sciendo

National University of Political Studies and Public Administration
Faculty of Management

Competing Confirmatory Factor Analysis Models in Management Research: Bifactor Modeling of the Employee Work Assessment Tool

Theophilus Ehidiamen OAMEN[1]
[1] Texila American University, 641038, Georgetown, GY; oamentheo.yahoo.com

Received: April 25, 2024
Revised: May 9, 2024
Accepted: May 19, 2024
Published: June 16, 2024

Abstract: Management researchers often use structural equation modeling to analyze data from questionnaire-based instruments. Usually, confirmatory factor analysis (CFA) is applied to confirm the hypothesized or theorized factor structure of the instrument. Most authors adopt a single CFA model without comparing it against other potentially valid models (general factor, correlated factor model, second-order hierarchical model, and bifactor model). Hence, the dimensionality and reliability of constructs using bifactor modeling to validate latent scores are often ignored. Also, this gap is widened by no unanimous agreement on the use of post hoc modification of CFA models to support fit to the data in covariance-based structural equation modeling (CB-SEM). The objective of the study was to explore model fit, dimensionality, and reliability of the Employee Work Assessment Tool (EWAT) using competing CFA models. The study used a published dataset on the EWAT instrument to illustrate the assessment of the dimensionality and model-based reliability of the tool using CB-SEM. Results showed that CFA statistics of the bifactor model were most adequate for the instrument ($\chi2$=70.053, df=19, RMSEA=0.082 [90% confidence interval; 0.062, 0.103], SRMR=0.036, CFI=0.963). The bifactor model ancillary measures supported the unidimensional structure of EWAT with justification for the use of total scores. The study concludes that the instrument is best described and applied as a unidimensional construct, and therefore, a single score can be used to rate employees' perceptions of their work conditions. The study presents both practical implications for management researchers and simplified reporting for bifactor modelling.

Keywords: bifactor model; general factor model; correlated factors model; hierarchical second order model; confirmatory factor analysis; dimensionality; structural equation modelling.

Introduction

The current trend in statistical analytical procedures in the management sciences commonly deploys the use of structural equation modeling in the form of partial least squares structural equation modeling (PLS-SEM) and covariance-based structural equation modeling (CB-SEM) (Hair et al., 2017). These methods have significantly improved the analysis of data including data obtained from self-reported questionnaires. In management research, researchers often use questionnaires as their instrument of choice to measure constructs, phenomena, and issues of interest. Constructs are derived or forged from the conceptualizations of the researcher based on theory to capture a desired meaning, research context, and/or goal (Oamen, 2023; Schuberth, 2021). They are not directly observable and therefore, are measured using observable variables or indicators. Usually, the core meanings of the constructs are embedded in indicators or measurement items that convey their meaning (Hair et al., 2017).

In psychometrics, they are also referred to as common factors since the measurement items or indicators used to measure them are derived based on the conceptualization of the researcher. In developing an instrument, the researcher may desire to test if the indicators measure the constructs they are supposed to measure, thereby informing the use of confirmatory factor analysis (CFA). CFA modeling is used to determine the construct validity of the factor structure of the instrument, CFA modeling is easily

How to cite
Oamen, T. E. (2024). Competing Confirmatory Factor Analysis Models in Management Research: Bifactor Modeling of the Employee Work Assessment Tool. *Management Dynamics in the Knowledge Economy, 12*(2), 101-115. DOI 10.2478/mdke-2024-0007
ISSN: 2392-8042 (online)
www.managementdynamics.ro
https://content.sciendo.com/view/journals/mdke/mdke-overview.xml

achieved using both CB-SEM and PLS-SEM for convergent and discriminant validity of the constructs used (Hair, et al., 2019; Henseler et al., 2015). Furthermore, Sarstedt et al. (2024) argued that reproducibility of results from research models developed using structural equation modeling and reliability of such findings by researchers is critical to avoid erroneous inferences and conclusions by management practitioners and policymakers. Hence, they recommended that when a choice is to be made among alternate or competing models, such evaluation should be critical, methodological and properly reported to avoid ambiguity in interpretation of results.

Presently, a large number of published studies utilize correlated and hierarchical higher-order CFA models to evaluate the measurement qualities and attributes of research questionnaires using structural equation modeling (Hair et al., 2019; Henseler et al., 2015; Oamen, 2023). CFA studies are statistical procedures to explore the underlying theoretical structures for defined measurement items presumed to measure one or more dimensions (Dunn & McCray, 2020; Gegenfurtner, 2022). CFA modeling achieves two main objectives; first, it enables the researcher to identify observed variables that do not sufficiently measure the construct. Secondly, it gives the researcher information on the probable factor structure of the model showing the sets of indicators explaining each construct (Gegenfurtner, 2022).

CFA models in management studies may be presented as a general factor, correlated factors models, or hierarchical or second-order models which present some inherent limitations. Whilst the general factor is limited to all indicators measuring a singular construct; the correlated factors model measures multiple constructs and excludes the possibility of a single-dimensional construct; the second order or hierarchical model does not support the independent assessment of specific and general factors or constructs (Savahl et al., 2023). The bifactor model resolves and addresses the limitations of the other CFA models. This is due to its ability to simultaneously estimate both the general and specific factors or constructs (Gegenfurtner, 2022; Torres-Vallejos et al., 2021; Zhang, et al., 2021). In other words, bifactor models give justification for a general factor and hence make a case for the use of total scores, and/or justify the use of independent scores for multiple factors explaining different aspects of general factor (Byllesby & Palmieri, 2023; Canivez, 2016; Reise, 2012; Torres-Vallejos et al., 2021).

In management research, during the analysis of self-reported data, researchers are faced with the question of whether a group of indicators can be measured by a single construct (that is, unidimensional), and/or by multiple constructs (multidimensional). This implies that a total summated score in the case of a unidimensional construct or specific scores of facets of a construct is applied with confidence. This necessitates the evaluation of competing CFA models such as the general factor, correlated factor model, second-order hierarchical model, and bifactor model. However, according to Canivez (2016) and Reise (2012), researchers may be faced with the uncertainty of which CFA model adequately determines the dimensions (unidimensional, multi-dimensional, or in-between) for a measurement scale. This justifies the use of competing CFA models, in particular, the bifactor model to address the issue of dimensionality and reliability (Luo & Al-Harbi, 2016).

The study explored the dimensionality and model reliability of the previously developed employee work assessment tool using competing models (a general factor model, a correlated two-factor model, a second-order model, and a bifactor model) using CB-SEM.

Vol.12 (2024) no.2, pp.101-115; DOI 10.2478/mdke-2024-0007

Literature review

Review of competing CFA models

Constructs or factors are measured by indicator items and may be unidimensional or multidimensional in nature. According to Reise, Moore, and Haviland (2010), CFA attempts to define the factor structure of an instrument which more often than not, is multi-dimensional. This assertion gives reasoning to the relevance of determining the actual unidimensionality, or perhaps multidimensionality of the instrument by exploring competing CFA models. There are four commonly used CFA models-general factor, correlated factors model, second-order hierarchical model, and bifactor model. A general factor model is made up of one general construct and the indicator items measuring it. In other words, all the indicators correlate and converge on a single factor; hence, it is unidimensional in nature with straightforward interpretation (Canivez, 2016; Gegenfurtner, 2022) (Figure 1). In the case of the correlated factors model, the factors or sub-constructs are not independent, and hence correlation exists between the multiple factors, thus, interpretation from the correlated factors model is not straightforward due to shared variance between the factors (Canivez, 2016; Reise, 2012; Tabachnick & Fidell, 2007) (Figure 2). Second-order factor models are derived from the correlated factors model as a higher-order or hierarchical factor model in which the correlated factors form a general or common factor known as a second-order factor (Savahl et al., 2023).

The second-order factor formed from the first-order factors is such that its (that is, second-order factor) influence on the indicators is implied and indirect (Canivez, 2014; Gignac, 2008) (Figure 3). However, a major shortcoming of the second-order or hierarchical model is that the second-order construct is indirectly inferred or measured by the indicator items or subscales. The fallout is that the relevance or meaning inferred from the first-order construct compared to the general and correlated factors models which are directly linked (Canivez, 2016). Another drawback is that the constructs in the correlated factors model and second-order hierarchical models, tend to show high intercorrelations (greater than 0.3). This raises the question of whether the constructs are actually measuring different subdomains or merely measuring one general construct (Tabachnick & Fidell, 2007).

Finally, the bifactor model is made up of a general factor directly linked to the indicator items and uncorrelated group factors (specific group factors) which are directly linked to the indicators related to them. That is, each indicator has a path from its group factor and one path from the general factor. The parsimonious and less ambiguous nature is such that the general factor and specific group factors are linked to the indicators (Canivez, 2016; Chen, Hayes, Carver, Laurenceau, & Zhang, 2012; Gignac, 2008) (Figure 4). From the literature, bifactor analysis is recommended to provide needed guidance on the dimensionality of a measurement model. Bifactor models provide evidence of the presence or absence of unidimensionality of the instrument as well as to determine if the specific factors or constructs are measured by their indicators (Canivez, 2016; Gignac, 2008; Torres-Vallejos et al., 2021). Hence, for a holistic analysis of an instrument, it is advisable to complement bifactor analysis using other unidimensional and multidimensional CFA models (Reise, Morizot, & Hays, 2007; Zhang et al., 2021).

Bifactor models for dimensionality and reliability estimation of instruments

Bifactor models are acknowledged to be the best strategy to identify and determine the feasibility of an instrument's dimensionality- whether unidimensional or multidimensional, that is, determine if the measurement items or indicators are measuring a single construct or specific sub-constructs or factors (Canivez, 2016; Gignac, 2008; Luo & Al-Harbi, 2016; Reise, 2012; Ventura-Leon et al., 2021; Zhang et al., 2021). Thereby, the feasibility of using total scores of all indicator items in the case of unidimensional models, and specific factor scores for multidimensional constructs is attained (Canivez, 2016; Reise, 2012). Bifactor models are unique in three main ways; 1) they are used to assess if a measurement model uses either general factors or specific factors on a group of indicators. 2) Bifactor models address the issue of scoring into total scores representing a general factor and/or scores of the specific factors, hence the requirement to confirm dimensionality through statistical modeling with structural equation modeling (Byllesby & Palmieri 2023; Reise, 2012; Ventura-Leon et al., 2021). 3) Ancillary bifactor measures are used to provide model-based reliability for good-of-fit statistics to the CFA measurement models. In other words, the goodness of fit measures such as root mean squared error of approximation-RMSEA, comparative fit index-CFI, and standardized root mean residual-SRMR from CFA, are not sufficient to establish or confirm dimensionality because good-of-fit tends to favor bifactor models compared to other competing models (Bonifay, Lane, & Reise, 2017; Dueber, 2017; Gignac, 2008).

Ancillary or supporting indices are used because the conventional Cronbach alpha test of internal reliability gives poor estimations of specific and general factors in a bifactor model (Chen et al., 2012). Ancillary bifactor measures cover two broad areas; 1) Dimensionality measures (explained common variance-ECV; Individual explained common variance-IECV, percentage of uncontaminated correlations-PUC, Average relative parameter bias-ARPB), and 2) model reliability (omega coefficient-ω, omega coefficient for specific factors-ωS, omega hierarchical-ωH, omega hierarchical for subscale-ωHS, percentage of reliable variance-PRV, construct replicability-H, and factor determinacy-FD) (Canivez, 2016; Reise, 2012; Rodriguez, Reise, & Haviland, 2016).

Critique of EWAT instrument

The current measurement model of the EWAT research instrument is a correlated factors model with two specific constructs for 9 measurement indicators or items. The fit of the model was obtained using modification indices which informed the correlation of error terms within constructs based on theoretically guided relationships (Oamen, 2021; Oamen, Idiake, & Omorenuwa, 2022; Oamen, Omorenuwa, & Moshood, 2022). This post hoc adjustment was deemed necessary because of the poor fit indices of the initial measurement model. A comparison of the studies by Oamen (2021), and Oamen et al. (2022a) showed that the model fit characteristics CFI=0.990, RMSEA=0.040 vs. CFI=0.953, RMSEA=0.084, gave mixed results in terms of difference between both post hoc correlated factor models (ΔCFI=0.006, ΔRMSEA=0.044) based on the recommendations of Cheung and Rensvold (2002; CFI > .01) and Chen (2007; RMSEA >0.015). Although the modifications were based on sound theoretical justifications or underpinnings (Cole, Ciesla, & Steiger, 2007; Collier, 2020), researchers (Landis, Edwards, & Cortina, 2010; Tomarken & Waller, 2003) argue that the post hoc practice of modification indices tend to alter the true inherent nature of the model, fails to address the core reasons for misfit, and the practice implies the use of chance to achieve desired model fit. This argument necessitated the need for a re-evaluation of the model in comparison to other CFA models as suggested by Reise et al. (2007). However, Landis et al. (2010) posited that a special case can be made for the use of modification indices for models where constructs are measured longitudinally.

A major focus of the study is that it attempts to answer the question of whether the EWAT instrument can be empirically tested to be valid and reliable using bifactor statistical modeling as: a) a single (unidimensional) construct, hence warranting the use of latent score for the overall construct of employee perception of the work conditions; b) several constructs or multidimensional, for example, a correlated two-factor model or hierarchical second-order model with specific factors D1 and D2 which independently measure facets of the instrument. Therefore, individual scores for D1 and D2 are considered applicable and thus, a single sum score representing D1 and D2 is not appropriate; c) a bifactor model that supports possible dimensionality of the general factor, and specific factors associated with the measurement items and d) to confirm the reliability of the empirically selected dimension using ancillary bifactor indices (see Figures 1, 2, 3, 4).

The overarching principles behind the dimensionality debate are essentially two-pronged:

1. The dimensionality of a research instrument is analogous to putting a square peg in a square hole and NOT a round peg in a square hole. In other words, instruments must be appropriately dimensioned so that the scores derived from them have meaningful relevance to managerial decisions
2. Ensuring the dimensionality of an instrument is critical to guarantee the applicability of the scores derived from empirical research instruments does not misguide decision-makers or policymakers. For instance, treating disability score as a single dimension or factor, when it could have been measured by several scores covering other facets of the construct.

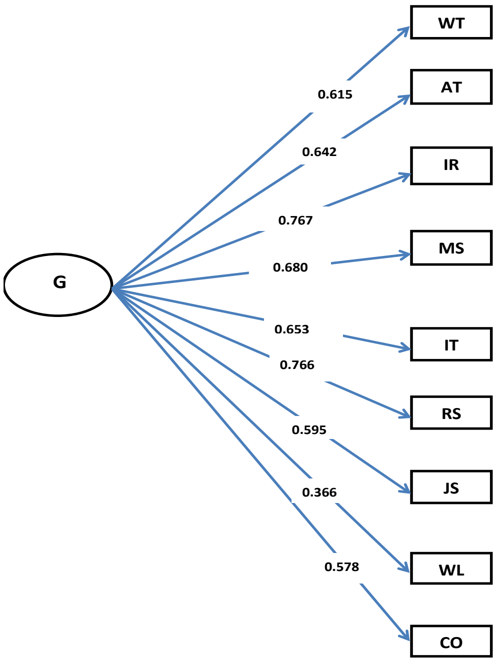

Figure 1. General factor model with standardized coefficients for the Employee Work Assessment Tool
Source: own processing

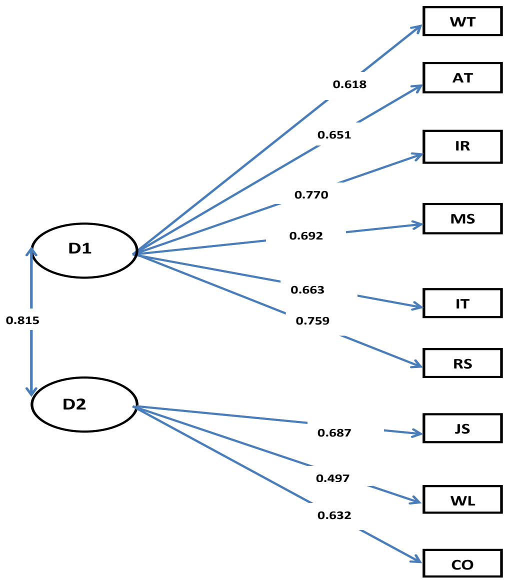

Figure 2. Correlated factors model with standardized coefficients for the Employee Work Assessment Tool
Source: own processing

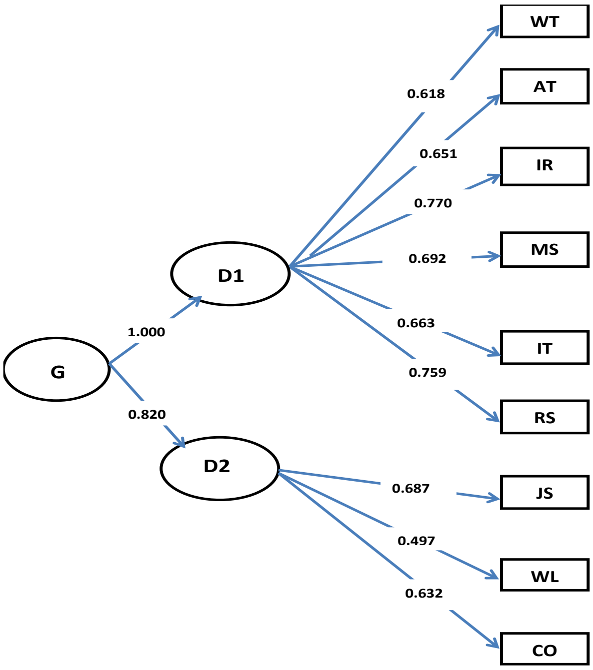

Figure 3. Hierarchical or second-order factor model with standardized coefficients for the Employee Work Assessment Tool
Source: own processing

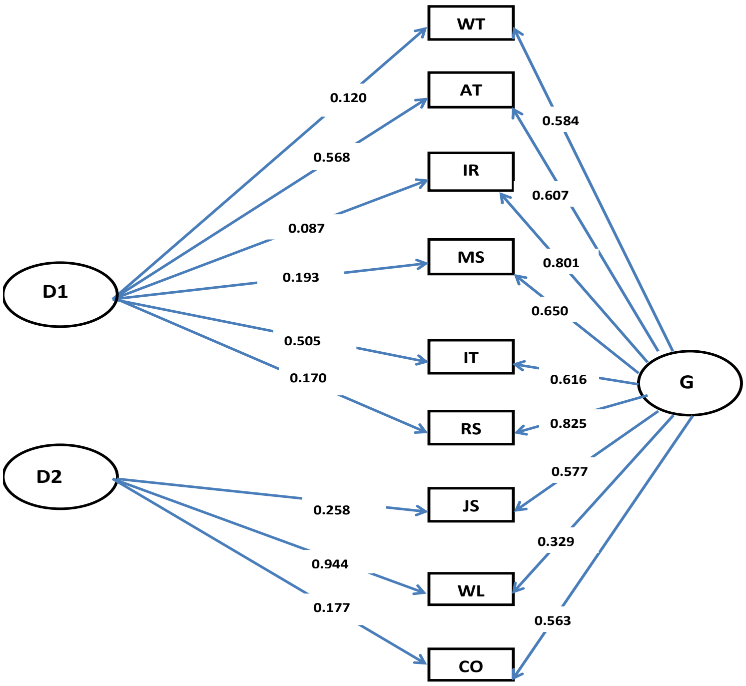

Figure 4. Bifactor measurement model with standardized coefficients for the Employee Work Assessment Tool
Source: own processing

Key to Figures 1 to 4:

WT-availability of work tools: **AT**-regular training: **IR**-availability incentive and remuneration scheme: **IT**-impactful training: **RS**-attractive reward system: **MS**- marketing support: **JS**-Job security: **WL**-work-life balance: **CO**-career opportunities: **D1**- first specific factor: **D2**-second specific factor, and **G**-general factor

Research methods

Study design and participants

To achieve the goal of comparing the competing CFA models, a dataset from recently published studies that developed the parsimonious self-reported employee work assessment tool (EWAT; Oamen, 2021; Oamen et al., 2022a; Oamen et al., 2022b) was used. The goal of the research instrument is to measure employee's perception of their working conditions. The short-form survey instrument has a total of 9 measurement items and is measured on a 3-point ordinal Likert scale of 1=poor, 3=fair, and 5 for good. From the initial construction and validation of the study, a correlated two-factor model was obtained; D1 (Work conditions with 6 indicators) and D2 (Implied working conditions with 3 items) with acceptable model fit after correlating error terms based on suggestions from modification indices output and theory. Construct validity measures-convergent and discriminant validity values of the instrument were within acceptable limits (Oamen, 2021; Oamen et al., 2022a). A total sample of 401 sales and marketing staff from over 20 pharmaceutical companies in Nigeria were used for the analysis. CB-SEM using analysis of moment structures (Arbuckle, 2016) with the maximum likelihood estimation method was used to develop and evaluate the competing CFA models. In analyzing data using CB-SEM, the global fit statistic-X^2 difference test tends to yield significant p-values because the method is sensitive to large sample sizes above 200. Therefore, absolute fit indices such as root mean error of approximation (RMSEA), Standardized root mean residual (SRMR), and comparative fit index (CFI) were used (Vandenberg, 2006). Ancillary or supporting bifactor measures of dimensionality and model-based reliability were computed using the online software tool called BifactorCalc developed by Ventura-Leon et al. (2021).

Measures of EWAT instrument

The measurement items were availability of work tools (WT), regular training (AT), availability incentive and remuneration scheme (IR), impactful training (IT), attractive reward system (RS), and marketing support (MS) forming the specific construct or factor (Work factors-D1). Also, Job security (JS), work-life balance (WL), and career opportunities (CO) formed the second specific factor (Implied factors-D2) (Oamen, 2021; Oamen et al., 2022a; Oamen et al., 2022b). Based on the objective of the study, four CFA models were developed: a general factor model, a correlated two-factor model, a second-order model, and a bifactor model according to the recommendation of Reise et al. (2007).

Dimensionality measures of the bifactor model

To determine whether the adopted model (best fitting) is unidimensional, multidimensional, or otherwise, ancillary measures from BifactorCalc (Ventura-Leon et al., 2021) can be generated. Dimensionality is determined using explained common variance-ECV (>0.85 threshold), if ECV is below 0.85 implies that the general factor explained less than 85% of item ECV. ECV above 0.85 suggests adequate unidimensionality of the instrument to necessitate a single-factor model (Stucky & Edelen, 2015). Individual extracted common variance-IECV of indicator items (IECV<0.5), If IECV is below 0.5, it means an item measures the specific factor better than the general factor. IECV is an index of unidimensionality (Reise et al., 2010). Absolute relative parameter biases-ARPB (<10-15% threshold) when above 10% implies that the instrument is multidimensional based on the comparison of the general factor from the bifactor model and the unidimensional model. While ARPB values below 10% suggest adequate unidimensionality of the instrument. Finally, the percentage of uncontaminated correlations (PUC) indicates if the structural coefficients are not contaminated by multicollinearity. Instruments with more specific factors and fewer indicator items tend to have high PUC values (Reise, Bonifay, & Haviland, 2013). PUC values above 0.8 are strongly indicative of unidimensionality. PUC values below 0.8

require ECV>0.6 to substantiate unidimensionality. Definition of indices of dimensionality are defined and explained mathematically in detail by Canivez (2016), Rodriguez et al. (2016), and Ventura-Leon et al. (2021).

Reliability measures of the bifactor model

In bifactor models, it is essential to determine the reliability of the general factor dimension and the specific factors dimension/s without the influence of others (that is, either the general factor or the specific factors) (Brunner, Nagy, & Wilhelm, 2012; Canivez, 2016; Zinbarg et al., 2006). In other words, they provide evidence that the total and subscale scores of a research instrument truly represent the target constructs of focus. Furthermore, model-based reliability measures support the use of structural equation modeling to generate latent cores, or summated scores from all the items measuring the construct can be used. Several measures are typically used namely Omega (ω) measures the reliability of total and subscale scores.

Omega hierarchical (ωH) with a range set at 0.5 to 0.75 is the reliability of a specific target construct with others (other constructs and general factors) removed (Canivez, 2016; Reise, 2012; Reise et al., 2013). ωH is the measure of the reliability of the total score. Omega Hierarchical subscale (ωHS) measures the reliability of the subscale scores. Generally, when the value of ωHS is < 0.5, it indicates that the influence of the general factor on the variance of the specific factor is large, hence validating unidimensionality. Construct replicability (H) is a measure of the adequacy of the indicator items measuring or defining a latent construct. A value of H>0.7 is an index of adequacy of the items measuring the construct (Dominguez-Lara, 2016). Factor determinacy (FD) indicates the assurance of obtaining similar factor scores from a construct. Acceptable FD values range from 0.8 to 0.9 and support the estimation of the general factor score (Grice, 2001; Gorsuch, 1983). The percentage of reliable variance (PRV) with the value set at >0.75 is an indication of the variance caused by the general factor (Hammer et al., 2018). Again, the definition of reliability indices are defined and explained mathematically in detail by Canivez (2016) and Ventura-Leon et al., (2021).

Results

As shown in Table 1, the interpretation of the CFA statistics of the competing CFA models is straightforward; only the bifactor model had acceptable fit measures (RMSEA<0.08, SRMR<0.06, CFI<0.95) based on the recommendations of Hu and Bentler (1999) and Kline (2016). Interestingly, the fit of the correlated factors model was identical to that of the second-order hierarchical model. This is majorly due to the strong correlation (β=0.815) between the two specific factors D1 and D2 (see Figures 1 and 3). This high correlation coefficient among the oblique or correlated factors D1 and D2 informs the development of a second-order construct or general factor found in the second-order hierarchical model (Canivez, 2016; Thompson, 2004).

Table 1. Analysis of fit statistics of competing CFA models

CFA Model	χ2	df	RMSEA [90% C.I]	SRMR	CFI
Unidimensional factor	252.267	27	0.144 [0.128, 0.161]	0.0732	0.835
Correlated Factors	228.317	26	0.139 [0.123, 0.156]	0.0691	0.852
Second Order	228.317	26	0.139 [0.123, 0.156]	0.0691	0.851
Bifactor	70.053	19	0.082 [0.062, 0.103]	0.0357	0.963

Note: C.I=confidence interval, CFA=confirmatory factor analysis, RMSEA=root mean error of approximation, SRMR=standardized root mean residual, CFI=comparative fit index, χ2=chi-square, df=degrees of freedom

Source: own processing

Therefore, since the bifactor model was chosen based on its most acceptable model fit characteristics, the ancillary bifactor measures for dimensionality and model-based reliability should then be assessed (Brunner et al., 2012; Canivez, 2016; Zinbarg et al., 2006) (Table 2).

Table 2. Dimensionality measures of EWAT

Factor	Measure	Obtained Value	Reference value	Inference
General factor (G)	ECV	0.820	>0.80	Unidimensional
Specific factor (D1)	ECV	0.190	>0.80	Unidimensional
Specific factor (D2)	ECV	0.140	> 0.80	Unidimensional
Indicator items	IECV	6 items>0.85	> 0.85	Unidimensional
General factor (G)	PUC	0.500	> 0.80	*Mixed
G. vs. Unidimensional	ARPB	0.060	< 10-15%	Unidimensional

Note. *Mixed refers to equal (50%) unidimensionality and (50%) multidimensionality, ECV=explained common variance, IECV=individual explained common variance of indicator items, PUC=percentage of uncontaminated variance, ARPB=average relative parameter bias

Source: own processing

Assessment of dimensionality of EWAT

As presented in Table 2, as regards the dimensionality of the EWAT instrument, the ECV for the general factor was 0.82 which implies that the general factor (D) explains 82% of the variance of the items; this is indicative of a tendency towards unidimensionality (ECV>0.8). Furthermore, the ECV for specific factor D1 and specific factor D2 has values of 0.19 and 0.14 respectively. This means that specific factors D1 and D2 respectively account for 19% and 14% of the common variance in the general factor. Similarly, an analysis of the IECV of the measurement items revealed that 6 items out of the 9-WT, IR, MS, RS, WL, and CO were strongly influenced or affected by the general factor (G), that is, they are very strong measures of the general factor (with IECV above 0.85 benchmarks) (Table 3).

Table 3. Model-based reliability measures of EWAT

Factor	Measure	Obtained Value	Reference value
General factor (G)	ω	0.880*	> 0.80
Specific factor (D1)	ω	0.890*	> 0.8
Specific factor (D2)	ωS	0.120	> 0.8
General factor (G)	ωH	0.800*	0.5 to 0.75
Specific factor (D1)	ωHS	0.120	ωHS>0.5 indicates multidimensionality
Specific factor (D2)	ωHS	0.070	ωHS< 0.5 indicates unidimensionality
General factor (G)	PRV	0.910*	> 0.75
Specific factor (D1)	PRV	0.140	> 0.75
Specific factor (D2)	PRV	0.130	> 0.75
General factor (G)	H	0.880*	> 0.7
Specific factor (D1)	H	0.480	>0.7
Specific factor (D2)	H	0.120	> 0.7
General factor (G)	FD	0.920*	> 0.8-0.9
Specific factor (D1)	FD	0.720	> 0.8- 0.9
Specific factor (D2)	FD	0.380	> 0.8- 0.9

Note: ω=Omega, ωS=Omega subscales, ωH=Omega Hierarchical, ωHS=Omega Hierarchical subscales, PRV=Percentage of reliable variance, H=Construct replicability, FD=Factor determinacy

Source: own processing

Also, the PUC value was 0.5 which implies that 50% of the correlations were contaminated by multidimensionality, and thus. 50% of the correlations were explained by the general factor (G). However, based on Reise et al. (2013), when PUC is below 80%, but obtained ECV for the general factor is > 80% and ωH is above 0.7, the presence of multidimensionality (contamination) is not overwhelming to disqualify the interpretation of unidimensionality. The ARPB was 0.06 which suggests that the disparity or difference between the factor loadings of the general factor of the bifactor model, and that of the unidimensional model is only 6% (acceptable range-ARPB=<10-15%). In conclusion, the aforementioned indicators of dimensionality (ECV, IECV, PUC, and ARPB) essentially describe or conceptualize the EWAT instrument as primarily a Unidimensional Instrument regardless of the presence of some multidimensionality. This implies that the total score of all the items can be used to score or rate employee perception of their work environment.

Assessment of reliability of EWAT

As presented in Table 3, the ω of the instrument was 0.88 indicating high reliability of total score while the (ωS) for specific factors (D1) and D2 were 0.89 and 0.12 respectively, which are indicative of acceptable and low reliability respectively. The ωH of the general factor was 0.80 which indicates that the general factor (D) compared to specific factors D1 and D2 account for the majority of variance in the model (baseline $\omega H>0.80$). Hence, affirms the total scores of the instrument as basically unidimensional. In this regard, the ωHS for specific factor D1 is 0.12 and D2 is 0.07 which can be considered low consistency of the specific factors based on the recommendations of Smits et al., 2014 ($\omega HS \geq .30$ is substantial; $0.20 \leq \omega HS < 0.30$ is moderate and $\omega HS < 0.20$ is low). The PRV of 0.91 indicates that 91% of the reliable variance is due to the general factor and only 14% and 13% of the reliable variance to the specific factors (D1-0.14 and D2-0.13). The H coefficient is equal to 0.88 in the general factor, which implies stability while the specific H for D1 (0.48) and D2 (0.11) were less than 0.70, providing evidence in favor of the general factor. Finally, the FD for the general factor is 0.92 and the two specific factors are 0.72 and 0.38 respectively, indicating that only the general factor score should be used for the analysis.

Discussions

This study was based on the recommendations of Canivez (2016) and Reise et al. (2008), comparing at least 2 or more competing models to obtain the best-fitting model, and confirm the dimensionality, and reliability of the EWAT instrument. An important finding of this study is that competing CFA models reduce the tendency for the researcher to achieve optimal model fit by drawing correlations between error terms based on the information suggested from post hoc modification indices available in CB-SEM software (LISREL, AMOS). This finding is in sync with the argument of Tomarken and Waller (2003), as well as Landis et al. (2010) that CFA models should be retained as they are. Rather, the model that best fits the data should be estimated in line with the empirical and theoretical underpinnings of the study. Based on the bifactor model examined and identified as the best-fitting model, the concern of replicability of the measurement model can be addressed (Ventura-Leon et al., 2021).

Furthermore, the study substantiated the use of bifactor models to explain the scale dimensionality and model reliability of the EWAT instrument (Tables 2 and 3). Compared to the initial correlated factors model for EWAT, the bifactor model improved the instrument in two aspects- a) the bifactor model had better-fit statistics, and b) the bifactor model substantiated the use of a single factor score to measure employee perception of the work environment. Therefore, the initial correlated model implied that the correlated factors- D1 and D2 cannot be used as individual latent scores to measure facets of employee perception of their work environment.

Interestingly, the value of bifactor analysis is strengthened by the affirmation of the fact that the unidimensional model is the best form of dimensionality even when the model fit attributes of the general factor model (Figure 1) were very poor (Table 1). This outcome is in sync with the findings of Hammer and Toland (2017) in which a bifactor analysis revealed that a single unidimensional construct of the Internalized stigma of mental illness scale (ISMI-9) was the most reliable model for the use of total scores with no support for use of specific factor scores (Hammer & Toland, 2017). By extension, the

Vol.12 (2024) no.2, pp.101-115; DOI 10.2478/mdke-2024-0007

use of multiple competing models supports the submission of Kline (2016) and Stone (2021) to ensure that the best-fitting model is selected instead of selecting a few indices of goodness of fit to report. Thus, the bifactor model of the instrument proposed in this study substantially improved the initial versions of the EWAT instrument.

The study provides a simplified illustration of scale dimensionality and model reliability testing of research instruments. Therefore, applied management researchers should use the EWAT instrument as a unidimensional construct, which implies that a single latent score can be used to estimate employee perception of their working conditions (that is, how well or bad they are). Management scholars should apply competing models compared against bifactor models to improve or support the use of single scores (unidimensional constructs) or multidimensional constructs (specific scores for different constructs) (Byllesby & Palmieri 2023; Reise et al., 2007). Also, the findings of the study can be applied as a management measurement tool to justify the use of latent scores to compare employees' perceptions before and after an intervention by management. Furthermore, the re-validated tool can be used as a reliable information source for human and operational resource management for gap analysis among employees. However, on a cautionary note according to Dunn and McGray (2020), Gegenfurtner (2022) and Rijmen (2010), there is a need to guide against the sole use of bifactor statistical modeling for analyzing research without ample consideration for underlying empirical fit, and theoretical foundations.

Conclusions

The study illustrated the use of bifactor modeling by comparing fit statistics from four competing models (general factor, correlated factor model, second-order hierarchical model, and bifactor model) examined in the study. The bifactor model provided the best fit of data to the model, and its ancillary measures gave evidence of a unidimensional construct suitable for the interpretation of the total score as a reliable measure of employee perception of their work environment. The use of specific factors or constructs was not justified as presented in the initially developed model of the instrument. The study concludes that the instrument is best described and applied as a unidimensional construct, and therefore, a single score can be used to rate employee's perception of their work conditions. In addition, the study presented practical guidance for management researchers and simplified reporting for bifactor modeling. Bifactor modeling assessment in management research ensures focus on the development of robust, fit and reliable measurement models to capture perceptions of respondents in survey research without undue manipulation of fit measures. Furthermore, researchers are encouraged to adopt the habit of evaluating competing CFA models as best practice in the development of self–reported questionnaires. Based on the availability of appropriately dimensioned research instruments, management practitioners are equipped with reliable tools that generate reliable scores to support dependable management decisions.

Implications of the findings of the study to management research and practice

The study provided empirical evidence which showed that the bifactor model provided a true representation of the EWAT instrument as a reliable, one-dimensional measure of employee perception. Therefore, a single latent total score adequately and reliably captures the dimension of the construct based on ancillary bifactor measures. In the context of research, researchers are encouraged to as a matter of practice critically examine competing CFA models which substantially improves the precision of the selection of appropriate models. As a result, the selected model adequately captures the dimensionality of the target construct/s of empirical interest. This approach is corroborated by the assertion of Sarstedt et al. (2024) that management researchers should adopt a critical assessment of competing CFA models so that measures obtained can be reliably applied by other researchers, management practitioners and policymakers. By extension, the use of bifactor models in the dimensionality and reliability assessment of research instruments helps to encourage the replicability of research which has been a concern for researchers in the social and management sciences (Block, Fisch, Kanwal, Lorenzen, & Schulze, 2023; Dau, Santangelo, & Van Witteloostuijn, 2021). Hence, researchers are advised to examine and report measures of bifactor modeling of any given research instrument.

The study substantially adds to the existing literature on the use of general factor, correlated factors model, second-order model, and bifactor models, to achieve the best dimensionality for any developed instrument. As a result, provides reasonable and logical information to support the critique of CFA models in published research. Hence, inferences for the use of factor scores of constructs or subscales factor scores can be empirically established. In other words, the justification to support confident use of latent construct scores and/ or composite scores is covered by using ancillary bifactor measures. As a result, management practitioners can confidently use latent scores to provide reliable quantitative means of measuring or representing perception from self-reported instruments. Therefore, the simplified presentation of latent scores provides quantitative evidence to support managerial decisions. Using the EWAT instrument as a case study, the perception scores of employees based on monthly, quarterly, or yearly evaluations can be compared. For instance, a hypothetical annual employee score in an organisation using EWAT with values of 500 in year 1, 700 in year 2, and 750 in 3 is suggestive of a growing trend of improved welfare conditions from the perspective of the employees.

Limitations of the study

The EWAT instrument was developed within the context of the Nigerian pharmaceutical marketing industry; hence, the instrument should be validated across other countries and industries. Due to the focus of the study on dimensionality assessment using bifactor modeling, measurement invariance evaluation of the instrument was not conducted. Hence, further studies involving questionnaires or psychometric tools in management research should include measurement invariance assessment and reporting in confirmatory factor analysis. This is relevant because measurement invariance assessment of a questionnaire ensures that the understanding, interpretation, and response of subgroups of a target sample are determined to be equivalent. Thereby, justifies the application of multigroup analysis without concern for instrumentation bias. Furthermore, the comparison of bifactor analysis models across cultural and national boundaries to enhance the transferability of findings is advocated.

References

Arbuckle, J. L. (2016). *AMOS 24.0 User's Guide*. IBM SPSS.

Block, J. H., Fisch, C., Kanwal, N., Lorenzen, S., & Schulze, A. (2023). Replication studies in top management journals: An empirical investigation of prevalence, types, outcomes, and impact. *Management Review Quarterly, 73*, 1109–1134. https://doi.org/10.1007/s11301-022-00269-6

Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science, 5*(1), 184-186. https://doi.org/10.1177/2167702616657069

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality, 80*(4), 796-846. https://psycnet.apa.org/doi/10.1111/j.1467-6494.2011.00749.x

Byllesby, B. M., & Palmieri, P. A. (2023). A bifactor model of general and specific PTSD symptom change during treatment. *Assessment. 30*(8), 2595-2604. https://doi.org/10.1177/10731911231156646.

Canivez, G. L. (2014). Construct validity of the WISC-IV with a referred sample: Direct versus indirect hierarchical structures. *School Psychology Quarterly, 29*(1), 38-51. http://doi.org/10.1037/spq0000032

Canivez, G. L. (2016). Bifactor modeling in construct validation of multifactored tests: Implications for multidimensionality and test interpretation. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements* (pp. 247–271). Hogrefe.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504. http://doi.org/10.1080/10705510701301834

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality, 80*(1), 219–251. https://doi.org/10.1111/j.1467-6494.2011.00739.x

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255. http://doi.org/10.1207IS15328007SEM0902_5

Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods, 12*(4), 381–398. https://doi.org/10.1037/1082-989X.12.4.381

Collier, J. (2020). *Applied structural equation modeling using AMOS: Basic to advanced techniques*. Routledge.

Dau, L. A., Santangelo, G. D., & van Witteloostuijn, A. (2022). Replication studies in international business. *Journal of International Business Studies, 53*(2), 215-230. https://doi.org/10.1057/s41267-021-00471-w

Dueber, D. M. (2017, April 10). *Bifactor indices calculator: a Microsoft excel-based tool to calculate various indices relevant to bifactor CFA models*. Digital Commons. https://doi.org/10.13023/edp.tool.01

Dunn, K. J., & McCray, G. (2020). The place of the bifactor model in confirmatory factor analysis investigations into construct dimensionality in language testing. *Frontiers in Psychology, 11,* 1357. https://doi.org/10.3389/fpsyg.2020.01357

Gegenfurtner, A. (2022). Bifactor exploratory structural equation modeling: A meta-analytic review of model fit. *Frontiers in Psychology, 13,* 1037111. https://doi.org/10.3389/fpsyg.2022.1037111

Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: G as superordinate or breadth factor? *Psychology Science Quarterly, 50,* 21-43.

Gorsuch, R. L. (1983). Two-and three-mode factor analysis. In R.L. Gorsuch (Ed.), *Factor Analysis* (2nd Ed.). Erlbaum.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6*(4), 430-450. https://doi.org/10.1037/1082-989x.6.4.430

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. 2017. *A primer on partial least squares structural equation modelling*. Sage.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and report the results of PLS-SEM. European Business Review, *31*(1), 2-24.

Hammer, J. H., & Toland, M. D. (2017). Internal structure and reliability of the Internalized stigma of mental illness scale (ISMI-29) and brief versions (ISMI-10, ISMI-9) among Americans with depression. *Stigma and Health*, *2*(3), 159-174. https://doi.org/10.1037/sah0000049.

Hammer, J. H., McDermott, R. C., Levant, R. F., & McKelvey, D. K. (2018). Dimensionality, reliability, and validity of the Gender-Role Conflict Scale-Short Form (GRCS-SF). *Psychology of Men & Masculinity, 19*(4), 570-583.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115-135.

Hu, L.-T., & Bentler, P. M. (1999). Cut-off 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

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th Ed). Guilford Press.

Landis, R. S., Edwards, B. D., & Cortina, J. M. (2010*). On the practice of allowing correlated residuals among indicators in structural equation models*. In C.E. Lance & R.J. Vandenberg (Eds.), *Statistical and methodological myths and urban legend* (pp. 213-236). Routledge.

Luo, Y., & Al-Harbi, K. (2016). The utility of the bifactor method for unidimensionality assessment when other methods disagree: an empirical illustration. *Sage Open, 6*(4), 1–7. https://doi.org/10.1177/2158244016674513

Oamen, T. E. (2021). The analysis of factors influencing pharmaceutical sales workforce engagement in pharmaceutical marketing in Nigeria: A structural equation modeling approach. *Global Journal of Pure and Applied Sciences*, *27*(4), 45-51. https://doi.org/10.4314/gjpas.v27i4.7.

Oamen, T. E. (2023). The moderating effect of contextual factors on the impact of competitive behavior on community pharmacists' performance in Nigeria. *International Journal of Economic Behavior, 13*(1), 93-108. https://doi.org/10.14276/2285-0430.3743.

Oamen, T. E., Idiake, J., & Omorenuwa, O. S. (2022b). Assessment of measurement invariance of psychometric tool for pharmaceutical sales executives: implications for social and behavioral pharmacy research. *Journal of Pharmaceutical Health Services Research*, *13*(4), 262-268. https://doi.org/10.1093/jphsr/rmac041.

Oamen, T. E., Omorenuwa, O. S., & Moshood, L. B. (2022a). A structural equation analysis of employment work assessment tool for pharmaceutical executives. *Journal of Social and Educational Research, 1*(1), 14-20.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667-696. https://doi.org/10.1080/00273171.2012.7

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment, 95,* 129-140. http://doi.org/10.1080/00223891.2012.725437

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*(6), 544–559, https://doi.org/10.1080/00223891.2010.496477

Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, *16*(1), 19-31. https://doi.org/10.1007/s11136-007-9183-7

Rijmen, F. (2010). Formal relations and an empirical comparison among the bifactor, the testlet, and a second-order multidimensional IRT model. *Journal of Educational Measurement*, *47*(3), 361–372. https://doi.org/10.1111/j.1745-3984.2010.0011

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137-150. https://doi.org/10.1037/met0000045

Sarstedt, M., Adler, S. J., Ringle, C. M., Cho, G., Diamantopoulos, A., Hwang, H., & Liengaard, B. D. (2024). Same model, same data, but different outcomes: Evaluating the impact of method choices in structural equation modelling. *Journal of Product Innovation management*, 1-17. https://doi.org/10.1111/jpim.12738

Savahl, S., Casa, F., & Adams, S. (2023). Considering a bifactor model of children's subjective well-being using a multinational sample. *Child Indicators Research*, *16*, 2253–2278. https://doi.org/10.1007/s12187-023-10058-6

Schuberth, F. (2021). Confirmatory composite analysis using partial least squares: Setting the record straight. *Review of Managerial Science*, *15*, 1311–1345.

Stone, B. M. (2021). The ethical use of fit indices in structural equation modeling: Recommendations for psychologists. *Frontiers in Psychology*, *12*, 783226. https://doi.org/10.3389/fpsyg.2021.783226

Stucky, B. D., & Edelen, M. O. (2015). Using hierarchical IRT models to create unidimensional measures from multidimensional data. In S.P. Reise & D.A.

Vol.12 (2024) no.2, pp.101-115; DOI 10.2478/mdke-2024-0007

Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (pp. 183-206). Routledge.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* American Psychological Association. https://doi.org/10.1037- 0694-000

Tomarken, A. J., & Waller, N. G. (2003). Potential problems with "well-fitting" models. *Journal of Abnormal Psychology, 112*(4), 578–598. https://doi.org/10.1037/0021-843X.112.4.578

Torres-Vallejos, J., Juarros-Basterretxea, J., Oyanedel, J. C., & Sato, M. (2021). A bifactor model of subjective well-being at personal, community, and country levels: a case with three Latin-American countries. *Frontiers in Psychology*, *12*, 641641. https://doi.org/10.3389/fpsyg.2021.641641

Vandenberg, R. J. (2006). Statistical and Methodological Myths and Urban Legends. *Organisational Research Methods*, *9*(2), 194-201.

Ventura-León, J., Quiroz-Burga, L., Caycho-Rodríguez, T., & Valencia, P. D. (2021). BifactorCalc: An online calculator for auxiliary measures of bifactor models. *Revista Evaluar, 21*(3), 1-14.

Zhang, B., Sun, T., Cao, M., & Drasgow, F. (2021). Using bifactor models to examine the predictive validity of hierarchical constructs: Pros, cons, and solutions. *Organizational Research Methods*, *24*(3), 530–571. https://doi.org/10.1177/1094428120915522

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for wh. *Applied Psychological Measurement, 30*(2), 121-144. http://doi.org/10.1177/0146621605278814