S., Wang, C., & Weiss, D.J. (2016). Sample size requirements for estimation of item parameters in the multidimensional graded response model. *Frontiers in Psychology*, *7*(Article:109), 1-10. https://doi.org/10.3389/fpsyg.2016.00109

Jorgensen, T.D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., ..., & Enders, C. (2016). *semTools: Useful Tools for Structural Equation Modeling. R package version 0.5-4*. Retrieved from https://cran.r-project.org/web/packages/semTools/index.html

Kleka, P., & Soroko, E. (2018). How to avoid the sins of questionnaires abridgement?. *Survey Research Methods, 12*(2), 147-160. https://doi.org/10.31234/osf.io/8jg9u

Kruyen, P.M., Emons, W.H., & Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing, 13*(3), 223-248. https://doi.org/10.1080/15305058.2012.703734

LaNoue, M., Harvey, A., Mautner, D., Ku, B., & Scott, K. (2015). Confirmatory factor analysis and invariance testing between Blacks and Whites of the Multidimensional Health Locus of Control scale. *Health Psychology Open, 2*(2), 1-16. https://doi.org/10.1177/2055102915615045

Leite, W.L., Huang, I.-C., & Marcoulides, G. A. (2008). Item selection for the development of short forms of scales using an Ant Colony Optimization Algorithm. *Multivariate Behavioral Research, 43*, 411–431. https://doi.org/10.1080/00273170802285743

Marcoulides, K.M., & Falk, C. (2018). Model specification searches in structural equation modeling with R. *Structural Equation Modeling, 25*(3), 484-491. https://doi.org/10.1080/10705511.2017.1409074

Nunnally, J.C. (1978). *Psychometric Theory* (2nd ed.). McGraw-Hill.

Olaru, G., Witthöft, M., & Wilhelm, O. (2015). Methods matter: Testing competing models for designing short-scale big-five assessments. *Journal of Research in Personality, 59*, 56-68. https://doi.org/10.1016/j.jrp.2015.09.001

Raborn, A.W., & Leite, W.L. (2018). ShortForm: An R package to select scale short forms with the ant colony optimization algorithm. *Applied psychological measurement, 42*(6), 516. https://doi.org/10.1177/0146621617752993

Raborn, A.W., Leite, W.L., & Marcoulides, K.M. (2020). A comparison of metaheuristic optimization algorithms for scale short-form development. *Educational and Psychological Measurement, 80*(5), 910-931. https://doi.org/10.1177/0013164420906600

Reise, S.P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behav. Res. 47*, 667–696. https://doi.org/10.1080/00273171.2012.715555

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA). *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Schroeders, U., Wilhelm, O., & Olaru, G. (2016). Meta-heuristics in short scale construction: Ant colony optimization and genetic algorithm. *PLoS One, 11*(11), 1-19. https://doi.org/10.1371/journal.pone.0167110

Singh, K., Junnarkar, M., & Kaur, J. (2016). M*easures of Positive Psychology: Development and Validation*. Springer.

Van Abswoude, A.A., van der Ark, L.A., & Sijtsma, K. (2004b). A comparative study of test data dimensionality assessment procedures under nonparametric IRT models. *Applied Psychological Measurement, 28*(1), 3-24. https://doi.org/10.1177/0146621603259277

Van Abswoude, A.A., Vermunt, J.K., Hemker, B.T., & van der Ark, L.A. (2004a). Mokken scale analysis using hierarchical clustering procedures. *Applied Psychological Measurement, 28*(5), 332-354. https://doi.org/10.1177/0146621604265510

Yang, Y., & Liang, X. (2013). Confirmatory factor analysis under violations of distributional and structural assumptions. *International Journal of Quantitative Research in Education, 1*(1), 61-84. https://doi.org/10.1504/ijqre.2013.055642