

BifactorCalc: An Online Calculator for Ancillary Measures of Bifactor Models

BifactorCalc: Una calculadora en línea para medidas auxiliares de modelos bifactor

José Ventura-León *¹, Luis Quiroz-Burga¹,
Tomás Caycho-Rodríguez¹, Pablo D. Valencia²

1 - Universidad Privada del Norte (UPN), Facultad de Ciencias de la Salud.

2 - Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Iztacala.

Introduction
Software
development
Conclusions
References

Recibido: 01/05/2021 Revisado: 15/06/2021 Aceptado: 24/06/2021

Abstract

The bifactor model allows examining the presence of a total score in a data set by modeling a general factor and two or more specific factors with an orthogonal relationship. These models tend to overestimate the goodness of fit (e.g., CFI, RMSEA, SRMR), hence there exist auxiliary measures that allow examining the dimensionality (ECV_{Gen} ; $ECV_{Specific}$; I-ECV, PUC, ARPB), and reliability (ω , ω_S , ω_H , ω_{HS} , PRV, H, and FD). The present study describes the operation, mathematical foundations, and application in psychological research of an online calculator called *BifactorCalc*. The results demonstrate that *BifactorCalc* is an online, user-friendly, and easy-to-use computer program for the calculation of the different auxiliary measures of bifactor models. It was concluded that the computer tool *BifactorCalc* is able to calculate the auxiliary measures of bifactor models in three simple steps and generate a *path* diagram.

Keywords: *software, bifactor, SEM, calculator, auxiliary measures*

Resumen

El modelo bifactor permite examinar la presencia de una puntuación total en un conjunto de datos a partir del modelamiento de un factor general y dos o más factores específicos con relación ortogonal. Estos modelos tienden a sobreestimar las bondades de ajuste (v.g., CFI, RMSEA, SRMR), y por esta razón es que existen medidas auxiliares que permiten examinar la dimensionalidad (ECV_{Gen} ; $ECV_{Specific}$; I-ECV, PUC, ARPB) y la fiabilidad (ω , ω_S , ω_H , ω_{HS} , PRV, H y FD). El presente estudio describe el funcionamiento, fundamentos matemáticos y aplicación en la investigación psicológica de una calculadora online denominada *BifactorCalc*. Los resultados demuestran que el *BifactorCalc* es un programa informático online, amigable y de fácil utilización para el cálculo de las diferentes medidas auxiliares de los modelos bifactor. Se concluye que el *BifactorCalc* es una herramienta informática que tiene la capacidad de calcular las medidas auxiliares de modelos bifactor en tres simples pasos y generar un diagrama *path*.

Palabras clave: *software, bifactor, SEM, calculadora, medidas auxiliares*

*Correspondence to: Dr. José Ventura-León, Facultad de Salud, Universidad Privada del Norte, Av. Alfredo Mendiola, 6062. Los Olivos, Perú. E-mail: jose.ventura@upn.pe

How to cite: 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. Retrieved from <https://revistas.unc.edu.ar/index.php/revaluar>

Participaron en la edición de este artículo: Eva Crasso, Stefano Macri, Juan Balverdi, Alicia Molinari, Mónica Serppe, Eugenia Maiorana, Benjamín Casanova, Ricardo Hernández.

Introduction

In the field of psychological evaluation, whether it be for diagnosis, intervention or research, the objective is to obtain the measurement of the construct (e.g., anxiety, depression, or stress) and the dimensions that comprise it. In this context, it is common to assume that the totality of items is influenced by the same latent variable; composed by specific factors that are integrated in a great general factor (Dominguez-Lara & Rodriguez, 2017). In that sense, the scores of the specific factors are previously summed to obtain a general score. However, it has recently been argued that to conduct this procedure, empirical evidence of the presence of a general factor must be obtained from statistical modeling (Reise, 2012).

In the context of psychology, it is common to use structural equation models (SEM), which are statistical techniques that assume the presence of a latent variable underlying the items and constitute an important alternative for evaluating psychological constructs, which are generally multidimensional (Bonifay, Lane, & Reise, 2017).

Bifactor models (see Figure 1) are within the so-called hierarchical models (Canivez, 2016; Reise, 2012), also called nested factor models (Gustafsson & Balke, 1993), and direct hierarchical models (Gignac, 2008). Its main characteristic is to evaluate the simultaneous effect of a general factor (GF), and specific factors (e.g., F1 and F2), on a set of indicators (Flores-Kanter, Dominguez-Lara, Trógolo, & Medrano, 2018). In that sense, the specific factors (SFs) are assumed to be orthogonal to each other (DeMars, 2013) because the shared variance between the specific factors is due to the general factor (Reise, 2012). Thus, the GF—in comparison to the SFs—is supposed to explain the items' greater amount of variance.

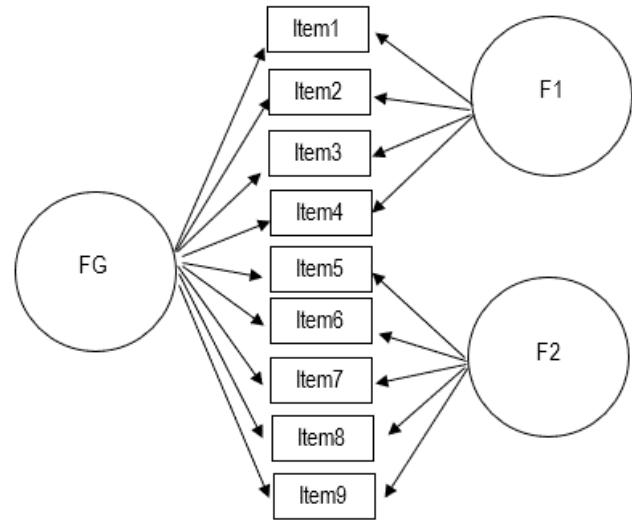


Figure 1
Diagram of bifactor model.

Although the bifactor model was originally described in the late 1930s (Holzinger & Swineford, 1937), it has been rediscovered in the past years (Reise, 2012) and it is increasingly used in psychological research conducted in diverse cultural contexts (Anderson & Marcus, 2019; Montes & Sanchez, 2019; Vuyk & Cudas, 2019). Nevertheless, the bifactor model is not immune to criticism. The evaluation of bifactor models with SEM techniques and the traditional goodness-of-fit indices only (e.g., CFI, RMSEA) can lead to false positives since it fails to evaluate the influence of the general factor and specific factors on the items (Bonifay et al., 2017; Dominguez-Lara & Rodriguez, 2017; Flores-Kanter et al., 2018). In fact, the evidence suggests that traditional goodness-of-fit indices may statistically favor bifactor models (Gignac, 2008; Morgan, Hodge, Wells, & Watkins, 2015).

Another important aspect is that the *interchangeability* of the specific factors in symmetric bifactor models (see Figure 1) is a prerequisite for its correct interpretation and the avoidance of anomalous models (see also Eid, Geiser, Koch, & Heene, 2017). The specialized literature provides

Table 1

Comparative information between BifactorCalc and other software.

Comparables	BifactorIndices Calculator	Excel® sheets*	BifactorCalc
Installation	R y RStudio	Excel®	None
Skills	R language programming	Using Excel	None
Diagram generation	No	No	Yes. Automatically
Report in APA format	No	No	Calculation output in APA format

Note. *Excel ® developed by [Dueber \(2017\)](#).

some examples of the correct use of the symmetric and *structurally different* bifactor models for the Beck Depression Inventory-II ([Heinrich, Zagorscak, Eid, & Knaevelsrud, 2018](#)) and ADHD/ODD symptoms ([Burns, Geiser, Servera, Becker, & Beauchaine, 2019](#)).

In this context, it is necessary to have a set of auxiliary measures that allow for a better evaluation of the bifactor model. Specifically software, which is needed to calculate all these measures quickly and easily. Currently, Excel® sheets are available ([Dueber, 2017](#)) and an R package called “BifactorIndicesCalculator” ([Dueber, 2020](#)) has recently been made available. The latter requires programming skills that are still not common among psychology professionals (comparative information in Table 1). This increases the need for develop a computer program for the calculation of the auxiliary measures of bifactor models, which provides a diagram with the factorial loads entries. In this sense, the objective of this research is to develop a software called *BifactorCalc* that allows for the calculation of the auxiliary measures of bifactor models in an easy, friendly way.

Omega Coefficients

For a bifactor structure, four types of omega coefficients can be calculated: Total Omega (ω),

Subscale Omega (ω_s), Hierarchical Omega (ω_H) and Hierarchical Omega for Subscale (ω_{HS}).

The omega coefficient (ω , [McDonald, 2013](#)) estimates what proportion of variance in the total observed score can be attributed to all common sources of variance ([Reise, Bonifay, & Haviland, 2013](#)). The ω is based on the factor loadings of a factorial model. Unlike other coefficients such as alpha, which is based on the assumption of equal loadings (tau-equivalent models), the omega coefficient is appropriate for cases in which the loadings of the items vary (congeneric models), an indication that is supported by several authors ([Dunn, Baguley, & Brunnsden, 2013](#); [Rodriguez, Reise, & Haviland, 2015](#)). The calculation of omega is as follows:

$$\omega = \frac{\left(\sum \lambda_{gen}\right)^2 + \sum_{k=1}^k \left(\sum \lambda_{grp_k}\right)^2}{\left(\sum \lambda_{gen}\right)^2 + \sum_{k=0}^k \left(\sum \lambda_{grp_k}\right)^2 + \sum (1-h^2)} \quad (1)$$

In the formula, the numerator expresses all common sources of variation of the total weighted score, and the denominator represents all common sources of total variance of the score plus the unique variance. High values of ω indicate high multidimensional composite reliability.

In the same way, the omega coefficient can

be calculated for the specific factors (ω_s) from the factor loadings and errors corresponding to each set of items that comprise the subscale. The following formula is used to calculate ω_s , when the variance of the general factor and the specific factors are combined to estimate reliability:

$$\omega_S = \frac{(\sum \lambda_{gen})^2 + (\sum \lambda_{k=1})^2}{(\sum \lambda_{gen})^2 + (\sum \lambda_{k=1})^2 + \sum (1-h^2)} \quad (2)$$

Both the ω and ω_s coefficients reflect the systematic variation, attributed to various common factors, that affects weighted composite scores (Rodríguez, Reise, & Haviland, 2016). In this context, it is important to determine the relative weight of the different factors that determine the variance of the composite scores. To that end, some alternate indices have been developed: hierarchical omega (ω_H), and hierarchical omega for the subscale (ω_{HS}). Both ω_H and ω_{HS} reflect the variance attributed to a single latent variable (Rodríguez et al., 2015).

Specifically, ω_H estimates the proportion of variance of the total scores that can be attributed to a single general factor and it is calculated by dividing the squared sum of the factor loadings on the general factor by the variance of the total scores (Reise, Moore, & Haviland, 2013).

$$\omega_H = \frac{(\sum \lambda_{gen})^2}{(\sum \lambda_{gen})^2 + \sum_{k=1}^k (\sum \lambda_{grp_k})^2 + \sum (1-h^2)} \quad (3)$$

The ω_H is sensitive to the number of items. A greater number of items is associated with an increase in the ω_H , which is also affected by

the relative size of the factor load of each item in the general factor, versus the specific factors (Rodríguez et al., 2015). A high ω_H ($\omega_H > .80$) would express that the scores can be considered essentially unidimensional, since *the general factor* is the main source of systematic variance compared to the influence of the *specific factors*.

The calculation of ω_H can be extended to subscales through the calculation of the hierarchical omega (ω_{HS}), which reflects the proportion of systematic variance of a subscale score after separating the variability attributed to the general factor (Reise, Bonifay et al., 2013). Hierarchical omega is calculated from the following formula:

$$\omega_{HS} = \frac{(\sum \lambda_{grp_k})^2}{(\sum \lambda_{gen})^2 + (\sum \lambda_{grp_k})^2 + \sum (1-h^2)} \quad (4)$$

Thus, there are some cut-off points in psychology that can be used as a reference: $\omega_{HS} \geq .30$ is substantial; $.20 \leq \omega_{HS} < .30$ is moderate and $\omega_{HS} < .20$ is low (Smits et al., 2014).

Percentage of Reliable Variance (PRV)

The percentage of reliable variance (PRV) is an indicator based on the logic of the bifactor model because it considers the variance explained by the general factor (Hammer et al., 2018). This index is the ratio of ω_H to ω ; and therefore, it can be conceptually understood as the percentage of the total reliability that can be attributed to the reliability of the general factor (Reise, Moore, et al., 2013). See the following equation:

$$PVR = \frac{\omega_H}{\omega} \times 100\% \quad (5)$$

Thus, some authors propose as a provisional cut-off point a $PVR > 50$, which would indicate that half of the reliable variation in test score is produced by the general factor (or the specific one, in which case ω_{HS} is replaced by ω_H in the numerator; [Li, 2015](#)).

Explained Common Variance (ECV)

The explained common variance (ECV) is an indicator of unidimensionality and expresses the proportion of the common variance that can be attributed to the general factor ([Reise, Moore, et al., 2013](#)). For its calculation, the factor loadings of the general and specific factors of a bifactor model are used on the following mathematical expression:

$$ECV_{GEN} = \frac{\sum \lambda_{GEN}^2}{\sum \lambda_{GEN}^2 + \sum \lambda_{grp_k}^2} \quad (6)$$

Where: $\sum \lambda_{GEN}^2$ is the sum of the squared factor loadings of the general factor; $\sum \lambda_{grp_k}^2$ is the sum of the squared factor loads of the specific groups. High ECV values, greater than .60, suggest that the common variance among the specific factors is small compared to the general factor; and therefore, that the data would fit an essentially unidimensional model ([Reise, Scheines, Widaman, & Haviland, 2013](#)).

For example, it has been observed that, when the ECV is greater than .60, the correlation between the general factor and a criterion variable is not substantially affected if only the general factor is modeled and not the specific factors. In other words, high ECV values indicate that it is possible to use a unidimensional model

even if the data fits better with a bifactor model. Other provisional cut-off points suggested by the literature are .70 or .80 ([Rodríguez et al., 2016](#)). However, the interpretation of ECV must be done in conjunction with that of the percentage of uncontaminated correlations (PUC), which is described in the following section ([Reise, Scheines, et al., 2013](#)).

$$ECV_{Specific} = \frac{\sum \lambda_{grp_k}^2}{\sum \lambda_{GEN}^2 + (\sum \lambda_{grp_k})^2} \quad (7)$$

In the case of the specific factors, a variant of the formula is made by positioning the loadings of the specific factors in the numerator and the loadings of the general factor plus the specific ones in the denominator. On the other hand, it is possible to obtain an ECV for each of the items (ECV-I) with the following mathematical expression:

$$ECV - I = \frac{\lambda_{GEN}^2}{\lambda_{grp_n}^2 + \lambda_{GEN}^2} \quad (8)$$

ECV-I expresses the proportion of true variance of each item that is explained by the general factor ([Stucky et al., 2013](#)). Values greater than .85 suggest an influence of the general factor on the variance of the item ([Stucky & Edelen, 2014](#)).

Percentage of Uncontaminated Correlations (PUC)

The percentage of uncontaminated correlations (PUC; [Reise, Scheines, et al., 2013](#)) expresses in percent the amount of correlations that are

not corrupted by multidimensionality (Rodríguez et al., 2015). In other words, it expresses what percentage of the total correlations between items occurs between items belonging to different specific factors. Therefore, the PUC together with the ECV provide information about the bias towards forcing multidimensional data into unidimensional models. Its mathematical expression is presented below:

$$PUC = \frac{\frac{I_G^*(I_G-1)}{2} - \left[\frac{I_{S1}^*(I_{S1}-1)}{2} + \frac{I_{S2}^*(I_{S2}-1)}{2} + \dots + \frac{I_{Sn}^*(I_{Sn}-1)}{2} \right]}{\frac{I_G^*(I_G-1)}{2}} \quad (9)$$

Where: I_1 is the number of items loaded onto the general factor; I_{S1} is the number of items loaded onto the specific factor 1; I_{S2} is the number of items loaded onto the specific factor 2; I_{Sn} is the number of items loaded onto the specific factor n .

The interpretation of the PUC must be conducted in conjunction with the ECV. In practical terms, it has been suggested that when the PUC is greater than .80, the ECV value is not very relevant; on the other hand, when the PUC is less than .80, the ECV *should be* greater than .60 in order to treat the instrument as if it were unidimensional (Reise et al., 2013). From another perspective, it has been suggested that when ECV and PUC are both greater than .70, the scale can be treated as if it were unidimensional (Rodríguez et al., 2015).

Factor Determinacy (FD)

Often, researchers do not only model a latent variable, but also seek to estimate everyone's score on that latent variable. These individual scores are called factor scores and, in their simplest form, they correspond to the sum of the items belonging to a factor (DiStefano, Zhu, &

Mîndrilă, 2009). There are, however, more refined methods, which are based on estimates from factor analysis.

Nevertheless, one problem with factor scores is that of the so-called indeterminacy. Although the details are technically complex, in simple terms this refers to the fact that from the same factorial solution, it is possible to obtain very dissimilar and even contradictory factor scores. In this sense, FD expresses the multiple correlations between the observed variables (items) and the factor (Grice, 2001). This value can be obtained with the following formula (Beauducel, 2011):

$$FD = \text{diag}(\Phi \Lambda^T \Sigma^{-1} \Lambda \Phi)^{\frac{1}{2}} \quad (10)$$

Under the conditions described in this work, this value is also equivalent to the correlation between factor scores and factors (Beauducel, 2011; Grice, 2001). For this reason, FD varies from 0 to 1, and values close to 1 indicate a better determinacy. In that case, values higher than .80 have been suggested to allow an estimate of the general factor score (Gorsuch, 1983). However, other authors argue for a higher cut-off point ($> .90$; Grice, 2001; Rodríguez et al., 2015)

Construct Replicability (H)

Another index that can help to better understand the quality of the measurement model is the construct replicability (Mueller & Hancock, 2008). The H index can be used to assess whether the set of items representing a latent variable is adequate. Therefore, it determines if the SEM model is adequate and replicable in all studies. The H index is calculated from the following mathematical formula:

$$H = \frac{1}{1 + \frac{1}{\sum_{i=1}^k \frac{\lambda_i^2}{1-\lambda_i^2}}} \quad (11)$$

As per the formula above, H is a function of the sum of the factor loading ratios of the squared items (proportion of variance explained by the latent variable), in a factor divided by 1, minus the factor loading squared (Rodríguez et al., 2015). In this sense, as the number of items and the factor loading increase, the H index approaches 1. Values of H greater than .70 suggest that the latent variable is well defined and is more likely to be stable in other studies (Dominguez-Lara, 2016); while low values suggest a poorly defined latent variable, which changes in other studies. The ease of calculating and interpreting the H index, makes it an ideal means of judging the viability of a measurement model based on a set of items.

Average Relative Parameter Bias (ARPB)

The ARPB is a measure for examining the difference between the factor loading of a unidimensional model and the general factor loading of the bifactor model (see equation 12). According to some authors, a maximum difference of .12 to .15 may be acceptable (Rodríguez et al., 2015).

$$ARPB = \sum \frac{\lambda_{GEN} - \lambda_{UNI}}{\lambda_{GEN}} \quad (12)$$

Average Factor Loading (mean)

A first approach to the bifactor model consists in the simple inspection of its factor loadings (Reise, Moore, & Haviland, 2010). If a scale has a strong general factor and a weak set of specific factors, then the factor loadings of the latter will be notoriously low, while the general factor loadings will tend to be higher. A simple way to examine this is by calculating the arithmetic mean of the items. Following other authors, means lower than .30 in the specific factors can be considered secondary evidence of unidimensionality (Ferrando & Lorenzo-Seva, 2017).

Software development

Description of BifactorCalc

The *BifactorCalc* calculator was developed with Python programming language, and all the previously presented formulas were entered. For this purpose, the summation and matrix multiplication calculations were produced with the Numpy library (Harris et al., 2020). The Django framework was used to deploy the web project and build a user-friendly interface in an online version without the need to install the software on a computer (Django Software Foundation, 2019). The styles of the online interface were made with the Bootstrap web style framework, which provided the visual characteristics of the buttons, colors, and frames; tables in APA format, the distribution of the content on the screen, and all the other components displayed on the interface. Finally, the graphic construction of the *BifactorCalc* and the integration with the calculations were performed with JavaScript.

For use, *BifactorCalc* will require the user and password (Figure 2), which can be requested to the authors of the article via email to store their

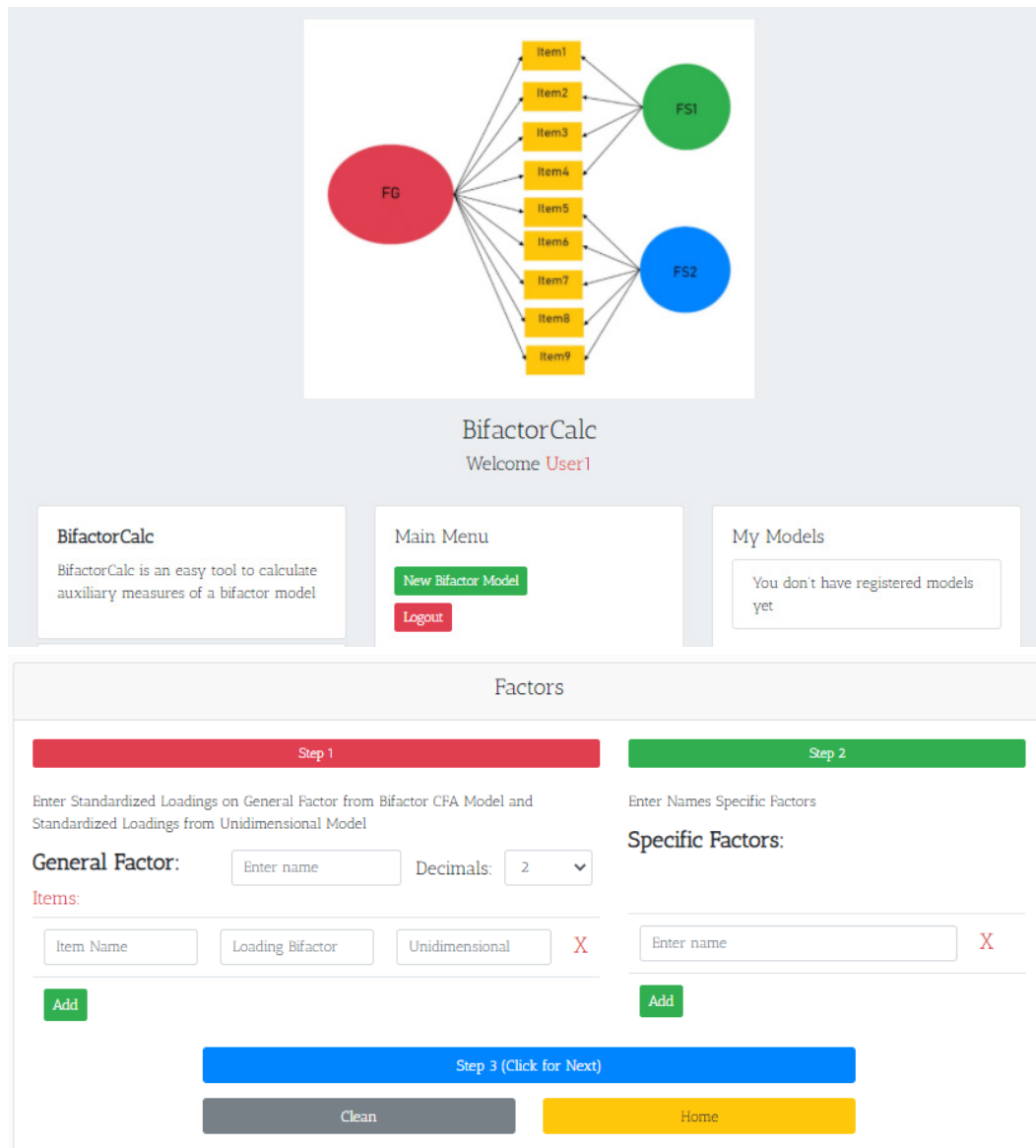


Figure 2
BifactorCalc Welcome Screen and information entry site.

bifactor models privately. Link to calculator:
<https://joseventuraleon.com/f/bifactorcalc>

In the main menu there are two options (Figure 2): *New Bifactor Model*, to generate new models and *Logout* to exit the application. In the *My Models* section, the models entered by the user will be displayed, identified by the name of the general factor assigned.

In the *New Bifactor Model* option you can

enter the factor loads of the model following the instructions provided in steps 1, 2 and 3. The procedure must be followed so that the calculator receives the data correctly and the information can be calculated satisfactorily.

In Step 1, you must enter the name of the General Factor, the items, and the general factor loadings. In addition, *BifactorCalc* allows for the entry of factor loadings from a unidimensional

model to calculate the ARPB, which measures the difference in the general factor loadings of the bifactor and unidimensional models. In Step 2, you must enter the names of the specific factors of the bifactor model. A maximum of two decimals should be used to enter the loads. Finally, it is necessary to *click* on Step 3, to continue entering the information.

Figure 3
First steps of BifactorCalc.

The figure shows two screenshots of the BifactorCalc software interface. The top screenshot is titled "Factors" and shows Step 1 and Step 2. Step 1 is "Enter Standardized Loadings on General Factor from Bifactor CFA Model and Standardized Loadings from Unidimensional Model". It features a "General Factor" dropdown set to "Ethnic Identity", a "Decimals" dropdown set to "2", and a table of items with their loadings for the general factor and a unidimensional model. Step 2 is "Enter Names Specific Factors". It features a "Specific Factors" section with two input fields: "Commitment" and "Exploration". The bottom screenshot is titled "Specific Factors" and shows Step 3: "Enter Standardized Loadings on Specific Factors from Bifactor CFA Model". It features two columns of items with their loadings for the specific factors "Commitment" and "Exploration".

Item	General Factor Loading	Unidimensional Loading
Item3	.71	.73
Item5	.90	.57
Item6	.70	.82
Item7	.72	.81
Item9	.71	.80
Item11	.89	.88
Item12	.66	.65
Item1	.53	.50
Item2	.46	.46
Item4	.53	.50
Item8	.66	.62
Item10	.77	.75

Item	Commitment Loading	Exploration Loading
Item3	.17	.42
Item5	.67	.14
Item6	.47	.41
Item7	.38	.44
Item9	.39	.03
Item11	.22	
Item12	.43	

Fourth, pressing the *Finish Bifactor Model* button automatically performs the calculation of the auxiliary measurements, which appear on a

new window, as shown in Figure 3.

Besides, *BifactorCalc* provides a diagram (see Figure 4) that can be copied and used in the scientific manuscript of the *BifactorCalc* user.

Validation of BifactorCalc

To demonstrate the operation of the calculator, information from Yap et al. (2014), was used as a sample, and the similarity of the results obtained in the *BifactorCalc* with the calculations of Rodriguez et al. (2016) was corroborated. Firstly, the factor loadings of the bifactor model provided by Yap et al. (2014) for its Ethnic Identity Scale (EIS) were entered in addition, unidimensional loadings were estimated from the inter-item correlation matrix with the R program (R Core Team, 2020). Secondly, the respective names were assigned to the specific and general factors (see Figure 3 on the left side). Thirdly, the factor loadings of the specific factors were entered (see Figure 3 on the right side).

As for the validation of *BifactorCalc*, the example explaining the “BifactorIndicesCalculator” package was run in R (Deber, 2020) and in *BifactorCalc*. The compared results similarity validate the correct functioning of the software.

Reporting BifactorCalc Results

In relation to the report of a bifactor model, this can be divided into two main moments: (a) Dimensionality, which consists of using the indexes ECV_{Gen} ; $ECV_{Specific}$; I-ECV, PUC and ARPB—to determine if the model is unidimensional or multidimensional—; and (b) Reliability, which consists of using the indexes ω , ω_s , ω_H , ω_{HS} , PRV, H and general and specific FD.

Using the information in the example, the

Model							
General Factor	N° Items	N° Specific Factors			Decimals	PUC	ARPB
Ethnic Identity	12	2			2	0.53	0.07

General Factor							
Name	Omega	OmegaH	PRV	H	FD	ECV	Mean
Ethnic Identity	0.93	0.81	0.87	0.92	0.94	0.75	0.65

Specific Factors							
Name	OmegaS	OmegaHS	PRV	H	FD	ECV	Mean
Commitment	0.93	0.22	0.24	0.64	0.79	0.26	0.39
Exploration	0.80	0.16	0.19	0.40	0.69	0.24	0.29

Items							
Name	Specific Factor	General Loading	Specific Loading	Standardized Residual Variance	Comunality	ECV-I	
Item3	Commitment	0.71	0.17	0.47	0.53	0.95	
Item5	Commitment	0.50	0.67	0.30	0.70	0.36	
Item6	Commitment	0.70	0.47	0.29	0.71	0.69	
Item7	Commitment	0.72	0.38	0.34	0.66	0.78	
Item9	Commitment	0.71	0.39	0.34	0.66	0.77	
Item11	Commitment	0.89	0.22	0.16	0.84	0.94	
Item12	Commitment	0.66	0.43	0.38	0.62	0.70	
Item1	Exploration	0.53	0.42	0.54	0.46	0.61	
Item2	Exploration	0.45	0.14	0.78	0.22	0.91	
Item4	Exploration	0.53	0.41	0.55	0.45	0.63	
Item8	Exploration	0.66	0.44	0.37	0.63	0.69	
Item10	Exploration	0.77	0.03	0.41	0.59	1.00	

Figure 4
Output of the results obtained with BifactorCalc.



Figure 5
Path Diagram of Bifactor Model.

following can be reported: In relation to the dimensionality of the Ethnic Identity Scale, it was observed that it presents an ECV_{Gen} .75, which suggests that the general factor explains 75% of the variance of the items, which could suggest a tendency towards unidimensionality ($ECV > .60$). In addition, the $ECV_{Specific1}$ and $ECV_{Specific2}$ presented a value of .26 and .24 respectively, which would indicate that the specific factor explains 26% and 24% of the common variance, respectively. In relation to the I-ECV it was observed that only items 3, 11, 2 and 10 are strongly influenced by the general factor ($I-ECV > .85$). The PUC was equal to .53. Therefore, 53% of the correlations are “contaminated” by the multidimensionality, leaving 47% of the correlations to be explained by the general factor alone. Finally, the ARPB is equal to .07, which indicates that the general factor loads of the bifactor model and factor loads of the unidimensional model are different only by 7%, being within the acceptable ranges.

In relation to the reliability of the IEE, it presented a ω of .93 and ω_s were .93 and .80 for the specific factor 1 and 2 respectively. All these values reveal an excellent composite reliability [the expressions suggested by Cicchetti (1994) used for Cronbach’s alpha are extrapolated]. With respect to the ω_H it is equal to .81 expressing that the general factor is the main source of variance in comparison with the specific factors. In this regard, ω_{HS} is .22, which can be considered a moderate consistency of factor 1; and .16 a low consistency of factor 2 (Smits et al., 2014). The PRV would indicate that 87% of the reliable variance is due to the general factor and only 24% and 19% of the reliable variance to the specific factors. The H coefficient is equal to .92 in the general factor, which implies stability in other studies; while the specific Hs are less than .70, providing evidence in favor of the general factor. Finally, the FD for

the general factor and the two specific factors are: .94, .79 and .69 respectively, indicating that only the general factor score should be used for the analysis.

Conclusions

This work was aimed at designing a user-friendly, online calculator for the auxiliary measures of the bifactor model. Understanding that multidimensional models are increasingly common (Montes & Sanchez, 2019; Vuyk & Cudas, 2019), and the presence of a general factor should be verified empirically (Dominguez-Lara & Rodriguez, 2017; Flores-Kanter et al., 2018). Since the assumption that a high correlation between factors indicates the presence of a total score is no longer sufficient, it is necessary to examine this structure with a bifactor model (Anderson & Marcus, 2019). Some authors state that the bifactor model’s goodness-of-fit tends to be positively biased (Bonifay et al., 2017; Gignac, 2008; Morgan et al., 2015) and thus, it is necessary to explore auxiliary measures (Reise et al., 2013; Rodriguez et al., 2016). Despite this, there is no software for the estimation of these measures in a quick and simple way (only in three steps). The most similar option is an R-package (Dueber, 2020), which requires programming skills and the installation of an Office program.

In that sense, *BifactorCalc* is an online software that through a user and password enables the storing of Bifactor models, the modification of factor loadings in case of errors, and the estimation of all auxiliary measures in only three steps. In relation to its validity, information from Yap et al. (2014), and estimates made by Rodriguez et al. (2016), were used to verify that *BifactorCalc*, which reported the same results. This same procedure was performed with

R-package *BifactorIndicesCalculator* (Dueber, 2020). Thus, *BifactorCalc* operation proved to be optimal.

In addition, this research provides an example of how the results obtained with *BifactorCalc* can be reported; framed in two major moments, the review of the dimensionality and the reliability. In this way, the users of this software will be able to easily incorporate the results in their scientific manuscript.

Finally, the software is expected to contribute to the scientific community in the field of psychology and to promote methodological best practices associated with the implementation of the bifactor models in the Spanish-speaking context.

References

- Anderson, A. E., & Marcus, D. K. (2019). A bifactor model of meanness, coldheartedness, callousness, and sadism. *Personality and Individual Differences*, *137*, 192-197. doi: [10.1016/j.paid.2018.09.006](https://doi.org/10.1016/j.paid.2018.09.006)
- Beauducel, A. (2011). Indeterminacy of factor score estimates in slightly misspecified confirmatory factor models. *Journal of Modern Applied Statistical Methods*, *10*(2), 583-598. doi: [10.22237/jmasm/1320120900](https://doi.org/10.22237/jmasm/1320120900)
- 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. doi: [10.1177/2167702616657069](https://doi.org/10.1177/2167702616657069)
- Burns, G. L., Geiser, C., Servera, M., Becker, S. P., & Beauchaine, T. P. (2019). Application of the bifactor S-1 model to multisource ratings of ADHD/ODD symptoms: An appropriate bifactor model for symptom ratings. *Journal of Abnormal Child Psychology*, *48*(7), 881-894. doi: [10.1007/s10802-019-00608-4](https://doi.org/10.1007/s10802-019-00608-4)
- Canivez, G. L. (2016). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schweizer & C. DiStefano (Eds.), *Principles and Methods of Test Construction: Standards and Recent Advancements* (pp. 247-271). Gottingen, Germany: Hogrefe Publishers.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284-290.
- DeMars, C. E. (2013). A tutorial on interpreting bifactor model scores. *International Journal of Testing*, *13*(4), 354-378. doi: [10.1080/15305058.2013.799067](https://doi.org/10.1080/15305058.2013.799067)
- DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation*, *14*. Retrieved from <https://scholarworks.umass.edu/pare>
- Django Software Foundation. (2019). *Django*. Retrieved from <https://djangoproject.com>
- Dominguez-Lara, S. (2016). Evaluación de la confiabilidad del constructo mediante el Coeficiente H: Breve revisión conceptual y aplicaciones. *Psychologia. Avances de la Disciplina*, *10*(2), 87-94. doi: [10.21500/19002386.2134](https://doi.org/10.21500/19002386.2134)
- Dominguez-Lara, S., & Rodriguez, A. (2017). Índices estadísticos de modelos bifactor. *Interacciones. Revista de Avances en Psicología*, *3*(2), 59-65. doi: [10.24016/2017.v3n2.51](https://doi.org/10.24016/2017.v3n2.51)
- Dueber, D. M. (2017). *Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models*. doi: [10.13023/edp.tool.01](https://doi.org/10.13023/edp.tool.01)
- Dueber, D. M. (2020). Package 'BifactorIndicesCalculator'. Retrieved from <https://github.com/ddueber/BifactorIndicesCalculator>
- Dunn, T. J., Baguley, T., & Brunsten, V. (2013). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399-412. doi: [10.1111/bjop.12046](https://doi.org/10.1111/bjop.12046)
- Eid, M., Geiser, C., Koch, T., & Heene, M. (2017).

- Anomalous results in G-factor models: Explanations and alternatives. *Psychological Methods*, 22(3), 541-562. doi: [10.1037/met0000083](https://doi.org/10.1037/met0000083)
- Ferrando, P. J., & Lorenzo-Seva, U. (2017). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement*, 78(5), 762-780. doi: [10.1177/0013164417719308](https://doi.org/10.1177/0013164417719308)
- Flores-Kanter, P. E., Dominguez-Lara, S., Trógolo, M. A., & Medrano, L. A. (2018). Best practices in the use of bifactor models: Conceptual grounds, fit indices and complementary indicators. *Revista Evaluar*, 18(3). doi: [10.35670/1667-4545.v18.n3.22221](https://doi.org/10.35670/1667-4545.v18.n3.22221)
- Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: g as superordinate or breadth factor? *Psychology Science*, 50(1), 21-43.
- Gorsuch, R. L. (1983). Two-and three-mode factor analysis. In R. L. Gorsuch (Ed.), *Factor Analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, 6(4), 430-450. doi: [10.1037/1082-989x.6.4.430](https://doi.org/10.1037/1082-989x.6.4.430)
- Gustafsson, J.-E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28(4), 407-434. doi: [10.1207/s15327906mbr2804_2](https://doi.org/10.1207/s15327906mbr2804_2)
- 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.
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature* 585, 357-362. doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2)
- Heinrich, M., Zagorscak, P., Eid, M., & Knaevelsrud, C. (2018). Giving G a meaning: An application of the Bifactor-(S-1) approach to realize a more symptom-oriented modeling of the Beck Depression Inventory-II. *Assessment*, 27(7), 1429-1447. doi: [10.1177/1073191118803738](https://doi.org/10.1177/1073191118803738)
- Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2(1), 41-54. doi: [10.1007/bf02287965](https://doi.org/10.1007/bf02287965)
- Li, C. (2015). *The short Grit scale: A dimensionality analysis* (Tesis de maestría). University of Kentucky, EE.UU. Retrieved from http://uknowledge.uky.edu/edp_etds/33
- McDonald, R. P. (2013). *Test theory: A unified treatment*. New York: Psychology Press. doi: [10.4324/9781410601087](https://doi.org/10.4324/9781410601087)
- Montes, S. A., & Sanchez, R. O. (2019). El factor p. ¿La estructura subyacente a la psicopatología? *Revista Evaluar*, 19(3), 20-41. doi: [10.35670/1667-4545.v19.n3.26774](https://doi.org/10.35670/1667-4545.v19.n3.26774)
- Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are fit indices biased in favor of bi-factor models in cognitive ability research?: A comparison of fit in correlated factors, higher-order, and bi-factor models via Monte Carlo simulations. *Journal of Intelligence*, 3(1), 2-20. doi: [10.3390/jintelligence3010002](https://doi.org/10.3390/jintelligence3010002)
- Mueller, R. O., & Hancock, G. R. (2008). Best practices in structural equation modeling. In J. Osborne (Ed.), *Best Practices in Quantitative Methods* (pp. 488-508). Thousand Oaks, California: Sage.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from <https://www.R-project.org>
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667-696. doi: [10.1080/00273171.2012.715555](https://doi.org/10.1080/00273171.2012.715555)
- 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(2), 129-140. doi: [10.1080/00223891.2012.725437](https://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. doi: [10.1080/00223891.2010.496477](https://doi.org/10.1080/00223891.2010.496477)
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2013). Applying unidimensional item response theory models to psychological data. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise & M. C. Rodriguez (Eds.), *APA Handbook of Testing and Assessment in Psychology, Vol. 1. Test theory and testing and assessment in industrial and organizational psychology* (pp. 101-119). American Psychological Association. doi: [10.1037/14047-006](https://doi.org/10.1037/14047-006)
- Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement*, 73(1), 5-26. doi: [10.1177/0013164412449831](https://doi.org/10.1177/0013164412449831)
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2015). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3), 223-237. doi: [10.1080/00223891.2015.1089249](https://doi.org/10.1080/00223891.2015.1089249)
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21(2), 137-150. doi: [10.1037/met0000045](https://doi.org/10.1037/met0000045)
- Smits, I. A., Timmerman, M. E., Barelds, D. P., & Meijer, R. R. (2014). The Dutch symptom checklist-90-revised. *European Journal of Psychological Assessment*, 31(4), 263-271.
- Stucky, B. D., & Edelen, M. O. (2014). Using hierarchical IRT models to create unidimensional measures from multidimensional data. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment* (1st ed., pp. 183-206).
- Stucky, B. D., & Edelen, M. O. (2014). Using hierarchical IRT models to create unidimensional measures from multidimensional data. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment*, 183-206. Routledge/Taylor & Francis Group.
- Vuyk, M. A., & Codas, G. (2019). Validación de la Escala de Esperanza Disposicional para Adultos en Paraguay. *Revista Evaluar*, 19(1), 59-71. doi: [10.35670/1667-4545.v19.n1.23880](https://doi.org/10.35670/1667-4545.v19.n1.23880)
- Yap, S. C. Y., Donnellan, M. B., Schwartz, S. J., Kim, S. Y., Castillo, L. G., Zamboanga, B. L., ... & Vazsonyi, A. T. (2014). Investigating the structure and measurement invariance of the Multigroup Ethnic Identity Measure in a multiethnic sample of college students. *Journal of Counseling Psychology*, 61(3), 437-446. doi: [10.1037/a0036253](https://doi.org/10.1037/a0036253)