Accuracy of Data-Model Fit Using Growing Levels of Invariance Models

Deyab A. Almaleki
damaleki@uqu.edu.sa
Department of Evaluation, Measurement and Research, Umm Al-Qura University, Makkah, Saudi Arabia

Abstract
The aim of this study is to provide empirical evaluation of the accuracy of data-model fit using growing levels of invariance models. Overall model accuracy of factor solutions was evaluated by the examination of the order for testing three levels of measurement invariance (MIV) starting with configural invariance (model 0). Model testing was evaluated by the Chi-square difference test ($\Delta\chi^2$) between two groups, and root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI) were used to evaluate the all-model fits. Factorial invariance result revealed that stability of the models was varying over increasing levels of measurement as a function of variable-to-factor ratio (VTF), subject-to-variable ratio (STV), and their interactions. There were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved. For VTF ratio (3:1, 6:1, and 9:1), the models started to show accuracy over levels of measurement when STV ratio was 6:1. Yet, the frequency of stability models over 1000 replications increased (from 69% to 89%) as STV ratio increased. The models showed more accuracy at or above 39:1 STV.

Key words:
Model- accuracy; Factorial-invariance; Level of measurement invariance; Factorial invariance

1. Introduction

To understand the impact of experimental design (ED) and the sampling design (SD)or other influences on WSV, a systematic structure for evaluating WSV changes is necessary. One possibility to evaluate WSV systematically is to use factorial invariance (FIV) [1]–[5]. Methods of FIV offer a structure that allows for disentangling measurement elements from structural elements in the factor model. Via FIV and evaluation of data-model fit, the impact of ED and SD on WSV can be compared among groups by examination of model precision [1], [6]–[9]. Previous research has investigated the precisions of factor solutions by the examination of Chi-square value (χ^2) and overall model fit indices (OMF) such as goodness-of- fit index (GFI), adjusted goodness-of-fit index (AGFI), Tucker-Lewis index (TLI), comparative fit index (CFI), and root mean square error of approximation (RMSEA) [10]–[14]. Overall model fit indices examined global measures of data-model fit.

There are three key features of design that are of paramount importance and generally overshadow all the technical decisions facing the researcher. These three features are: (a) the selection of and number of indicator variables, (b) the nature and size of the sample, and (c) the communality magnitude. Understanding the impact of variable-to-factor ratio (VTF), sample size or subject-to-variable ratio (STV), and communalities (h2) magnitude in FA analyses is relevant because these features affect the model precision and operationalized (measured) latent variable (factor) variance, which determines model invariance of FA findings.

The benefit of FA is based on its ability to produce a well-built, reliable, and understandable estimates of factor loadings [15, p. 154]. Therefore, understanding how VTF, STV, and h2 interact in FA and how they possibly influence or change the model precision and operationalized (measured) latent variable (factor) variance is the basic problem investigated in this study[16]–[29].

The model precision in this research is operationalized along psychometric lines, not statistical. Statistically, precision is inversely related to the standard error of the sampling distribution and is related to the minimizing the standard error of a statistic. Psychometrically, precision can mean this, but additionally, in a reliability context it can also refer to the accuracy of the estimator to be near (or the same) as the theoretical latent variable (e.g., the true score) [1], [10]. Thus, as the standard error of measurement decreases the precision/accuracy of the observed scores converges to the true score. However, no comprehensive study has been found in the existing literature that has systematically examined the incremental or combined impacts of two features of ED and SD and how best to estimate the model. Therefore, evaluating the impact of ED and SD effects on WSV in FA findings is the basis of the proposed Monte Carlo simulation study[30], [31].

Three major concerns have emerged repeatedly in the literature related to the use and interpretation of FA in social science research: (a) determining an adequate number of indicator variables to describe the latent trait; (b) factoring a sufficient sample size to have reasonable confidence in the stability of the model estimate; and (c) establishing minimum communality levels to determine which indicator variables

Manuscript received December 5, 2021
Manuscript revised December 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.12.23

can represent a latent trait, especially in simulation studies [8, 10–13].

Factor Analysis (FA) assumes that the indicator variables used should be linearly related to one another. Otherwise, the number of extracted factors will be the same as the number of original variables [2, 15]. Survey instrument length and number of variables differ based on discipline, purpose, sample frame, and method of data collection. Recently, the online survey has become an important method of data collection for many researchers and scholars for a variety of reasons (e.g., online surveys are easy to design, conduct, and sometimes they are the only option for data collection). According to SurveyMonkey the median length of its paid surveys was 9 questions [37]. While industry-specific surveys and market-research surveys tend to have more questions, event surveys and just-for-fun surveys tend to be shorter [37]. If the length of the survey is about 9 questions or fewer, it can lead to a higher completion rate and increase the likelihood that people will choose to take the researcher's surveys in the future. More recent studies of factor analysis in the literature do not include the VTF ratio 9:1 in their investigations [5, 12, 15–17], nor how this number is relative to sample size when factor analysis is conducted.

2. Theoretical Consideration

Researchers should determine an adequate number of indicator variables that is required to produce a stable and precise model to describe the latent trait. Fabrigar et al. [41] investigated the effects of indicator variables on pattern recovery to determine the sufficient number of indicator variables that is likely to produce patterns that closely approximate the population pattern. They reported that the number of indicator variables can strongly affect the degree to which a sample pattern reproduces the population pattern, and a minimum of three variables per factor is critical. The information about the adequate number of indicator variables that is required to produce a stable and precise model can be used in the design of a study and, retrospectively, in the evaluation of an existing study.

A larger sample size is better than a smaller sample size because it is minimizing misfit and the probability of errors. In many cases, increasing the sample size may not be possible. In medical research, it is very difficult to collect a large sample of patients suffering from a certain disease [22–24]. Investigating the minimum STV ratio or small absolute sample size to obtain the stability of the model is necessary. Only a very limited number of studies on the role of sample size in factor analysis have investigated real or simulated small sample size. De Winter, Dodou, and Wieringa [40] investigated the minimum sample size necessary to obtain reliable factor solutions under various conditions. They concluded that under the conditions of high communality, high number of observed variables, and small number of factors, FA yields a stable estimates model for sample sizes below 50.

Previous research has investigated the stability of factor solutions by the examination of chi-square value (χ^2) and overall model fit indices (OMF) such as goodness-of- fit index (GFI), adjusted goodness-of-fit index (AGFI), Tucker-Lewis index (TLI), comparative fit index (CFI), root mean square error of approximation (RMSEA), and root mean square residual (RMR)[8, 10, 18, 25–27]. Overall model fit indices examined global measures of data-model fit. Examinations of measurement invariance (MIV) (configural, weak, and strong) were used to evaluate model stability.

Selecting the adequate sample size is an important decision in study design. A researcher must determine how large the sample should be and what is the most appropriate sampling frame. Literature has proposed tremendous guidelines for estimating an adequate sample size for FA [2, 4, 11, 36, 40].

3. Methods

Simulation data are used in social science to answer a particular research question, solve a statistical problem, or improve analysis procedures techniques. Statistical program developers and research designers usually perform simulation data techniques for several reasons: gathering real data may be difficult, time-consuming, expensive, or real data sometime violate distributional assumptions. Simulation data often leads to greater understanding of an analysis and the results one can expect from various oddities of real-life data [10]. Simulation may approximate real-world results yet requires less time and effort and gives the researcher a chance to experiment with data under various conditions.

3.1 Research procedure

The study was designed to investigate empirical evaluation of accuracy of model fit in growing levels of invariance. This study manipulated: (a) variable-to-factor ratio (3:1, 6:1, and 9:1) that were randomly sampled from a population of 1000 indicator variables, (b) subject-to-variable ratio of 3:1 to 39:1 in multiple of 2 (3:1, 6:1, 9:1, 12:1, and 39:1), and (c) communality magnitude (high, moderate, low, and mixed). These factors were varied in a known factor structure with: (a) continuous variables (measurement scale), (b) normal distribution, (c) 6-factor solutions (common factor), and (d) orthogonal solution (factor structure).

3.2 Invariance

The present study used (MGCFA) multiple-group confirmatory factor analysis model to exam invariance of the effectiveness scale across students' classifications (gender and status). Table 1 illustrates the order for testing measurement invariance starting with configural invariance (model 0). Model testing was evaluated by the chi-square difference test ($\Delta\chi^2$) between two groups, and RSMA, CFI, and TLI were used to evaluate all of the model fits. As previously referenced, the following criteria values suggested were used in this study: RMSEA: 0.00 - 0.05 very decent fit, CFI > 0.95 decent fit, and TLI ≥ 0.96 decent fit.

Table 1. Procedure for Testing Stability Among Models

M	Test Name	H_0	Symbol	$\Delta\chi^2$ Test	Test Statistics Guide
M0	Configural invariance	$H_0: \lambda^1_{group} = \lambda^2_{group} = \cdots = \lambda^g_{group}$	λ : The number of factor patterns across g^{th} groups		If $\Delta\chi^2$ NS, model shows configural factorial invariance in place
M1	Weak measurement invariance	$H_0: \lambda_j^{group1} = \lambda_j^{group2} = \cdots = \lambda_j^{groupg}$	λ_j^{group1} : The factor loading of j^{th}indicator variable in the group	$\Delta\chi^2{}_{M1-M0}$	If $\Delta\chi^2$ NS, model shows weak factorial invariance in place
M2	Strong measurement invariance	$H_0: \tau_j^{group1} = \tau_j^{group2} = \cdots = \tau_j^{groupg}$	τ : The indicator variables intercept (means) of j^{th} indicator variable in the group	$\Delta\chi^2{}_{M2-M1}$	If $\Delta\chi^2$ NS, model shows strong factorial invariance in place

4. Results

Table 2a to 2c presents complete findings of measurement invariance for mixed communality among levels of STV over 1000 replications where significant p-value marked with "*".

VTF (3:1). examination of the measurement invariance, beginning with configural (M0) to weak (M1) to strong (M2). The findings revealed that for VTF ratios (6:1), $\chi^2{}_{M_0}$ showed statistically significant results when testing configural invariance: (3:1 with 39:1); (3:1 with 12:1); (3:1 with 9:1); and (3:1 with 6:1). Thus, non-invariance was established precluding further invariance testing, e.g., weak, strong, and structural. However, at higher STV ratios, e.g., groups (6:1 with 39:1); (6:1 with 12:1); (6:1 with 9:1); (9:1 with 39:1), (9:1 with 12:1); (9:1 with 39:1) and (12:1 with 39:1), $\chi^2{}_{M_0}$ was not statistically significant indicating configural invariance was established. Given the presence of configural invariance, testing for weak invariance was conducted. Again, chi-square difference between $\Delta\chi^2{}_{M1-M0}$ was not statistically significant supporting the hypothesis of weak factorial invariance between the two groups. After weak invariance was supported, examination of the indicator intercepts was tested. Results again supported the finding of strong invariance, e.g., the $\Delta\chi^2{}_{M2-M1}$ was not statistically significant. In conclusion, there were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved as described above.

VTF (6:1). The model beginning with configural (M0) to weak (M1) to strong (M2). The findings revealed that for VTF ratios (6:1), $\chi^2{}_{M_0}$ showed statistically significant results when testing configural invariance: (3:1 with 39:1); (3:1 with 12:1); (3:1 with 9:1); (3:1 with 6:1); and (6:1 with 9:1). Thus, non-invariance was established precluding further invariance testing, e.g., weak, strong, and structural. However, at higher STV ratios, e.g., groups (6:1 with 39:1); (6:1 with 12:1); (12:1 with 39:1), (9:1 with 12:1); (9:1 with 39:1) and (12:1 with 39:1), $\chi^2{}_{M_0}$ was not statistically significant indicating configural invariance was established. Given the presence of configural invariance, testing for weak invariance was conducted. Again, chi-square difference between $\Delta\chi^2{}_{M1-M0}$ was not statistically significant supporting the hypothesis of weak factorial invariance between the two groups. After weak invariance was supported, examination of the indicator intercepts was tested. Results again supported the finding of strong invariance, e.g., the $\Delta\chi^2{}_{M2-M1}$ was not statistically significant. In conclusion, there were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved as described above.

VTF (9:1). The findings revealed that for VTF ratios (9:1), $\chi^2{}_{M_0}$ showed statistically significant results when testing configural invariance: (3:1 with 39:1); (3:1 with 12:1); (3:1 with 9:1); (3:1 with 6:1); (6:1 with 9:1); (6:1 with 12:1); and (6:1 with 39:1). Thus, non-invariance was established precluding further invariance testing, e.g., weak, strong, and structural. However, at higher STV ratios, e.g., groups (9:1 with 12:1); (9:1 with 39:1) and (12:1 with 39:1), $\chi^2{}_{M_0}$ was not statistically significant indicating configural invariance was established. Given the presence of configural invariance, testing for weak invariance was conducted. Again, chi-square difference between $\Delta\chi^2{}_{M1-M0}$ was not statistically significant supporting the hypothesis of weak factorial invariance between the two groups. After weak invariance was supported, examination of the indicator intercepts was tested. Results again supported the finding of strong invariance, e.g., the $\Delta\chi^2{}_{M2-M1}$ was not statistically significant. In conclusion, there were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved as described above.

Table 2a. Examination for factorial-invariance (measurement and structural) across levels of education groups

VTF	Between Groups	χ^2	M	$\Delta\chi^2$	p-value	RMSEA CFI TLI

						GFI
3:1	STV=3:1 & STV=39:1	369.18	M0	---	0.0301*	0.0201 0.9893 0.9874
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=12:1	371.33	M0	---	0.0252*	0.0285 0.9791 0.9754
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=9:1	375.77	M0	---	0.0172*	0.0399 0.9602 0.9530
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=6:1	386.26	M0	---	0.0065*	0.0573 0.9245 0.9104
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=6:1 & STV=39:1	340.29	M0	---	0.2084	0.0117 0.9952 0.9950
		356.07	M1-M0	15.78	0.3968	0.0117 0.9950 0.9951
		370.63	M2-M1	14.56	0.4835	0.0114 0.9950 0.9954
	STV=6:1 & STV=12:1	342.43	M0	---	0.1859	0.0164 0.9907 0.9902
		358.02	M1-M0	15.59	0.4098	0.0163 0.9904 0.9904
		372.60	M2-M1	14.58	0.4820	0.0159 0.9905 0.9910
	STV=6:1 & STV=9:1	346.89	M0	---	0.1444	0.0237 0.9825 0.9808
		362.81	M1-M0	15.92	0.3873	0.0236 0.9819 0.9810
		377.46	M2-M1	14.65	0.4769	0.0229 0.9820 0.9820
	STV=9:1 & STV=12:1	329.85	M0	---	0.3402	0.0082 0.9971 0.9978
		345.07	M1-M0	15.22	0.4356	0.0081 0.9970 0.9979
		360.13	M2-M1	15.06	0.4471	0.0080 0.9969 0.9979
	STV=9:1 & STV=39:1	331.88	M0	---	0.3120	0.0115 0.9946 0.9957
		346.96	M1-M0	15.08	0.4456	0.0113 0.9945 0.9958
		362.06	M2-M1	15.1	0.4442	0.0111 0.9945 0.9960
		325.34	M0	---	0.4066	0.0063
						0.9980 0.9990
	STV=12:1 & STV=39:1	340.40	M1-M0	15.06	0.4471	0.0062 0.9980 0.9990
		355.28	M2-M1	14.88	0.4600	0.0061 0.9980 0.9991

Table 2b. Examination for factorial-invariance (measurement and structural) across levels of education groups

VTF	Between Groups	χ^2	M	$\Delta\chi^2$	p-value	RMSEA CFI TLI GFI
6:1	STV=3:1 & STV=39:1	1257.15	M0	---	0.0006*	0.0152 0.9942 0.9938
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=12:1	1262.68	M0	---	0.0004*	0.0213 0.9888 0.9879
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=9:1	1262.65	M0	---	0.0004*	0.0213 0.9888 0.9879
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=6:1	1310.82	M0	---	<0.0001*	0.0425 0.9579 0.9544
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=6:1 & STV=39:1	1170.54	M0	---	0.0686	0.0093 0.9975 0.9973
		1200.54	M1-M0	30	0.4656	0.0092 0.9975 0.9974
		1230.60	M2-M1	30.06	0.4625	0.0090 0.9975 0.9975
	STV=6:1 & STV=12:1	1176.06	M0	---	0.0549	0.0130 0.9952 0.9949
		1206.12	M1-M0	30.06	0.4625	0.0128 0.9951 0.9950
		1236.47	M2-M1	30.35	0.4478	0.0127 0.9951 0.9951
	STV=6:1 & STV=9:1	1191.49	M0	---	0.0279*	0.018 0.9905 0.9898
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---

	STV=9:1 & STV=12:1	1137.82	M0	---	0.2086	0.0062 0.9986 0.9987
		1167.98	M1-M0	30.16	0.4574	0.0061 0.9986 0.9987
		1197.93	M2-M1	29.95	0.4682	0.0061 0.9986 0.9987
	STV=9:1 & STV=39:1	1143.33	M0	---	0.1773	0.0086 0.9975 0.9975
		1173.42	M1-M0	30.09	0.4610	0.0085 0.9975 0.9976
		1203.61	M2-M1	30.19	0.4559	0.0084 0.9974 0.9976
	STV=12:1& STV=39:1	1122.39	M0	---	0.3133	0.0044 0.9991 0.9993
		1152.28	M1-M0	29.89	0.4712	0.0044 0.9991 0.9993
		1182.31	M2-M1	30.03	0.4641	0.0044 0.9991 0.9994

Table 2c. Examination for factorial-invariance (measurement and structural) across levels of education groups

VTF	Between Groups	χ^2	M	$\Delta\chi^2$	p-value	RMSEA CFI TLI GFI
9:1	STV=3:1 & STV=39:1	2640.05	M0	---	<0.0001*	0.01241 0.9948 0.9946
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=12:1	2654.33	M0	---	<0.0001*	0.0174 0.9899 0.9894
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=9:1	2681.11	M0	---	<0.0001*	0.0244 0.9806 0.9796
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=3:1 & STV=6:1	2748.81	M0	---	<0.0001*	0.0345 0.9621 0.9601
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=6:1 & STV=39:1	2464.25	M0	---	<0.0001*	0.0076 0.9978 0.9978
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=6:1 & STV=12:1	2478.53	M0	---	<0.0001*	0.0108 0.9958 0.9956
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=6:1 & STV=9:1	2505.31	M0	---	0.0059*	0.0154 0.9918 0.9914
		---	M1-M0	---	---	--- --- ---
		---	M2-M1	---	---	--- --- ---
	STV=9:1 & STV=12:1	2396.55	M0	---	0.1648	0.0047 0.9989 0.9990
		2441.69	M1-M0	45.14	0.4661	0.0047 0.9989 0.9990
		2486.29	M2-M1	44.6	0.4470	0.0046 0.9989 0.9990
	STV=9:1 & STV=39:1	2410.83	M0	---	0.1190	0.0069 0.9980 0.9980
		2455.99	M1-M0	45.16	0.4652	0.0068 0.9980 0.9980
		2500.68	M2-M1	44.69	0.4432	0.0067 0.9980 0.9980
	STV=12:1& STV=39:1	2369.77	M0	---	0.2782	0.0034 0.9993 0.9995
		2413.98	M1-M0	44.21	0.5053	0.0033 0.9993 0.9995
		2458.99	M2-M1	45.01	0.4715	0.0033 0.9993 0.9995

5. Conclusion

In general, this study provided empirical evaluation of the accuracy of the data-model fit over increasingly levels of factorial invariance for different feature of design in factor analysis. The study concluded that accuracy of the models was varying over increasingly levels of measurement as a function of VTF, STV, and their interactions. There were invariant factor loadings and invariant intercepts among the groups indicating that measurement invariance was achieved. For VTF ratio (3:1, 6:1, and 9:1) the models started to showed stability over levels of measurement when STV ratio was 3:1. Yet, the frequency of stability models over 1000 replications increased (from 69% to 89%) as STV ratio increased. The models showed more accuracy at or above 39:1 STV.

References

[1] M. J. Allen and W. M. Yen, "Introduction to measurement theory, Monterey, CA: Brooks/Cole, 1979," *Google Scholar*, 1979.

[2] K. M. Marcoulides, N. Foldnes, and S. Grønneberg, "Assessing model fit in structural equation modeling using appropriate test statistics," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 27, no. 3, pp. 369–379, 2020.

[3] Y. Rosseel, "Small sample solutions for structural equation modeling," *SMALL SAMPLE SIZE SOLUTIONS*, p. 226, 2020.

[4] S. C. Smid, D. McNeish, M. Miočević, and R. van de Schoot, "Bayesian versus frequentist estimation for structural equation models in small sample contexts: A systematic review," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 27, no. 1, pp. 131–161, 2020.

[5] S. Zitzmann and M. Hecht, "Going beyond convergence in Bayesian estimation: Why precision matters too and how to assess it," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 26, no. 4, pp. 646–661, 2019.

[6] L. Crocker and J. Algina, *Introduction to classical and modern test theory.* ERIC, 1986.

[7] E. G. W. Velicer, "Relation of sample size to the stability of component patterns," *Psychological Bulletin*, vol. 103, pp. 265–275, 1988.

[8] Y. A. Wang and M. Rhemtulla, "Power analysis for parameter estimation in structural equation modeling: A discussion and tutorial," 2020.

[9] J. C. Westland, "Lower bounds on sample size in structural equation modeling," *Electronic commerce research and applications*, vol. 9, no. 6, pp. 476–487, 2010.

[10] D. L. Bandalos and P. Gagné, "Simulation methods in structural equation modeling.," 2012.

[11] T. A. Brown, *Confirmatory factor analysis for applied research*. Guilford publications, 2015.

[12] P. F. M. Bullón, "Failing to replicate: Hypothesis testing as a crucial key to make direct replications more credible and predictable," 2015.

[13] O. P. John and S. Srivastava, "The Big Five trait taxonomy: History, measurement, and theoretical perspectives," *Handbook of personality: Theory and research*, vol. 2, no. 1999, pp. 102–138, 1999.

[14] J. Wang and X. Wang, *Structural equation modeling: Applications using Mplus*. John Wiley & Sons, 2019.

[15] D. Moody, "The Method Evaluation Model: A Theoretical Model for Validating Information Systems Design Methods," *ECIS 2003 Proceedings*, Jan. 2003, [Online]. Available: https://aisel.aisnet.org/ecis2003/79.

[16] D. Almaleki, "Empirical Evaluation of Different Features of Design in Confirmatory Factor Analysis," 2016.

[17] C. S. Wardley, E. B. Applegate, A. D. Almaleki, and J. A. Van Rhee, "A comparison of Students' perceptions of stress in parallel problem-based and lecture-based curricula," *The Journal of Physician Assistant Education*, vol. 27, no. 1, pp. 7–16, 2016.

[18] C. S. Wardley, E. B. Applegate, A. D. Almaleki, and J. A. Van Rhee, "Is Student Stress Related to Personality or Learning Environment in a Physician Assistant Program?," *The Journal of Physician Assistant Education*, vol. 30, no. 1, pp. 9–19, 2019.

[19] D. Almaleki, "Examinee Characteristics and their Impact on the Psychometric Properties of a Multiple Choice Test According to the Item Response Theory (IRT)," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6889–6901, 2021.

[20] D. Almaleki, "Stability of the Data-Model Fit over Increasing Levels of Factorial Invariance for Different Features of Design in Factor Analysis," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 6849–6856, 2021.

[21] D. Almaleki, "The Precision of the Overall Data-Model Fit for Different Design Features in Confirmatory Factor Analysis," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6766–6774, 2021.

[22] D. A. Almaleki, "Challenges Experienced Use of Distance-Learning by High School Teachers Responses to Students with Depression," *International Journal of Computer Science and Network Security*, vol. 21, no. 5, pp. 192–198, May 2021, doi: 10.22937/IJCSNS.2021.21.5.27.

[23] D. A. Almaleki, "The Psychometric Properties of Distance-Digital Subjective Happiness Scale," *International Journal of Computer Science and Network Security*, vol. 21, no. 5, pp. 211–216, May 2021, doi: 10.22937/IJCSNS.2021.21.5.29.

[24] D. A. Almaleki, R. A. Alhajaji, and M. A. Alharbi, "Measuring Students' Interaction in Distance Learning Through the Electronic Platform and its Impact on their Motivation to Learn During Covid-19 Crisis," *International Journal of Computer Science and Network Security*, vol. 21, no. 5, pp. 98–112, May 2021, doi: 10.22937/IJCSNS.2021.21.5.16.

[25] D. A. Almaleki, W. W. Khayat, T. F. Yally, and A. A. Al-hajjaji, “The Effectiveness of the Use of Distance-Evaluation Tools and Methods among Students with Learning-Difficulties from the Teachers’ Point of View,” *International Journal of Computer Science and Network Security*, vol. 21, no. 5, pp. 243–255, May 2021, doi: 10.22937/IJCSNS.2021.21.5.34.

[26] “Evaluating Psychological Experiences of Saudi Students in Distance-Learning.” https://scholar.google.com/citations?view_op=view_citation&hl=en&user=RWnye6UAAAAJ&citation_for_view=RWnye6UAAAAJ:kNdYIx-mwKoC (accessed Oct. 13, 2021).

[27] “Factor Structure, Validity and Reliability of The Teacher Satisfaction Scale (TSS) In Distance-Learning During Covid-19 Crisis: Invariance Across Some Teachers’ Characteristics.” https://scholar.google.com/citations?view_op=view_citation&hl=en&user=RWnye6UAAAAJ&citation_for_view=RWnye6UAAAAJ:MXK_kJrjxJIC (accessed Oct. 13, 2021).

[28] “The Effect of Methods of Estimating the Ability on The Accuracy and Items Parameters According to 3PL Model.” https://scholar.google.com/citations?view_op=view_citation&hl=en&user=RWnye6UAAAAJ&citation_for_view=RWnye6UAAAAJ:8k81kl-MbHgC (accessed Oct. 13, 2021).

[29] “The Psychometric Properties of Distance-Digital Subjective Happiness Scale.” https://scholar.google.com/citations?view_op=view_citation&hl=en&user=RWnye6UAAAAJ&citation_for_view=RWnye6UAAAAJ:UebtZRa9Y70C (accessed Oct. 13, 2021).

[30] L. H. Lam, T. D. H. Phuc, and N. H. Hieu, “Simulation Models For Three-Phase Grid Connected PV Inverters Enabling Current Limitation Under Unbalanced Faults,” *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5396–5401, Apr. 2020, doi: 10.48084/etasr.3343.

[31] A. H. Blasi and M. Alsuwaiket, “Analysis of Students’ Misconducts in Higher Education using Decision Tree and ANN Algorithms,” *Engineering, Technology & Applied Science Research*, vol. 10, no. 6, pp. 6510–6514, Dec. 2020, doi: 10.48084/etasr.3927.

[32] D. Almaleki, “The Precision of the Overall Data-Model Fit for Different Design Features in Confirmatory Factor Analysis,” *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6766–6774, Feb. 2021, doi: 10.48084/etasr.4025.

[33] B. Thompson, “Ten commandments of structural equation modeling.,” in *US Dept of Education, Office of Special Education Programs (OSEP) Project Directors’ Conference, 1998, Washington, DC, US; A previous version of this chapter was presented at the aforementioned conference and at the same annual conference held in 1999.*, 2000.

[34] K. B. Coughlin, “An analysis of factor extraction strategies: A comparison of the relative strengths of principal axis, ordinary least squares, and maximum likelihood in research contexts that include both categorical and continuous variables,” 2013.

[35] J. J. Hox, C. J. Maas, and M. J. Brinkhuis, “The effect of estimation method and sample size in multilevel structural equation modeling,” *Statistica neerlandica*, vol. 64, no. 2, pp. 157–170, 2010.

[36] A. J. Morin, N. D. Myers, and S. Lee, “Modern Factor Analytic Techniques: Bifactor Models, Exploratory Structural Equation Modeling (ESEM), and Bifactor-ESEM,” *Handbook of sport psychology*, pp. 1044–1073, 2020.

[37] D. Almaleki, “Empirical Evaluation of Different Features of Design in Confirmatory Factor Analysis,” 2016.

[38] R. K. Henson and J. K. Roberts, “Use of exploratory factor analysis in published research: Common errors and some comment on improved practice,” *Educational and Psychological measurement*, vol. 66, no. 3, pp. 393–416, 2006.

[39] A. W. Meade and D. J. Bauer, “Power and precision in confirmatory factor analytic tests of measurement invariance,” *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 14, no. 4, pp. 611–635, 2007.

[40] J. C. de Winter*, D. Dodou*, and P. A. Wieringa, “Exploratory factor analysis with small sample sizes,” *Multivariate behavioral research*, vol. 44, no. 2, pp. 147–181, 2009.

[41] L. R. Fabrigar, D. T. Wegener, R. C. MacCallum, and E. J. Strahan, “Evaluating the use of exploratory factor analysis in psychological research.,” *Psychological methods*, vol. 4, no. 3, p. 272, 1999.

[42] R. Jacobucci, A. M. Brandmaier, and R. A. Kievit, “A practical guide to variable selection in structural equation modeling by using regularized multiple-indicators, multiple-causes models,” *Advances in methods and practices in psychological science*, vol. 2, no. 1, pp. 55–76, 2019.

[43] E. Guadagnoli and W. F. Velicer, “Relation of sample size to the stability of component patterns.,” *Psychological bulletin*, vol. 103, no. 2, p. 265, 1988.

[44] A. B. Costello and J. Osborne, “Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis,” *Practical assessment, research, and evaluation*, vol. 10, no. 1, p. 7, 2005.

[45] X. An and Y.-F. Yung, “Item Response Theory: What It Is and How You Can Use the IRT Procedure to Apply It,” p. 14.

[46] F. B. Bryant and P. R. Yarnold, "Principal-components analysis and exploratory and confirmatory factor analysis.," 1995.
[47] C. M. Ringle, M. Sarstedt, R. Mitchell, and S. P. Gudergan, "Partial least squares structural equation modeling in HRM research," *The International Journal of Human Resource Management*, vol. 31, no. 12, pp. 1617–1643, 2020.

Deyab A. Almaleki is an Associate Professor in the Department of Evaluation, Measurement and Research. Dr. Almaleki received his Ph.D. from Western Michigan University (USA) in 2016 in Evaluation, Measurement and Research. Since 2011, Dr. Almaleki has authored and co-authored in more than 20 peer-reviewed journal articles, and over 30 peer-reviewed presentations at professional conferences. Dr. Almaleki has extensive experience in educational and psychological research, research design, applied statistics, structural equation modeling, design, and analysis of psychological measurement.