

The Precision of the Overall Data-Model Fit for Different Design Features in Confirmatory Factor Analysis

Deyab Almaleki

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

Abstract-Factor Analysis (FA) is the study of variance within a group. Within-Subject Variance (WSV) is affected by multiple features in a study context such as the Experimental Design (ED) or the Sampling Design (SD). The aim of this study is to provide an empirical evaluation of the influence of different aspects of ED and SD on WSV in the context of FA in terms of model precision. The study results showed that the precisions of the overall model fit indices TLI and CFI, as functions of VTF, STV, h^2 , and their interaction, varied, as did the precisions of the overall model fit indices GFI, AGFI, and RMSEA as functions of VTF, STV, and their interactions. Overall, when the VTF is 4:1 or 7:1, the required STV is 16:1 or above 32:1 or above to show precision in factor solution.

Keywords-model-precision; factor-analysis; model-fit; model-design

I. INTRODUCTION

Factor Analysis (FA) is a useful and flexible analytic family of methods that plays a critically important role in many empirical applications. FA is the study of variance within a group, as opposed to statistical analysis, which focuses on partitioning variance among groups [1]. FA is extensively applied in education and behavioral science research. A recent two-year analysis (from 2003 to 2005) of the use of FA methods as indexed by the PsycINFO database revealed that more than 1,700 studies used some form of FA [2]. However, FA is a generic term for a family of statistical techniques, including Exploratory FA (EFA) and Confirmatory FA (CFA). The fundamental purpose of EFA is to identify unknown latent constructs in a relatively large set of measured indicator variables that can summarize (or reproduce) the observed covariance or correlation pattern among a set of indicator variables. EFA is often used when the researcher may not have specific expectations of the number of constructs or factors underlying the dimensional structure of the observed correlational pattern, or even in cases when the researcher has emergent ideas about the underlying dimensional structure of the observed correlational pattern among a larger set of indicator variables [1-3]. When conducting an EFA study or analysis, the researcher faces a multitude of methodological and technical decisions. For example, there are two different

EFA statistical models to choose from: (a) the full component model, or (b) the common factor model [4-6]. The implications of choosing either one need to be well understood. CFA is a form of factor analysis that tests hypotheses regarding how well the measured indicator variables represent the number of constructs [7, 8]. CFA is a confirmatory method that can be used to examine, evaluate, and/or test the number of hypothesized factors underlying the variance/covariance in a set of measured indicator variables. CFA allows testing hypothetical and plausible alternative latent variable structures for the observed indicator variance/covariance [9, 10]. More recently, CFA has also been used in exploratory analysis too. Both EFA and CFA attempt to understand the variance of the observed indicator variables through studying the Within-Subject Variance (WSV). It is well understood that WSV is affected by many features of the study conduct, such as the study Experimental Design (ED) and the Sampling Design (SD). Thus, anything that influences or changes variance may affect the conclusions related to FA. Previous studies have isolated one or two elements within ED and SD [8, 11, 12]. However, to the best of our knowledge, no study provided a comprehensive examination of multiple WSV factors, yet this is precisely what a researcher must do when planning a WSV study.

To understand the impact of ED and 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) [13-17]. The FIV methods offer a structure that allows disentangling measurement elements from structural elements in the factor model. Via FIV and the evaluation of data-model fit, the impact of ED and SD on WSV can be compared among groups by examination of model precision [1, 12, 13, 18, 19]. Previous researches investigated the precision of factor solutions by the examination of chi-square value (χ^2) and Overall Model Fit (OMF) indices such as Goodness-of-Fit Index (GFI), Adjusted GFI (AGFI), Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA) [3, 9, 20-22]. The overall model fit indices examine the global measures of data-model fit.

Corresponding author: Deyab Almaleki

There are three key features of design that are of paramount importance and generally overshadow all the technical decisions a researcher faces. These three features are: (a) the selection and the number of indicator variables, (b) the nature and size of the sample, and (c) the communality magnitude. Understanding the impact of Variable-To-Factor (VTF) ratio, sample size or Subject-To-Variable (STV) ratio, and communalities (h^2) magnitude in FA analysis is relevant because these features affect the model precision and operationalized (measured) latent variable (factor) variance, which determines the model invariance of FA findings. The benefit of FA is based on its ability to produce well-built, reliable, and understandable estimates of factor loadings [23]. Therefore, understanding how VTF, STV, and h^2 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.

The model precision in this research is operationalized along psychometric and not statistical lines. Statistically, precision is inversely related to the standard error of the sampling distribution and related to the minimizing of 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) [9, 13]. 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 to have systematically examined the incremental or combined impacts of two features of ED and SD and the optimum way 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 [24, 25].

II. ED: VARIABLE-TO-FACTOR RATIO

Previous research indicates that CFA yields more precise results when each common factor is represented by multiple indicator variables in the analysis [5, 21, 26]. While this concern is related to model identification [1, 14, 18, 27, 28], this is not the focus given here. Specifically, given an identified model, the way the sample and number of indicator variables inform the understanding of the latent factor is an important question which has not received great attention in the CFA literature [2, 12, 17, 28-30]. Authors in [18] concluded that the VTF ratio was important for factor stability with more indicator variables per factor yielding more stable result. However, the researchers who investigated VTF ratio have not reached a mutual decision on an optimal VTF. Authors in [18, 31] concluded that a VTF of 3:1 is sufficient for factorial precision, whereas authors in [11] found that 24.6% of the studies published in Journal of Personality and Social Psychology (JPSP) and 34.4% of the studies published in Journal of Applied Psychology (JAP) have VTF ratio of 4:1 or less. However, other researchers who have investigated the effects of indicator variable sampling did not attempt to systematically manipulate conditions that could potentially affect the pattern stability of CFA findings. The issue of variable sampling has been used extensively in conceptual development, but the

existing literature has received almost no empirical evaluation that generally has sampled indicator variables at random from the universe of variables. The assumption of random sampling is useful to minimize sampling issues and for developing generalizability rather than a prescription for applied research procedures.

III. UNDERLYING FACTOR STRUCTURE

One of the broad domains in social science is the Big Five Personality Traits, which is used to describe human personality and includes extraversion, agreeableness, openness, conscientiousness, and neuroticism. Most commonly used in academic psychology, this model incorporates five different factors into a conceptual model for describing personality traits [3, 8, 32]. It has been selected as the structural model for this study because of its wide use in social sciences. The model structure of the Big Five Personality Traits theory has received favorable attention from researchers in the psychological discipline. The author in [3] concluded that the measurement structure of the Big Five Personality Traits is an orthogonal solution and the variation on each one of the Big Five Personality Trait dimensions is commonly proposed to be independent of variation on each of the others.

IV. SD: SUBJECT-TO-VARIABLE RATIO

Authors in [4, 12, 13, 15, 18-19, 22, 26, 33-36] found a mixed range in FA sample sizes described by either the absolute size of the sample or the STV ratio. Authors in [37] reviewed 60 studies utilizing FA and found the average minimum sample size was 42, and the minimum STV ratio was 3.25:1, with most of the studies using a STV ratio less than 5:1. Similarly, authors in [11] reviewed FA studies published in JPSP and JAP (from 1991 to 1995) and found that 18.9% of the published articles utilizing FA in JPSP and 13.8% in JAP had an average minimum sample size of 100 or less. Authors in [2] examined publications utilizing FA from the PsychINFO database between 2003 and 2005. They found that 15.4% of the studies reported STV from 10 to 20:1 and only 3% of the studies used a STV of 20-100:1. High absolute sample size or STV ratio is important to predict precise outcomes, increase the generalizability of the findings, and maximize the accuracy of population estimates [2, 4, 7, 22, 29]. There are many common practice rules of sample size in the literature, most of them not empirically based [18]. Moreover, a limited number of studies have empirically investigated the effect of STV on model precision [38]. Selecting the adequate sample size is an important decision in study design. The researcher must determine how large the sample should be and what is the most appropriate sampling frame. There are tremendous guidelines for estimating an adequate sample size for FA [2, 4, 11, 35, 39].

V. SD: COMMUNALITIES (h^2) MAGNITUDE

The communality of a variable can be interpreted as the proportion of variation estimated by the common factors. Communalities range from 0 to 1, where 0 means that the factors do not explain any of the variances and 1 means that all the variances are explained by the factors. A large value of communality suggests a strong effect by an underlying

construct [7, 20, 40]. Communality is the sum of the squared factor loadings for observed variables variances accounted for by all the factors [17, 19]. Communality measures the percentage of variance in the observed variables explained by all the factors. The larger the communality for each variable, the more successful the factor analysis solution is, and the smaller the communality, the more questionable the solution [4, 31, 39, 41]. Communality ranges between 0 and 1. If the communality exceeds 1.0, there is something wrong with the data, which may reflect model specification or SD problems. Low values of communalities across the set of observed variables indicate the variables are marginally related to each other and the factors provide little explanation of the variance in the observed variables [2, 20, 26, 39].

VI. METHODS

The current study was designed to investigate the empirical evaluation of the influence of different aspects of ED and SD on WSV in terms of model precision and operationalized (measured) latent variable (factor) variance relative to a known factor structure via Monte Carlo simulations. The study manipulated: (a) the VTF ratio (4:1, 7:1, and 10:1) randomly sampled from a population of 100 indicator variables, (b) the STV ratio from 2:1 to 32:1 in multiples of 2 (2:1, 4:1, 8:1, 16:1, and 32:1), and (c) the 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) 5-factor solutions (common factor), and (d) orthogonal solution (factor structure). The precision of factor solution was evaluated by the examination of CFA for an orthogonal 5-factor (common factor) model. Chi-squared value and overall model fit indices criteria were used to evaluate the models for all conditions. Chi-squared value and four model fit indices were treated as Dependent Variables (DVs) in three-way Analysis of Variance (ANOVA) using the level of h^2 , VTF ratio, and STV ratio as independent variables. Figure 1 illustrates the study design.

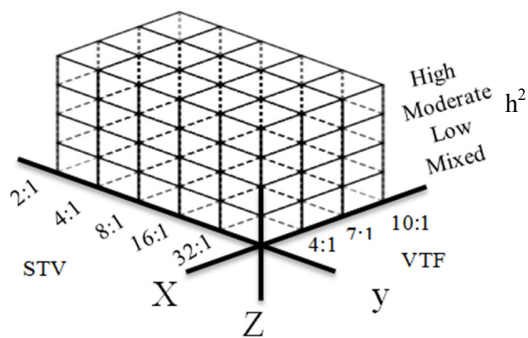


Fig. 1. Design of interaction conditions.

VII. RESULTS

The Monte Carlo simulation populated 60 three-way cells between subjects' factorial design matrix: h^2 (high, moderate, low, and mixed) by VTF ratio (4:1, 7:1, and 10:1) and STV ratio (2:1, 4:1, 8:1, 16:1, and 32) with 1000 replicate samples (see Figure 1). Chi-squared values and overall model fit indices

(GFI, RMSEA, TLI, and CFI) were treated as DVs in parallel three-way ANOVA. To control for multiplicity among DVs a type I error rate was adjusted by a Bonferroni correlation: $0.05/7=0.007$.

A. Chi-squared Test (χ^2)

The descriptive statistics for the chi-squared values among all design cells are presented in Table I.

TABLE I. DESCRIPTIVE STATISTICS FOR χ^2 AVERAGED OVER 1000 REPLICATIONS

VTF	STV	h^2			
		High	Moderate	Low	Mixed
4:1	2:1	220.0 23.8	219.86 24.60	219.37 23.41	220.42 23.53
	4:1	190.8 20.9	190.00 20.53	189.98 20.92	189.54 20.48
	8:1	179.3 19.2	179.20 19.64	180.09 19.59	178.70 19.13
	16:1	174.2 19.2	174.41 18.70	174.69 19.18	174.07 18.33
	32:1	171.7 18.6	171.53 19.68	172.81 18.08	171.63 17.72
7:1	2:1	709.5 43.6	709.71 42.05	705.16 43.12	709.22 42.58
	4:1	620.0 38.2	621.19 36.03	618.77 37.49	622.59 37.60
	8:1	588.4 34.5	589.28 36.56	587.06 34.58	589.76 36.09
	16:1	573.3 34.5	574.42 34.97	573.77 34.94	574.10 34.05
	32:1	566.2 33.9	566.33 33.26	568.73 33.68	568.56 33.57
10:1	2:1	1471.88 61.4	1469.46 59.60	1473.04 60.34	1472.52 60.92
	4:1	1298.02 55.8	1297.31 52.89	1296.04 53.45	1296.67 52.45
	8:1	1231.67 49.5	1230.17 49.97	1228.86 51.12	1228.82 50.92
	16:1	1203.03 48.4	1203.64 50.24	1201.30 50.37	1201.90 48.66
	32:1	1186.37 47.6	1187.40 49.00	1188.81 50.89	1187.65 49.20

The overall ANOVA findings for the χ^2 revealed the statistically significant model, $F(59,59940)=137546, p<0.0001$. As can be seen in Table II, there were statistically significant results for the main effects of VTF and STV, and a statistically significant interaction between VTF*STV. Surprisingly, no statistically significant main effect or interaction involving communality was found.

TABLE II. THREE-WAY ANOVA FOR χ^2 BY CONDITIONS

Conditions	df	F	p-value
h^2	3	0.49	0.6876
VTF	2	3967031	<0.0001
h^2 *VTF	6	1.39	0.2143
STV	4	32825.9	<0.0001
h^2 *STV	12	0.94	0.5024
VTF*STV	8	6224.92	<0.0001
h^2 *VTF*STV	24	0.70	0.8589

Post hoc analysis of the VTF*STV interaction focused on splitting out the levels of STV and examined five one-way ANOVAs for the levels of VTF. Visual examination of the mean χ^2 values suggests a small, but significant decrease in mean χ^2 values as STV increases among all levels of VTF that decreased in magnitude as VTF increased (see Figure 2). Table III presents the simple ANOVA effect on VTF for each STV level. Pairwise comparisons among the levels of VTF at each level of STV, e.g. VTF 4:1 vs. 7:1 @ STV=2:1, were statistically significant with all p-values <0.0001, and the directional pattern in means confirmed the decreasing trends seen in Figure 2.

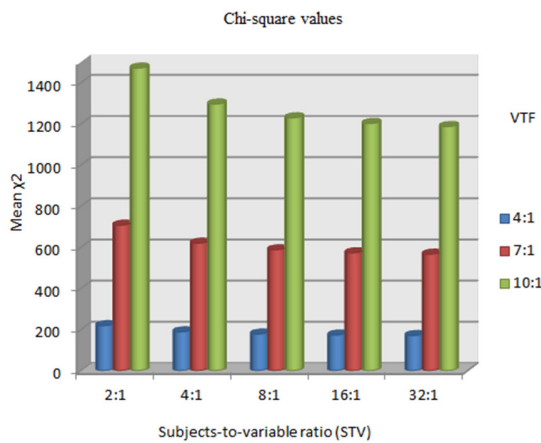


Fig. 2. χ^2 mean values of interaction between STV and VTF ratios.

TABLE III. ONE-WAY ANOVA SIMPLE EFFECT OF STV BY LEVELS OF VTF

STV	df	SS	MS	F	p-value
2:1	2	3184	1592	104503	<0.0001
4:1	2	2490	1245	817419	<0.0001
8:1	2	2243	1121	736161	<0.0001
16:1	2	2149	1074	705261	<0.0001
32:1	2	2096	1048	688051	<0.0001

B. Goodness-of-Fit Index

The averaged GFI statistics over 1000 replications for the 60 cells in the three-way design are presented in Table IV. The standard threshold of acceptable fit percentage is 0.90.

The overall ANOVA findings for GFI revealed an overall statistically significant model, $F(59,59940)=138368, p<0.0001$. As can be seen in Table V, there were statistically significant main effects of VTF and STV, and a statistically significant interaction between VTF*STV. Once again, the communality factor was not significant.

Post hoc analysis of VTF*STV interaction focused on splitting out the levels of STV and examined five one-way ANOVAs for the levels of VTF. Visual examination of the mean GFI values suggests a significant increase in mean GFI values as STV increases mixed with decreasing mean GFI as VTF ratios increased (see Figure 3).

TABLE IV. DESCRIPTIVE STATISTICS FOR GFI AVERAGED OVER 1000 REPLICATIONS

VTF	STV	h^2			
		High	Moderate	Low	Mixed
4:1	2:1	0.69	0.69	0.69	0.69
	4:1	10.82	0.82	0.82	0.82
	8:1	0.90	0.90	0.90	0.90
	16:1	65.60	65.70	61.60	68.00
	32:1	0.94	0.94	0.94	0.94
			100.00	100.00	100.00
7:1	2:1	0.68	0.68	0.68	0.68
	4:1	0.81	0.81	0.81	0.81
	8:1	0.90	0.90	0.90	0.90
	16:1	28.80	30.90	31.70	28.60
	32:1	0.94	0.94	0.94	0.94
			100.00	100.00	100.00
10:1	2:1	0.67	0.67	0.67	0.67
	4:1	0.80	0.80	0.80	0.80
	8:1	0.90	0.90	0.90	0.90
	16:1	7.60	7.80	9.10	8.80
	32:1	0.94	0.94	0.94	0.94
			100.00	100.00	100.00

TABLE V. THREE-WAY ANOVA FOR GFI BY CONDITIONS

Conditions	df	F	p-value
h^2	3	1.27	0.2820
VTF	2	6671.43	<0.0001
h^2 *VTF	6	1.52	0.1683
STV	4	2036352	<0.0001
h^2 *STV	12	1.41	0.1534
VTF*STV	8	610.99	<0.0001
h^2 *VTF*STV	24	0.85	0.6733

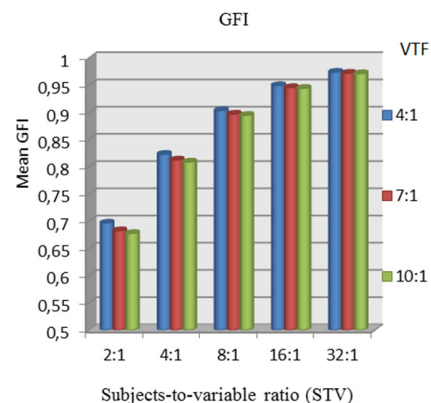


Fig. 3. Goodness-of-fit index of interaction between STV and VTF ratios.

Table VI presents the simple ANOVA effect on the VTF for each STV level. Pairwise comparisons among the levels of

VTF at each level of STV, e.g. VTF 4:1 vs. 7:1 @ STV=2:1, were statistically significant with all p-values <0.0001, and the directional pattern in means confirmed the patterns seen in Figure 3.

TABLE VI. ONE-WAY ANOVA SIMPLE EFFECT OF STV BY LEVELS OF VTF

STV	df	SS	MS	F	p-value
2:1	2	0.7851	0.3925	4941.23	<0.0001
4:1	2	0.4301	0.2150	2707.26	<0.0001
8:1	2	0.1641	0.0820	1032.93	<0.0001
16:1	2	0.0537	0.0268	338.14	<0.0001
32:1	2	0.0152	0.0076	95.83	<0.0001

C. Root Mean Square Error of Approximation

RMSEA statistics were averaged over 1000 replications for the 60 cells in the three-way design and are presented in Table VII. The standard threshold of acceptable fit is 0.05. The overall ANOVA findings for RMSEA revealed an overall statistically significant model, $F(59,59940)=4901.85$, $p<0.0001$. As can be seen in Table VIII, there were statistically significant main effects VTF and STV, and a statistically significant interaction between VTF*STV. Once again, the communality factor was not significant.

TABLE VII. DESCRIPTIVE STATISTICS FOR RMSEA AVERAGED OVER 1000 REPLICATIONS

VTF	STV	h ²			
		High	Moderate	Low	Mixed
4:1	2:1	0.08	0.08	0.08	0.08
		7.50	8.80	7.10	6.70
	4:1	0.03	0.03	0.03	0.03
		73.80	75.50	75.70	76.00
	8:1	0.01	0.01	0.01	0.01
		99.90	99.70	99.60	99.90
7:1	16:1	0.00	0.00	0.00	0.00
		100.00	100.00	100.00	100.00
	32:1	0.00	0.00	0.00	0.00
		100.00	100.00	100.00	100.00
	2:1	0.06	0.06	0.06	0.06
		11.50	8.80	13.20	10.60
10:1	4:1	0.02	0.02	0.02	0.02
		100.00	100.00	100.00	100.00
	8:1	0.01	0.01	0.01	0.01
		100.00	100.00	100.00	100.00
	16:1	0.00	0.00	0.00	0.00
		100.00	100.00	100.00	100.00
10:1	32:1	0.00	0.00	0.00	0.00
		100.00	100.00	100.00	100.00
	2:1	0.05	0.05	0.05	0.05
		47.20	45.20	47.20	46.70
	4:1	0.02	0.02	0.02	0.02
		100.00	100.00	100.00	100.00
10:1	8:1	0.00	0.00	0.00	0.00
		100.00	100.00	100.00	100.00
	16:1	0.00	0.00	0.00	0.00
		100.00	100.00	100.00	100.00
	32:1	0.00	0.00	0.00	0.00
		100.00	100.00	100.00	100.00

Post hoc analysis of VTF*STV interaction focused on splitting out the levels of STV and examined five one-way

ANOVAs for the levels of VTF. The visual examination of the mean RMSEA values suggests a significant decrease in mean RMSEA values as STV increases and an increasing mean RMSEA as VTF ratios decreased (see Figure 4).

TABLE VIII. THREE-WAY ANOVA FOR RMSEA BY CONDITIONS

Conditions	df	F	p-value
h ²	3	0.41	0.7467
VTF	2	6698.38	<0.0001
h ² *VTF	6	1.45	0.1919
STV	4	66127.9	<0.0001
h ² *STV	12	0.75	0.7077
VTF *STV	8	1408.03	<0.0001
h ² * VTF *STV	24	0.72	0.8337

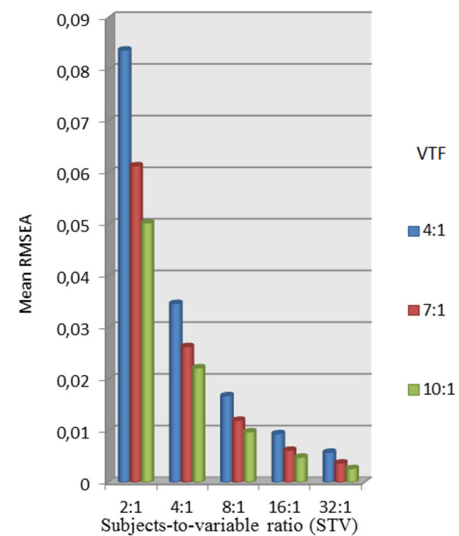


Fig. 4. RMSEA of interaction between STV and VTF ratios.

Table IX presents the simple ANOVA effect on the VTF for each STV level. Pairwise comparisons among the levels of VTF at each level of STV, e.g. VTF 4:1 vs. 7:1 @ STV=2:1, were statistically significant with all p-values <0.0001, and the directional pattern in means confirmed the patterns seen in Figure 4.

TABLE IX. ONE-WAY ANOVA SIMPLE EFFECT OF STV BY LEVELS OF VTF

STV	df	SS	MS	F	p-value
2:1	2	2.3236	1.1618	10188.3	<0.0001
4:1	2	0.3222	0.1611	1413.01	<0.0001
8:1	2	0.1021	0.0510	447.68	<0.0001
16:1	2	0.0436	0.0218	191.23	<0.0001
32:1	2	0.0206	0.0103	90.34	<0.0001

D. Non-Normed-Fit Index

Non-normed-fit index (or Tucker-Lewis Index-TLI) statistics were averaged over 1000 replications for the 60 cells in the three-way design and are presented in Table X. Additionally, the percentage of TLI statistics meeting the standard threshold of acceptable fit, e.g. 0.96, was tabulated.

The overall ANOVA findings for TLI revealed an overall statistically significant model, $F(59,59940)=4556.25$,

$p < 0.0001$. As can be seen in Table XI, there were statistically significant main effects and interactions, including a statistically significant triple interaction between $h^2 * VTF * STV$. The visual examination of the mean TLI values reveals a differential increase in mean TLI within levels of STV as a function of h^2 magnitude. In the h^2 =high condition, the mean TLI values show minimal gains even between 2:1 and 4:1 STV levels. Within the mixed and moderate h^2 conditions, the mean TLI values show asymptotic gains after $STV > 4:1$. However, in the low h^2 condition, the mean TLI values were markedly lower in the 2:1 and 4:1 STV levels, only showing asymptotic values when $STV > 8:1$ (see Figure 5).

TABLE X. DESCRIPTIVE STATISTICS FOR TLI AVERAGED OVER 1000 REPLICATIONS

VTF	STV	h^2			
		High	Moderate	Low	Mixed
4:1	2:1	0.92 15.90	0.81 4.60	0.53 1.90	0.82 3.90
	4:1	0.98 93.10	0.95 48.50	0.86 22.90	0.96 50.90
	8:1	0.99 100.00	0.98 91.30	0.96 48.20	0.99 94.40
	16:1	0.99 100.00	0.99 99.90	0.99 77.20	0.99 100.00
	32:1	0.99 100.00	0.99 100.00	0.99 95.90	0.99 100.00
	7:1	2:1	0.93 10.50	0.86 0.20	0.60 0.00
4:1		0.98 99.90	0.97 71.80	0.88 14.70	0.97 94.50
8:1		0.99 100.00	0.99 100.00	0.97 63.20	0.99 100.00
16:1		0.99 100.00	0.99 100.00	0.99 95.60	0.99 100.00
32:1		0.99 100.00	0.99 100.00	0.99 100.00	0.99 100.00
10:1	2:1	0.94 7.30	0.88 0.00	0.68 0.00	0.91 0.20
	4:1	0.98 100.00	0.97 87.60	0.91 10.80	0.98 99.90
	8:1	0.99 100.00	0.99 100.00	0.98 86.50	0.99 100.00
	16:1	0.99 100.00	0.99 100.00	0.99 100.00	0.99 100.00
	32:1	0.99 100.00	0.99 100.00	0.99 100.00	0.99 100.00

TABLE XI. THREE-WAY ANOVA FOR TLI BY CONDITIONS

Conditions	df	F	p-value
h^2	3	13451.2	<0.0001
VTF	2	1437.33	<0.0001
$h^2 * VTF$	6	211.73	<0.0001
STV	4	36383.3	<0.0001
$h^2 * STV$	12	5964.08	<0.0001
VTF * STV	8	659.74	<0.0001
$h^2 * VTF * STV$	24	80.84	<0.0001

Post hoc analysis of the 3-way interaction of $h^2 * VTF * STV$ first focused on the simple effect interaction of $h^2 * VTF$ after blocking on STV, specifically the test of 2-way interaction $h^2 * VTF$ at each level of STV. The results revealed statistically significant 2-way interactions at all STV levels: $STV=(2:1)$,

$F(6,11988)=160.34, p < 0.0001$; $STV=(4:1)$, $F(6,11988)=40.07, p < 0.0001$; $STV=(8:1)$, $F(6,11988)=18.51, p < 0.0001$; $STV=(16:1)$, $F(6,11988)=3.48, p=0.0019$; and $STV=(32:1)$, $F(6,11988)=0.43, p=0.0115$. Further analysis of the four by three 2-way interactions focused on the simple-simple effects of STV blocking on the $STV * h^2$ interaction. Specifically, the analysis examined differences in mean TLI values among VTF levels for each $STV * h^2$ interaction. Table XII presents these findings.

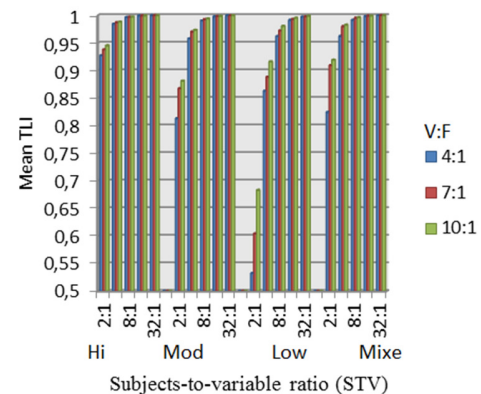


Fig. 5. Non-normal-fit index mean values for the interaction between STV and VTF ratios at different levels of communalities.

TABLE XII. SIMPLE-SIMPLE EFFECT OF $STV * VTF * h^2$ SLICED BY $STV * h^2$

h^2	STV	df	F	p-value	VTF		
					4:1vs7:1	4:1vs10:1	7:1vs10:1
Low	2	2	344	<0.001	<0.0001	<0.0001	<0.0001
	4	2	69.7	<0.001	<0.0001	<0.0001	<0.0001
	8	2	28.9	<0.001	<0.0001	<0.0001	0.0023
	16	2	5.1	0.005	0.3846	0.0041	0.1474
	32	2	3.4	0.031	0.6692	0.0266	0.1945
Moderate	2	2	446	<0.001	<0.0001	<0.0001	<0.0001
	4	2	85.3	<0.001	<0.0001	<0.0001	0.0563
	8	2	19.7	<0.001	<0.0001	<0.0001	0.2112
	16	2	3.7	0.023	0.1040	0.0251	0.8368
High	2	2	0.1	0.932	--	--	--
	2	2	156	<0.001	<0.0001	<0.0001	0.2111
	4	2	35.2	<0.001	<0.0001	<0.0001	0.0470
	8	2	6.3	0.001	0.1169	0.0011	0.2606
	16	2	0.6	0.522	--	--	--
Mixed	32	2	0.2	0.797	--	--	--
	2	2	1304	<0.001	<0.0001	<0.0001	0.0008
	4	2	204	<0.001	<0.0001	<0.0001	0.0322
	8	2	43.2	<0.001	<0.0001	<0.0001	0.1836
	16	2	9.26	<0.001	0.0025	0.0002	0.7672
32	2	1.02	0.3604	--	--	--	

E. Comparative-Fit Index (CFI)

CFI values were averaged over 1000 replications for the 60 cells in the three-way design and are presented in Table XIII. Additionally, the percentage of CFI values meeting the standard threshold of acceptable fit, e.g. 0.95 was tabulated. The overall ANOVA findings for CFI revealed an overall statistically significant model, $F(59,59940)=6535.47, p < 0.0001$. As can be seen in Table XIV, there were statistically

significant main effects and interactions, including the statistically significant triple interaction $h^2*VTF*STV$.

TABLE XIII. DESCRIPTIVE STATISTICS FOR CFI AVERAGED OVER 1000 REPLICATIONS

VTF	STV	h^2			
		High	Moderate	Low	Mixed
4:1	2:1	0.93 28.90	0.83 6.10	0.57 2.30	0.84 5.60
	4:1	0.98 99.20	0.95 61.90	0.85 25.70	0.96 67.00
	8:1	0.99 100.00	0.98 97.90	0.95 55.40	0.98 98.90
	16:1	0.99 100.00	0.99 100.00	0.98 85.80	0.99 100.00
	32:1	0.99 100.00	0.99 100.00	0.99 99.60	0.99 100.00
7:1	2:1	0.94 30.90	0.87 0.70	0.62 0.00	0.91 6.60
	4:1	0.98 100.00	0.97 91.20	0.89 18.50	0.98 99.40
	8:1	0.99 100.00	0.99 100.00	0.96 67.40	0.99 100.00
	16:1	0.99 100.00	0.99 100.00	0.98 99.20	0.99 100.00
	32:1	0.99 100.00	0.99 100.00	0.99 100.00	0.99 100.00
10:1	2:1	0.94 41.00	0.88 0.00	0.69 0.00	0.92 3.00
	4:1	0.98 100.00	0.97 98.70	0.91 18.90	0.98 100.00
	8:1	0.99 100.00	0.99 100.00	0.97 96.20	0.99 100.00
	16:1	0.99 100.00	0.99 100.00	0.99 100.00	0.99 100.00
	32:1	0.99 100.00	0.99 100.00	0.99 100.00	0.99 100.00

TABLE XIV. THREE-WAY ANOVA FOR CFI BY CONDITIONS

Conditions	df	F	p-value
h^2	3	21591.5	<0.0001
VTF	2	2208.58	<0.0001
h^2*VTF	6	359.17	<0.0001
STV	4	51446.3	<0.0001
h^2*STV	12	8444.73	<0.0001
VTF*STV	8	626.81	<0.0001
$h^2*VTF*STV$	24	87.90	<0.0001

Visual examination of the mean CFI values reveals a differential increase in mean CFI within the levels of STV as a function of h^2 magnitude. In the h^2 = high condition, the mean CFI values evidence minimal gains even between 2:1 and 4:1 STV levels. Within the mixed and moderate h^2 conditions, the mean CFI values show asymptotic gains after $STV>4:1$. However, in the low h^2 condition, the mean CFI values were markedly lower in the 2:1 and 4:1 STV levels, only showing asymptotic values when $STV>8:1$ (see Figure 6).

Post hoc analysis of the 3-way interaction $h^2*VTF*STV$ first focused on the simple effect interaction h^2*VTF after blocking on STV, specifically the test of 2-way interaction h^2*VTF at each level of STV. The results revealed statistically significant 2-way interactions at all STV levels: $STV=(2:1)$,

$F(6,11988)=168.36, p<0.0001$; $STV=(4:1)$, $F(6,11988)=93.29, p<0.0001$; $STV=(8:1)$, $F(6,11988)=109.95, p<0.0001$; $STV=(16:1)$, $F(6,11988)=96.87, p<0.0001$; and $STV=(32:1)$, $F(6,11988)=92.47, p<0.0001$. Further analysis of the four by three 2-way interactions focused on the simple-simple effects of STV blocking on the $STV*h^2$ interaction. Specifically, the analysis examined the differences in mean TLI values among VTF levels for each $STV*h^2$ interaction. Table XV presents these findings.

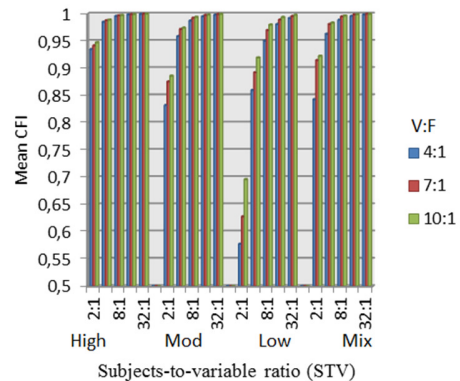


Fig. 6. Comparative-fit index mean values for the interaction between STV and VTF ratios at different levels of communalities.

TABLE XV. SIMPLE-SIMPLE EFFECT OF $STV*VTF*h^2$ SLICED BY $STV*h^2$

h^2	STV	df	F	p-value	VTF		
					4:1vs7:1	4:1vs10:1	7:1vs10:1
Low	2	2	354	<0.001	<0.0001	<0.0001	<0.0001
	4	2	163	<0.001	<0.0001	<0.0001	<0.0001
	8	2	179	<0.001	<0.0001	<0.0001	<0.0001
	16	2	156	<0.001	<0.0001	<0.0001	<0.0001
	32	2	150	<0.001	<0.0001	<0.0001	<0.0001
Moderate	2	2	355	<0.001	<0.0001	<0.0001	<0.0001
	4	2	142	<0.001	<0.0001	<0.0001	0.0345
	8	2	131	<0.001	<0.0001	<0.0001	0.0012
	16	2	134	<0.001	<0.0001	<0.0001	0.0105
	32	2	138	<0.001	<0.0001	<0.0001	0.0127
High	2	2	95.3	<0.001	<0.0001	<0.0001	<0.0001
	4	2	50.4	<0.001	<0.0001	<0.0001	0.0327
	8	2	59.6	<0.001	<0.0001	<0.0001	0.0015
	16	2	81.8	<0.001	<0.0001	<0.0001	0.0004
	32	2	87.1	<0.001	<0.0001	<0.0001	<0.0001
Mixed	2	2	1151	<0.001	<0.0001	<0.0001	<0.0001
	4	2	334	<0.001	<0.0001	<0.0001	0.0120
	8	2	251	<0.001	<0.0001	<0.0001	0.0024
	16	2	216	<0.001	<0.0001	<0.0001	0.0164
	32	2	206	<0.001	<0.0001	<0.0001	0.0037

VIII. DISCUSSION

The study findings refuted some of the guidelines found in the literature, e.g. authors in [18] reported that sample size was not an important factor in determining model stability, and authors in [4] reported that the subject-to-variable ratio should be no lower than 5. The results of the current study revealed that sample size did have a strong effect on both stability and precision of the simulated models. For instance, when the VTF ratio was 4:1, the mean values related to data-model fit indices

were adequate at STV ratio $\geq 4:1$. However, looking at the frequency of rejections based on conventional thresholds over the 1000 replications depicted a different conclusion. The percentage of stable (invariant) models ranged from 77% at 4:1 STV to 91% at 32:1 STV clearly indicating that larger STV ratios are related to higher stability levels with a model. These findings validated the authors in [42], who reported that the percentage of invariance tests varied depending on the sample size of the group. The findings of the current study do agree with those in [43], where in some models a STV of 30:1 was needed to produce a stable model and minimize the amount of misfit.

The study findings also contradicted some previous research on the effect of VTF ratio on the precision and stability of factor solutions. Authors in [18] concluded that the VTF ratio was important for factor stability with more indicator variables per factor yielding more stable result. In the current study, there was a trend in the findings over all RQ1 analyses that suggests that data-model fit diminished as VTF increased. Most probably this is a result of the increasing complexity of the measurement models, i.e. the number of paths. For example, in the 10:1 VTF condition there are 100 estimated paths, whereas in the 4:1 condition there are 40 estimated paths. The accumulation of many small deviations from the population correlation matrix impacted negatively the global fit statistics more in the high VTF conditions than in the low VTF conditions. The overall model fit indices TLI and CFI all varied as functions of VTF, STV, h^2 , and their interaction. Chi-square value and overall model fit indices GFI, and RMSEA only varied as functions of VTF, STV, and their interactions.

IX. CONCLUSION

This study provided empirical evaluation of the influence of ED and SD on WSV in terms of model precision relative to a known factor structure via Monte Carlo simulations. The results revealed that the means of χ^2 and RMSEA values significantly decrease as STV increases and increase as VTF ratio decreased. However, the examination suggests significant increase in mean GFI values as STV increased, and decreasing mean GFI values as VTF ratio increased.

Overall, when the VTF is 4:1 an STV ratio of 16:1 or above is required to show precision in factor solution and stability of the model as indicated by four fit indices. When the VTF ratio is 7:1 or 10:1, an STV ratio of 32:1 or above is required to show precision in factor solution and the stability of the model. If a researcher is interested in minimizing misfits or meeting more than four overall model fit indices criteria, STV and VTF ratios more than 32:1 and 10:1 respectively would be necessary in order to gain a precise and stable model.

REFERENCES

- [1] L. Crocker and J. Algina, *Introduction to Classical and Modern Test Theory*. Bristol, UK: ERIC, 1986.
- [2] 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. 10, 2005, Art. no. 7.
- [3] T. A. Brown, *Confirmatory Factor Analysis for Applied Research*, Second edition. London, UK: The Guilford Press, 2015.
- [4] F. B. Bryant and P. R. Yarnold, "Principal-components analysis and exploratory and confirmatory factor analysis," in *Reading and understanding multivariate statistics*, Washington, DC, US: American Psychological Association, 1995, pp. 99–136.
- [5] I. P. Cobham, "The effects of subject-to-variable ratio, measurement scale, and number of factors on the stability of the factor model," Ph.D. dissertation, University of Miami, FL, USA, 1999.
- [6] L. Hatcher, *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling*, 1st edition. North Carolina, USA: SAS Publishing, 1994.
- [7] K. 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," Ph.D. dissertation, University of South Florida, FL, USA, 2013.
- [8] L. R. Goldberg, "The development of markers for the Big-Five factor structure," *Psychological Assessment*, vol. 4, no. 1, pp. 26–42, 1992, <https://doi.org/10.1037/1040-3590.4.1.26>.
- [9] D. L. Bandalos and P. Gagne, "Simulation methods in structural equation modeling," in *Handbook of structural equation modeling*, New York, USA: The Guilford Press, 2012, pp. 92–108.
- [10] H.-M. Lin, M.-H. Lee, J.-C. Liang, H.-Y. Chang, P. Huang, and C.-C. Tsai, "A review of using partial least square structural equation modeling in e-learning research," *British Journal of Educational Technology*, vol. 51, no. 4, pp. 1354–1372, 2020, <https://doi.org/10.1111/bjet.12890>.
- [11] L. Fabrigar, D. Wegener, R. MacCallum, and E. Strahan, "Evaluating the Use of Exploratory Factor Analysis in Psychological Research," *Psychological Methods*, vol. 4, no. 3, pp. 272–299, Sep. 1999.
- [12] Y. A. Wang and M. Rhemtulla, "Power Analysis for Parameter Estimation in Structural Equation Modeling: A Discussion and Tutorial," in *Advances in Methods and Practices in Psychological Science*, California, USA: University of California, 2020.
- [13] M. J. Allen, *Introduction to measurement theory*. CA, USA: Cole Publishing, 1979.
- [14] 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, May 2020, <https://doi.org/10.1080/10705511.2019.1647785>.
- [15] Y. Rosseel, "Small sample solutions for structural equation modeling," in *Small Sample Size Solutions*, London, UK: Routledge, 2020, pp. 226–238.
- [16] S. C. Smid, D. McNeish, M. Miocevic, 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, Jan. 2020, <https://doi.org/10.1080/10705511.2019.1577140>.
- [17] 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, Jul. 2019, <https://doi.org/10.1080/10705511.2018.1545232>.
- [18] E. Guadagnoli and W. F. Velicer, "Relation of sample size to the stability of component patterns," *Psychological Bulletin*, vol. 103, no. 2, pp. 265–275, 1988, <https://doi.org/10.1037/0033-2909.103.2.265>.
- [19] J. C. Westland, "Lower bounds on sample size in structural equation modeling," *Electronic Commerce Research and Applications*, vol. 9, no. 6, pp. 476–487, Nov. 2010, <https://doi.org/10.1016/j.elerap.2010.07.003>.
- [20] P. F. M. Bullon, "Failing to replicate: Hypothesis testing as a crucial key to make direct replications more credible and predictable," Ph.D. dissertation, Western Michigan University, MC, USA, 2015.
- [21] O. P. John and S. Srivastava, "The Big Five Trait taxonomy: History, measurement, and theoretical perspectives," in *Handbook of personality: Theory and research*, New York, NY, USA: Guilford Press, 1999, pp. 102–138.
- [22] J. Wang and X. Wang, *Structural Equation Modeling: Applications Using Mplus*, 2nd edition. New York, NY, USA: Wiley, 2019.

- [23] D. L. Moody, "The method evaluation model: a theoretical model for validating information systems design methods," in *11th European Conference on Information Systems*, Naples, Italy, Jun. 2003, pp. 1–17, Accessed: Jan. 21, 2021.
- [24] 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, <https://doi.org/10.48084/etasr.3343>.
- [25] 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, <https://doi.org/10.48084/etasr.3927>.
- [26] R. O. Mueller and G. R. Hancock, "Structural Equation Modeling," in *The Reviewer's Guide to Quantitative Methods in the Social Sciences*, London, UK: Routledge, 2018, pp. 445–456.
- [27] M. K. Cain and Z. Zhang, "Fit for a Bayesian: An Evaluation of PPP and DIC for Structural Equation Modeling," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 26, no. 1, pp. 39–50, Jan. 2019, <https://doi.org/10.1080/10705511.2018.1490648>.
- [28] J. S. Tanaka, "How Big Is Big Enough?: Sample Size and Goodness of Fit in Structural Equation Models with Latent Variables," *Child Development*, vol. 58, no. 1, pp. 134–146, 1987, <https://doi.org/10.2307/1130296>.
- [29] K. H. Kim and P. M. Bentler, "Data Modeling: Structural Equation Modeling," in *Handbook of complementary methods in education research*, J. L. Green, G. Camilli, and P. B. Elmore, Eds. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2006, pp. 161–175.
- [30] B. O. Muthen and A. Satorra, "Complex Sample Data in Structural Equation Modeling," *Sociological Methodology*, vol. 25, pp. 267–316, 1995, <https://doi.org/10.2307/271070>.
- [31] K. Y. Hogarty, C. V. Hines, J. D. Kromrey, J. M. Ferron, and K. R. Mumford, "The Quality of Factor Solutions in Exploratory Factor Analysis: The Influence of Sample Size, Communalities, and Overdetermination," *Educational and Psychological Measurement*, vol. 65, no. 2, pp. 202–226, Apr. 2005, <https://doi.org/10.1177/0013164404267287>.
- [32] A. J. S. Morin, N. D. Myers, and S. Lee, "Modern Factor Analytic Techniques," in *Handbook of Sport Psychology*, New York, NY, USA: John Wiley & Sons, 2020, pp. 1044–1073.
- [33] B. S. Everitt, "Multivariate Analysis: The Need for Data, and other Problems," *The British Journal of Psychiatry*, vol. 126, no. 3, pp. 237–240, Mar. 1975, <https://doi.org/10.1192/bjp.126.3.237>.
- [34] G. D. Garson, *Factor Analysis*. NC, USA: Statistical Associates Publishers, 2013.
- [35] 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, Jul. 2020, <https://doi.org/10.1080/09585192.2017.1416655>.
- [36] B. Thompson, "Ten commandments of structural equation modeling," in *Reading and understanding MORE multivariate statistics*, Washington, DC, US: American Psychological Association, 2000, pp. 261–283.
- [37] 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, Jun. 2006, <https://doi.org/10.1177/0013164405282485>.
- [38] A. H. Soomro, A. S. Larik, M. A. Mahar, A. A. Sahito, and I. A. Sohu, "Simulation-based Analysis of a Dynamic Voltage Restorer under Different Voltage Sags with the Utilization of a PI Controller," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5889–5895, Aug. 2020, <https://doi.org/10.48084/etasr.3524>.
- [39] D. Almaleki, "Empirical Evaluation of Different Features of Design in Confirmatory Factor Analysis," Ph.D. dissertation, Western Michigan University, MI, USA, 2016.
- [40] C. Wardley, E. Applegate, A. Almaleki, and J. V. 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, Mar. 2019, <https://doi.org/10.1097/JPA.0000000000000241>.
- [41] 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, Mar. 2016, <https://doi.org/10.1097/JPA.0000000000000060>.
- [42] 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, Oct. 2007, <https://doi.org/10.1080/10705510701575461>.
- [43] S. V. Paunonen and D. N. Jackson, "What Is Beyond the Big Five? Plenty!," *Journal of Personality*, vol. 68, no. 5, pp. 821–835, 2000, <https://doi.org/10.1111/1467-6494.00117>.