Psicothema

Psicothema 2017, Vol. 29, No. 3, 396-401
doi: 10.7334/psicothema2016.369

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

Rationale and Applicability of Exploratory Structural Equation Modeling (ESEM) in psychoeducational contexts

Cristiano Mauro Assis Gomes[1], Leandro S. Almeida[2] and José Carlos Núñez[3]
[1] Universidade Federal de Minas Gerais, [2] Universidade do Minho and [3] Universidad de Oviedo

Abstract

Background: In last few years, the use of confirmatory factor analysis (CFA) has become dominant in structural validation of psychological tests. However, the requirement of latent variables only loading on specific target items introduces some constraints on the solutions found, namely a factor solution that links some items only in one specific dimension. The most recent use of exploratory structural equation modeling (ESEM), which allows items to be predominantly related to a factor, with non-zero loadings on other factors, has been identified as the one that best respects the proper functioning of the assessed psychological attributes. **Method:** In this study we compared the two approaches to structural validity using the answers of a sample of 2,478 first-year higher education students to a multidimensional questionnaire of academic expectations. **Results:** The results show clear gains in information collected when combining CFA and ESEM. **Conclusions:** In conclusion, some implications are highlighted for research and practice of psychological assessment.

Keywords: Structural equation modeling, confirmatory factor analysis, exploratory structural equation modeling, academic expectations assessment.

Resumen

Fundamentos y aplicabilidad del Modelado Exploratorio de Ecuaciones Estructurales en contextos psicoeducativos. **Antecedentes:** en los últimos años, el uso del Análisis Factorial Confirmatorio (AFC) se ha convertido en un tipo de análisis predominante en la validación de tests psicológicos. Sin embargo, el requisito de que las variables latentes únicamente carguen sobre algunas de las respectivas dimensiones de destino conlleva algunas restricciones a las soluciones obtenidas; es decir, una solución factorial que requiere la vinculación de ciertos ítems solo en una dimensión. El uso más reciente del Modelo Exploratorio de Ecuaciones Estructurales (ESEM), que permite que los ítems puedan ser predominantemente relacionados con un factor y con cargas diferentes a cero en otros factores, ha sido identificado como aquel que mejor respeta el buen funcionamiento de los atributos psicológicos evaluados. **Método:** en este estudio, con las respuestas de una muestra de 2.478 estudiantes de primer año de la enseñanza superior a un cuestionario multidimensional de expectativas académicas, hemos comparado los dos enfoques de validez estructural. **Resultados:** los resultados muestran claros beneficios en la información recopilada al combinar el AFC y el ESEM. **Conclusiones:** como conclusión se señalan algunas implicaciones para la investigación y la práctica de evaluación psicológica.

Palabras clave: modelo de ecuaciones estructurales, análisis factorial confirmatorio, modelo exploratorio de ecuaciones estructurales, evaluación de las expectativas académicas.

Confirmatory factor analysis (CFA) has notably improved the empirical investigation of theories and the comparison of different models. Through this technique, researchers can investigate the relationships between latent variables and their causal role to explain the variance of certain observable variables. As the name implies, exploratory factor analysis (EFA) is used to explore this relationship, whereas CFA is used to confirm it, and the researcher can a priori define all the relations between latent and observable variables in a specific model (Caro & García, 2009).

Received: November 30, 2016 • Accepted: March 24, 2017
Corresponding author: Leandro Almeida
Instituto de Educação
Universidade do Minho
4715 Braga (Portugal)
e-mail: leandro@ie.uminho.pt

CFA is a subset of structural equation modeling, because the former establishes the measurement model, which is the relationship among certain latent variables with certain observable variables, whereas the latter possesses that same feature (measurement model) and also defines the relationship between all latent variables from different measurement models. For example, if a researcher a priori defines the relations between a latent variable of depression and some items that are supposedly explained by this latent variable, he or she would apply CFA. However, if this same researcher also defines a priori the relations between a latent variable of anxiety and some items that are supposedly explained by that latent variable, and moreover, defines the relationship between the two latent variables, anxiety and depression, then he or she would apply structural equation modeling.

In light of its theoretical rationale, CFA demands that each latent variable of the measurement model loads exclusively on at least two items related to that latent variable. These two items must not load on any other latent variables of the measurement model. For example, if a measurement model defines that 10 items of a questionnaire are related to two latent variables, and that item 1 to item 5 relate to the first latent variable, while item 6 to item 10 relate to the second latent variable, it is mandatory that at least two items from item 1 to item 5 load exclusively on the first latent variable, and that the second variable has zero loadings on both these items. The independent-cluster solution is an extreme case of CFA where all items related, in theory, to a specific latent variable must be loaded exclusively by their target latent variable, and must possess zero loadings from the other latent variables of the measurement model. Thus, if a questionnaire is supposed to measure two latent variables, for example, career expectancy and personal development expectancy, in the independent-cluster CFA solution the career expectancy target items can only load on the career expectancy latent variable, whereas personal development expectancy items can only load on the personal development expectancy latent variable. Hence, the career latent variable must have zero loadings on items that are markers of the personal development latent variable, and the personal development latent variable must have zero loadings on items that are markers of the career latent variable.

Although the constraint present in IC-CFA makes it easier to interpret and produce the scores related to each dimension because of the imposition that a latent variable is only related to its target items and not to other latent variables of the model, this imposition introduces several difficulties in the empirical verification. The reason is simple and straightforward: reality is not so pure. In the example of the latent variables, depression and anxiety, even though depression plays a preponderant role in the explanation of the variance of its target items, it is very hard to explain those items only by depression. Anxiety can very probably explain some of these items' variance, despite the fact that depression plays the strongest role their explanation.

Exploratory structural equation modeling (ESEM) is a new technique that aims to overcome the above-mentioned limitation, allowing cross-loadings among different latent variables and items of some questionnaires. ESEM does not impose any constraint that some items must be exclusively loaded by a specific latent variable. As we said, CFA determines that each latent variable of the measurement model has an exclusive relationship with a minimum of two items. Technical aspects of ESEM are presented in Asparouhov and Muthén (2009). This technique has the advantages of CFA, such as the calculation of model data fit, the a priori definition of the relation among latent variables and items, group invariance analysis, and so on (Morin & Maïano, 2011). The strict difference is that ESEM relaxes the constraint that, at least, two target items must exclusively load on their latent variables. Through this technique, the researcher can define which items will load on a latent variable, as well as which items will load on this same latent variable with loadings as close as possible to zero. This latter aspect is the fundamental difference between CFA or IC-CFA and ESEM. The latter has been very productive in situations where solutions from traditional exploratory factor analysis show a bad fit with CFA, particularly IC-CFA because of the strict condition of the latter of not allowing cross-loadings (Marsh, Morin, Parker, & Kaur, 2014).

ESEM has been applied mainly in the field of personality and the five-factor-approach (FFA), where the items of a questionnaire rarely load exclusively on one latent variable (Furnham, Guenole, Levine, & Chamorro-Premuzic, 2013). As commented by Marsh, Nagengast, and Morin (2012), "Confirmatory factor analyses (CFAs) conducted at the item level often do not support a priori FFA structures, due in part to the overly restrictive assumptions of CFA models" (Marsh et al., 2012, p. 1).

Although ESEM has been applied principally in psychological areas, its approach is applicable in any scientific field. Education, for example, can benefit from this approach. For example, it makes sense that a problem-solving item of mathematics should load both on a problem-solving latent variable and on a reading comprehension latent variable. In this case, clearly IC-CFA does not seem to be an adequate approach.

Despite their differences, CFA and ESEM perform similar tasks. Both "test how the data fit with a priori expectations, to systematically investigate the degree to which a measurement or predictive model is invariant across meaningful subgroups of participants, and to assess relations between constructs corrected for measurement errors" (Howard, Gagné, Morin, & Forest, 2016, p. 4). Moreover, it is impossible to state that CFA is better than ESEM or vice versa. CFA presents many positive aspects, and yet, also limitations, and the same holds true for ESEM. We do not intend to present a list of advantages or limitations, seeing that such an endeavor has been performed before efficiently by Marsh et al. (2014) and Howard et al. (2016), for example. However, the literature presents much evidence of the general ineffectiveness of CFA, especially IC-CFA, to analyze the fit of models related to psychological instruments using multiple latent variables. IC-CFA seems to work in very specific contexts for psychological instruments, and its use should be integrated with ESEM. This idea is not original, as it is clearly present in the following claim: "Over and above the intuitive appeal of clearly defined concepts, measured by a small number of items perfectly designed to assess a single construct, has come a recent recognition that the ideals pursued through a CFA approach are often impossible to achieve in applied research" (Howard et al., 2016, p. 4).

Aiming to employ ESEM in the psychoeducational area, the present paper applies this approach to the *Academic Perceptions Questionnaire - Expectations* (APQ-E; Almeida et al., 2012). This questionnaire is described in the Instruments section.

Two models are tested. The first one assumes that the specific target items exclusively load on the first-order factors (latent variables) and have zero loadings on the other factors. The factors may correlate with each other. Of course, this model corresponds to an IC-CFA. On the other hand, the second model relaxes the constraint of the first model and permits cross-loadings. This model defines the target items from each of the seven factors of theory, and establishes which items have a loading as close as possible to zero in other factors.

Thus, this article aims to compare the implications of CFA, particularly the broadly used IC-CFA, and ESEM in the structural analysis of a multidimensional psychoeducational questionnaire. These implications can be related to theoretical and practical aspects of psychoeducational assessment, namely the discrepancies between empirical versus internal approaches to test validity.

Method

Participants

The sample is composed of 2,478 first-year students of Minho University, a public higher education institution in the north of

Portugal. Most of the students are female (55.8%) and age mean is 18.65 years old (*SD* = 3.34). The fathers' educational level is predominantly elementary school (50.2%), followed by high school (29.2%), and higher education (20.6%).A similar pattern is found in mothers' educational level, with a tendency towards a higher academic level: most of the mothers had elementary school (42.8%), followed by high school (30.3%), and higher education (26.9%).

Instrument

Academic Perceptions Questionnaire - Expectations (APQ-E; Almeida et al., 2012). This instrument explores students' beliefs and aspirations in the transition to higher education, namely, what they expect to find and to develop. Its items combine cognitive and motivational aspects of academic experience that seems be related to students' academic engagement and adjustment. A total of 42 items divided into seven dimensions (6 items per dimension) are included: (a) Career: training for job and career development (e.g., professional preparation to get a good job); (b) Development: personal and social development (developing maturity and autonomy); (c) Mobility: student mobility (using Erasmus or similar programs to gain academic or practicum experiences in other countries); (d) Citizenship: political engagement and citizenship (discussing the world or country's socio-economic problems); (e)Pressure: social pressure (considering parents' and society's investment in their education); (f) Course: training quality (taking part in an interesting scientific graduation program); and (g) Living together: social interaction (participating in student parties and leisure activities). Students rate their agreement with item content on a 6-point Likert-type scale ranging from 1 (*completely disagree*) to 6 (*completely agree*). Response categories 1 to 3 (c1, c2, c3) represent negative judgments or low expectations about the statement of the item, while categories 4 to 6 (c4, c5, c6) indicate positive appraisals or optimistic expectations about the statement of the item. Reliability and structural validity analysis were conducted with a sample of first-year students after six months of academic experience (during the second semester), obtaining adequate psychometric coefficients for each dimension and for the internal structure of the questionnaire (Deaño et al., 2015). In the present sample, the following Cronbach alpha coefficients were obtained: Career (α = .82), Development (α = .84), Mobility (α = .89), Citizenship (α = .86), Pressure (α = .82), Course (α = .77), and Living together (α = .86).

Procedures

In this study, the questionnaire was applied when the candidates arrived at the university to enroll. After the 12th grade exams in June and July, students must wait for a place in a graduation course and institution according to a numerus clausus system (students choose six pairs of options combining courses and institutions). Early in September, the students and the higher education institutions receive an electronic list from the Education Government with the placements, after which the students have one week to enroll. Thus, the questionnaire was completed just when students start to confirm their interest in the course and institution where they were placed. The study objectives and procedures were presented to each student, and confidentiality was assured, in order to obtain their formal consent. Students were then invited into a room where a small group of psychologists handed out the questionnaire and answered any questions.

Data analysis

Both one Model 1 and Model 2 were run using statistical software Mplus 7.0 (Muthén & Muthén, 1998-2014). Model 1 was investigated through CFA, whereas Model 2 was tested with ESEM. The syntax used for performing ESEM shares many aspects with the CFA syntax, but it has added the cross-loadings with values as close as possible to zero. Moreover, as this approach uses an exploratory strategy, it is necessary to apply a rotation technique, in this case oblique target rotation. Target rotation was used seeing that this technique combines the best aspects of exploratory and confirmatory approaches. It emphasizes the confirmatory approach in ESEM, as it "provides a stronger a priori model, gives the researcher greater control in specifying the model, and facilitates interpretation of the results" (Marsh et al., 2014, p. 90).

The estimator used for all analyses was the weighted least squares estimation with robust mean and variance corrected chi-square statistic (WLSMV). This estimator treated the data as ordered categorical data. Both models had seven continuous latent variables representing the seven first-order factors, which presupposes the theoretical basis of the questionnaire. We applied WLSMV since the majority of items presented asymmetric patterns with a concentration of answers in the superior range of the scale.

Data fit of models was inspected through the comparative fit index (CFI) and root mean square error of approximation (RMSEA). A CFI value equal to or above .95 and a RMSEA value equal to or below .05 indicate a good model fit (Bentler, 1990; Browne & Cudeck, 1993; Hu & Bentler, 1999). Whereas RMSEA is a fit index that defines the lack of model fit, CFI is a fit index that aims toward the perfect fit, contrasting the null model with the tested model (Schumacker & Lomax, 2004).

Results

Model 1 determines that only the target items (six items for each factor) can load on their corresponding seven first-order factors (citizenship, career, development, mobility, pressure, course, and living together). The first-order factors of the model are allowed to correlate with each other. Figure 1 shows the relations in Model 1 among the factors and items of questionnaire. The correlations between the factors were omitted in Figure 1 but can be seen in Table 2. The data fit for Model 1(χ^2[798]=11,924.59, *CFI*=.884, *RMSEA*=.075, 90% CI [.074, .076]) shows that it should be rejected, because it did not achieve the minimum acceptable CFI value (.90). The literature recommends that values below this are unacceptable (Bentler, 1990).

Model 2 relaxes the constraint that the target items must only load on their factors. Data fit of Model 2 was good (χ^2[588]=4,137.934, *CFI*=.963, *RMSEA*=.049, 90% CI [.048, .051]) because the CFI value was higher than .95, and the RMSEA was below .05. Thus, the model cannot be rejected.

Table 1 shows the loadings of the items of the questionnaire on the first-order factors. Values below .20 were omitted, and the target items related to each factor are shaded in gray. The target items of the mobility and citizenship factors had adequate loadings (values of at least .40). Five of their six target items had adequate loadings

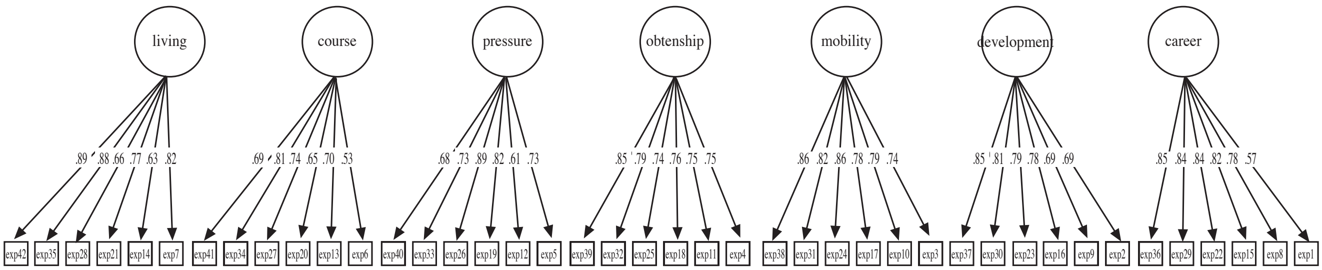

Figure 1. Latent variables, items and their loadings in Model 1.The correlations between the factors were omitted (see Table 2).
Legend: Note: Career = training for job and career development, Development = personal and social development, Mobility = student mobility, Citizenship = political engagement and citizenship, Pressure = social pressure, Course = training quality, Living [together] = social interaction

on the career, development, living together, and pressure factors. The course factor was the worst, because only three of its six target items had adequate loadings. Two items (1 and 27) showed very low loadings as target marker items, with values below .20.

The career factor was the latent variable with the highest number of cross-loadings with a value at least of .20. Career factor had loadings from nine non-target items (loadings equal to or above .20). Some of these items loaded on career factor more than on their theoretical target factor, for instance, Item 23, with a loading of .382 on career factor and of .250 on development factor (its target factor). The same thing was observed with Item 34, with a loading of .460 on career factor and a value of .241 on its target factor, course. Therefore, Items 23 and 34 seem to be explained more by career factor than by their target factors. Item 23 (“To have goals in life and to know what I want to achieve”) focuses on well-defined life goals, and Item 34 (“To obtain academic achievement that enriches my curriculum vitae”) focuses on grades to enrich the personal curriculum vitae.

Only four non-target items loaded on the development factor (with values at least equal to .20). This factor did not receive the highest loading from any of these four items. The same occurred with the mobility factor, which only received one loading from a non-target item. The course factor also received loadings from five non-target items; the citizenship factor received loadings from four non-target items, and none of these items had the highest loading.

The living together factor presented a different pattern, because no non-target items loaded on it with a value equal to or above .20. The pressure factor was similar to the career factor, because the pressure factor also received some relevant loadings from non-target items (with values at least of .20). Two of these items are best explained by this factor than by any other factor: Item 1 (a target item of the career factor) loaded on the pressure factor with a value of .326, and Item 27 (a target item of the course factor) loaded on the pressure factor with a value of .392. Item 1 focuses on achieving a profession valued by society, and Item 27 focuses on achieving academic success to match society’s investment in the student. Both of these items seem to represent expectancies that are related to society and its demands.

Summing up, Model 2 shows that four items are better explained by factors other than their target factors, and the career and pressure factors better explain those items than their original target factors. Out of the total 42 items, 4 items (10% of the items) are not better explained by their target factor.

Beyond the fact that Model 1 shows unacceptable data fit and Model 2 shows a good data fit, there is another important difference between these two models. Because Model 1 constrains

Table 1
First-order factor loadings on items in model 2

Items	career	development	mobility	citizenship	pressure	course	living
1					.326		
2		.696					
3			.821				
4		.295		.427		.268	
5					.858		
6						.620	
7							.703
8	.516	.221				.226	
9		.716					
10			.909				
11				.686			
12					.545		
13	.294					.423	
14						.237	.301
15	.566	.228				.206	
16		.565					
17			.749				
18		.206		.461		.229	
19					.894		
20				.242		.599	
21							.683
22	.569						
23	.382	.250					
24	.316		.469				
25				.582			
26	.304				.328		
27	.263			.290	.392		
28							.867
29	.606						
30	.281	.437		.323			
31			.921				
32				.512			
33					.734		
34	.460					.241	
35							.925
36	.407		.215				
37	.259	.465		.321			
38			.860				
39				.758			
40					.550		
41	.213				.217	.307	
42							.881

Note: Career = training for job and career development, Development = personal and social development, Mobility = student mobility, Citizenship = political engagement and citizenship, Pressure = social pressure, Course = training quality, Living together = social interaction the loadings and Model 2 relaxes this condition, this produces a strong change in the values of the correlations among the first-order factors. Model 1 inflates the correlations among the factors, which is attenuated by Model 2. Table 2 shows the correlations among the factors in Model 1 and Model 2, as well as the difference of correlations of these two models.

Model 1 presents a mean correlation of .66 and a standard deviation of .14, and Model 2 presents a mean correlation of .40 and a standard deviation of .09. The difference of these means is .24, with a standard deviation of .14. This difference is remarkable because Model 1 increases the correlations of Model 2 by approximately 58%. This is a consequence of the constraint introduced in Model 1, where non-target items must have zero loadings on other factors. All these zero loadings inflate the first-order factors' correlations. In contrast, in Model 2, the occurrence of non-zero loadings of non-target items estimates the true correlations among the factors more correctly, because part of the estimated correlations among factors in Model 1 goes directly to the relation between the factors and the non-target items in Model 2. In other words, if the model does not allow non-target items to load on the factors, these loadings are automatically transferred to the correlations among the factors, inflating these correlations.

Discussion

This study showed the applicability of the ESEM in the educational field, employing this technique for a psychoeducational questionnaire about academic expectations. The model that constrained the non-target items to have zero loadings (IC-CFA) was refuted, reinforcing the large body of evidence showing that IC-CFA is too restrictive in the case of rating instruments. As commented by Marsh (2007), "it is almost impossible to get an acceptable fit (e.g., CFI, TLI >.90/RMSEA <.05) for even 'good' multifactor rating instruments when analyses are done at the item level and there are multiple factors (e.g., 5-10) […]" (Marsh, 2007, p. 785).

Besides obtaining a good data fit, the small cross-loadings in Model 2 were very important because they attenuated the inflated correlation among the first-order factors from Model 1, considerably improving the discriminant validity of the factors. Model 1 had inflated by 58% the mean correlations among factors. This result is in accordance with the results of previous studies, such as Ferrando and Lorenzo-Seva (2000), which state that the misfits in the measurement model in CFA concerning the correct relationships among the latent variables and the items bring an incorrect increase

Table 2
Correlation matrixes of first-order factors of model 1 and model 2 and their difference

	Development	Career	Mobility	Citizenship	Pressure	Course	Living together
Model 1							
Development							
Career	.881						
Mobility	.544	.540					
Citizenship	.790	.669	.580				
Pressure	.647	.656	.335	.508			
Course	.845	.892	.588	.846	.727		
Living together	.696	.661	.507	.597	.613	.646	
Model 2							
	Development	**Career**	**Mobility**	**Citizenship**	**Pressure**	**Course**	**Living together**
Development							
Career	.441						
Mobility	.397	.248					
Citizenship	.486	.356	.479				
Pressure	.450	.364	.253	.345			
Course	.546	.261	.469	.491	.374		
Living together	.533	.353	.494	.457	.543	.383	
Difference							
	Development	**Career**	**Mobility**	**Citizenship**	**Pressure**	**Course**	**Living together**
Development							
Career	.440						
Mobility	.147	.292					
Citizenship	.304	.313	.101				
Pressure	.197	.292	.082	.163			
Course	.299	.631	.119	.355	.353		
Living together	.163	.308	.013	.140	.070	.263	

Note: Career = training for job and career development, Development = personal and social development, Mobility = student mobility, Citizenship = political engagement and citizenship, Pressure = social pressure, Course = training quality, Living together = social interaction in the correlations of the latent variables. With such in mind, our results reinforce the claims of Marsh stating that "[…] strategies often used to compensate for these problems in CFA (e.g., parceling, ex post facto modifications such as ad hoc correlated uniquenesses) tend to be counterproductive, dubious, misleading, or simply wrong" (Marsh et al., 2014, p. 88). In other words, if parceling or aggregation of certain items could ameliorate the problem of cross-loadings, this strategy and others do not solve the problem of the factors' correlations and their positive bias in CFA.

Even being partly an exploratory strategy of analysis, ESEM can maximize its confirmatory part. The present study employed the strategy of maximizing the confirmatory approach in ESEM using target rotation, a recommendation of Marsh et al., who claim that the use of this kind of rotation is "based on partial knowledge of the factor structure, and is consistent with the view that ESEM is more typically used for confirmatory rather than exploratory purposes" (Marsh et al., 2014, p. 89).

Some interesting results have been revealed in the cross-loadings. Four items of the questionnaire did not obtain the highest loading in their respective theoretical target factors. However, as commented in the section of results, the two non-target items that had the highest loading on the career factor could be interpreted as having a conceptual relationship with career. The same is true for the two non-target items that obtained the highest loading on the pressure factor. The empirical results may provide new perspectives for these items, which were not considered previously. For practical reasons, it is important to allow the possibility of correlations among latent dimensions and behaviors (items) on psychological assessment techniques when in real life situations, those are normally expected. In those cases, the information obtained can increase the validity of psychoeducational instruments for practical uses if the internal factorial structure is more in accordance with reality or with the subjects' psychosocial functioning.

References

Almeida, L. S., Deaño, M., Araújo, A. M., Costa, A. R., Conde, A., & Alfonso, S. (2012). *Questionário de Perceções Académicas: Versão Expectativas (QPA-E)* [Questionnaire of Academic Perceptions: Expectations Version (QPA-E]. Braga: Universidad de Vigo-Ourense.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*(3), 397-438. http://dx.doi.org/10.1080/10705510903008204

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246. http://dx.doi.org/10.1037/0033-2909.107.2.238

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Deaño, M., Diniz, A., Almeida, L. S., Alfonso, S., Costa, A. R., García-Señorán, M., Conde, A., Araújo, A., Iglesias-Sarmiento, V., Gonçalves, P., & Tellado, F. (2015). Propiedades psicométricas del Cuestionario de Percepciones Académicas para la evaluación de las expectativas de los estudiantes de primer año en Enseñanza Superior [Psychometric properties of the Questionnaire of Academic Perceptions to assess 1st-year higher education students' expectations]. *Anales de Psicología*, *31*(1), 964-973. http://dx.doi.org/10.6018/analesps.31.1.161641

Ferrando, P. J., & Lorenzo-Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicológica, 21*(3), 301-323.

Furnham, A., Guenole, N., Levine, S. Z., & Chamorro-Premuzic, T. (2013). The NEO Personality Inventory-Revised: Factor structure and gender invariance from exploratory structural equation modeling analyses in a high-stakes setting. *Assessment, 20*(1), 14-23. http://dx.doi.org/10.1177/1073191112448213

García, J. A. M., & Caro, L. M. (2009). El análisis factorial confirmatorio y la validez de escalas en modelos causales [Confirmatory factor analysis and the validity of the measurement scales within a causal modelling framework]. *Anales de Psicología, 25*(2), 368-374.

Howard, J. L., Gagné, M., Morin, A. J. S., & Forest, J. (2016). Using bifactor exploratory structural equation modeling to test for a continuum structure of motivation. *Journal of Management*, 1-27. http://dx.doi.org/10.1177/0149206316645653

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

Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd edition, pp. 774-798). New York: Wiley.

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of Big-Five Factors over the life span: ESEM tests of gender, age, plasticity, maturity and la dolce vita effects. *Developmental Psychology*, *49*(6), 1194-1218. http://dx.doi.org/10.1037/a0026913

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 1,* 85-110. http://dx.doi.org/10.1146/annurev-clinpsy-032813-153700

Morin, A. J. S., & Maïano, C. (2011). Cross-validation of the short form of the Physical Self-Inventory (PSI-S) using exploratory structural equation modeling (ESEM). *Psychology of Sport and Exercise, 12*(5), 540-554. http://dx.doi.org/10.1016/j.psychsport.2011.04.003

Muthén, L. K., & Muthén, B. O. (1998-2014). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. London: Erlbaum.