Psicothema

Psicothema 2017, Vol. 29, No. 2, 236-240
doi: 10.7334/psicothema2016.304

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

Program FACTOR at 10: Origins, development and future directions

Pere Joan Ferrando and Urbano Lorenzo-Seva
Universitat Rovira i Virgili

Abstract

Background: We aim to provide a conceptual view of the origins, development and future directions of FACTOR, a popular free program for fitting the factor analysis (FA) model. **Method:** The study is organized into three parts. In the first part we discuss FACTOR in its initial period (2006-2012) as a traditional FA program with many new and cutting-edge features. The second part discusses the present period (2013-2016) in which FACTOR has developed into a comprehensive program embedded in the framework of structural equation modelling and item response theory. The third part discusses expected future developments. **Results:** at present FACTOR has attained a degree of technical development comparable to commercial software, and offers options not available elsewhere. **Discussion:** We discuss several shortcomings as well as points that require changes or improvements. We also discuss the functioning of FACTOR within its community of users.

Keywords: Exploratory factor analysis, semi-confirmatory factor analysis, item analysis, structural equation modelling, item response theory.

Resumen

10 años del programa FACTOR: una revisión crítica de sus orígenes, desarrollo y líneas futuras. **Antecedentes:** se pretende dar una visión conceptual del origen, desarrollos y futuras líneas de investigación de FACTOR, un popular programa no comercial de análisis factorial (AF). **Método:** el estudio se organiza en tres partes. En la primera se discute FACTOR en su etapa inicial (2006-2012) como un programa AF tradicional con opciones novedosas. En la segunda se discute la etapa actual (2013-2016) en la que FACTOR se presenta ya como un programa general enmarcado tanto en los modelos de ecuaciones estructurales como en la teoría de respuesta a los ítems. En la tercera parte, finalmente se discute la esperada evolución futura del programa. **Resultados:** en la actualidad FACTOR ha alcanzado un grado de desarrollo técnico comparable al software comercial, ofreciendo opciones no disponibles en otros programas. **Discusión:** se discuten algunas limitaciones, así como varios puntos que requieren cambios o mejoras. Se discute también el funcionamiento del programa dentro de la comunidad de usuarios.

Palabras clave: análisis factorial exploratorio, análisis factorial semi-confirmatorio, análisis de ítems, modelos de ecuaciones estructurales, teoría de respuesta al ítem.

We need both Exploratory and Confirmatory
(but we need Exploratory much more than Confirmatory)
John W. Tukey (1980)

FACTOR is a comprehensive, free, and user-friendly stand-alone program for fitting exploratory (EFA) and semi-confirmatory (SCFA) factor analytic (FA) models. Since it was initially proposed 10 years ago (Lorenzo-Seva & Ferrando, 2006) the program has become quite popular and has been used in such different domains as Ornithology (Leveau, 2013) and Animal Genetics (Parés Casanova, Sinfreu Blasi, & Villalba Mata, 2012). In spite of its versatility, however, FACTOR was mainly designed for psychometric applications and, more specifically, for item analysis and individual scoring purposes (e.g., Izquierdo, Olea, & Abad, 2014; Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco 2014). In *Psicothema*, for example, 23 FACTOR-based psychometric applications have been published to date (e.g., Peña-Suárez, Muñiz, Fonseca-Pedrero, & García-Cueto, 2013; Fernández, Cueto, Vázquez, & González, 2012; Menéndez, García, & Viejo, 2010).

Received: October 4, 2016 • Accepted: December 22, 2016
Corresponding author: Pere Joan Ferrando
departamento Psicologica
Universitat Rovira i Virgili
43007 Tarragona (Spain)
e-mail: perejoan.ferrando@urv.cat

Between 2006 and 2016 FACTOR has evolved continuously, going from release 5.1 to release 10.4. More important than the continuous updates, however, has been its general evolution, which has transformed it from an almost traditional EFA program with a few extra features, to a comprehensive cutting-edge program that includes exploratory and semi-confirmatory procedures, and is embedded in the frameworks of structural equation modelling (SEM) and Item Response Theory (IRT). In this article we aim to provide a critical review of the origins, evolution, and future directions of FACTOR, and a discussion on the rationale and background for the procedures it implements. The review is conceptual rather than technical and is expected to be useful for both current and (hopefully) future users.

We decided to create and develop FACTOR mainly as the result of our stand on the exploratory vs. confirmatory controversy and our dissatisfaction with the technical treatment generally given to EFA by standard programs. This treatment is very poor and reflects the still dominant position that CFA is the way to go, and that EFA is, at best, a rough precursor of CFA that is useful only when there is no ‘a priori’ assumption regarding the structure of the item responses (e.g., Bollen, 2002). Our position, however, is that modelling item responses using a fully confirmatory solution in which all the items behave as markers of a single factor (i.e. a strict independent-cluster solution, McDonald, 2000) is unrealistic, especially with personality and attitude items. And, if an overly restrictive solution of this type is forced on the data, two potential main problems are expected to arise: (a) poor fit, especially as the test becomes longer and the sample increases, and (b) biased parameter estimates, particularly the inter-factor correlations (see Ferrando & Lorenzo-Seva, 2000). Overall, we feel that most of the items found in practice are factorially complex (e.g. Cattell, 1986), and that a flexible EFA or a SCFA solution is generally much more appropriate and ‘natural’ than a strict CFA solution.

With regards to the second point above our position is methodologically standard: EFA is a structural equation model that can be fitted and assessed in the same ways as any other model of this type (including, indeed, the CFA model). This obvious position however, is not usually reflected in EFA implementations. While CFA solutions can be fitted with such sophisticated procedures as robust maximum likelihood or weighted least squares, and assessed with goodness-of fit indices such as RMSEA, CFI, GFI or NNFI, EFA solutions generally have to make do with principal axis extraction procedures and approximate model-data fit criteria such as ‘proportion of explained variance’ or ‘eigenvalues greater than one’. This ‘discriminatory’ treatment tends to widen the gap between the ‘technically sophisticated’ CFA model and the ‘second class’ approximate EFA model. This gap, however, is patently false.

The first period: 2006–2012

The structure of FACTOR follows the conventional structure of the FA modules implemented in packages such as SPSS, BMDP, SAS or SYSTAT. It has six parts: (a) descriptive statistics, (b) assessment of sampling adequacy, (c) estimation of the factor parameters (factor loadings and inter-factor correlations), (d) transformations of the structural solution (i.e. rotations), (e) goodness-of-fit assessment, and (f) individual scoring.

Although the above organization is quite standard, each of the parts in FACTOR had novelties which made FACTOR a potentially better option. For example, in (c) FACTOR included not only the standard estimation procedures – principal axis factoring (PAF), unweighted least squares (ULS) and maximum likelihood (ML) – but also minimum-rank factor analysis (MRFA). In (d) as many as 30 analytical rotation procedures were available (16 orthogonal and 14 oblique; we are not aware of any programs with so many). In (e) FACTOR included three auxiliary procedures – the minimum average partial (MAP), Optimal Implementation of Parallel Analysis (Timmerman & Lorenzo-Seva, 2011), and the Hull method (Lorenzo-Seva, Timmerman, & Kiers, 2011) – as well as three proper goodness-of-fit indices: the chi-square test (for both ML and ULS estimation), the gamma-goodness of fit index (GFI) and the root mean square of the residuals together with the Thurstone-Kelley critical value (Kelley, 1935, page 146). We also included several analytical and graphical descriptives of the distribution of residuals. Finally, in (f) FACTOR provided factor scores for each respondent and the marginal reliability estimate for these scores.

Most of the extra features mentioned above were the result of the psychometric orientation of the program, and our own ideas on FA, so some comments about them are in order. First, the main advantage of including MRFA is that it allows the percentage of explained common variance to be estimated. In principle, we consider the proportion of explained variance to be a useful auxiliary criterion for judging model adequacy. However, the explained proportion of total variance based on principal component analysis (the usual criterion) is potentially misleading in the case of test items which generally have a large amount of measurement error. So, if the user wants to report the proportion of explained variance, then it makes sense that this should at least be the common proportion.

With regards to the overwhelming number of rotation methods, as we have explained above we believe that items are, inherently, factorially complex. And we also believe that, in the presence of a complex structure, it is worth attempting to reach the maximally simple solution in Thurstonian terms (Lorenzo-Seva, 2003). Now, no single analytical rotation procedure automatically achieves this, so it is useful to try different methods that use different criteria. As a default, FACTOR uses Promin (Lorenzo-Seva, 1999) which usually achieves considerable simplicity without being computationally too demanding.

Let us now discuss the emphasis placed on goodness-of-fit assessment. Because we consider FA to be a SEM, we are strongly against the approximate ‘rules of thumb’ that still plague the applied field on the key issue of deciding the most appropriate number of factors. And we believe that the best starting point in this respect is to use the same rationale as that used in SEM. In this regard, we fully agree with McDonald (2000; McDonald & Ho, 2002) that a careful inspection of the residuals, the RMSR and the GFI is the best approach for assessing the appropriateness of the FA solution under any estimation procedure. To see why inspecting the residuals is necessary, consider that misfit may be caused not only by a small number of model misspecifications that give rise to a few large discrepancies but also by a wide general scatter of discrepancies not associated with any particular misspecification. A single scalar index is unable to inform us exactly what the cause is. Only a detailed inspection of the residuals can.

Finally, let us take a look at scoring. As the literature reveals, most FA applications based on item responses are solely concerned with the estimation of the structural parameters. Once a satisfactory item set has been obtained, sum scores are then computed, and the remaining analyses use a descriptive, classical test-theory approach based on these sum scores. This practice is not incorrect, but is far from optimal because FA-based scores are generally more informative and accurate than simple sum scores (Ferrando & Lorenzo-Seva, 2016). So, from the beginning, we tried to promote the use of factor scores in FACTOR as a natural second-step analysis once the structural part of the model had been fitted.

As well as the points summarized above, the main feature that distinguished FACTOR from the existing FA programs was that two types of modelling were available: standard FA based on the product-moment correlation matrix, and FA based on the tetrachoric/polychoric correlation matrix. This feature reflected the psychometric orientation of FACTOR, because most item scores are ordered categorical with few response points.

The controversy regarding the appropriateness of one modelling system or the other when fitting test items goes back more than 75 years in the psychometric literature and is still alive nowadays. Our position on this issue has been discussed in several publications (see e.g., Ferrando & Lorenzo-Seva, 2014) and will be summarized here only briefly. We believe that the greater appropriateness of one modelling or the other is an empirical matter that depends on such conditions as the distribution and discriminating power of the items, the size of the sample, and the number of response points. So, it is the researcher's task to carefully assess his/her dataset in order to decide which modelling is 'a priori' most appropriate (see e.g. Lloret-Segura et al. 2014).

A less debated aspect of which product-moment vs. polychoric FA modelling is most relevant concerns its relations with IRT modelling. Fitting the standard FA model to the product-moment inter-item correlation matrix is analogous to fitting a linear IRT model intended for continuous scores (Ferrando, 2009a). On the other hand, the standard FA modelling of the inter-item tetrachoric correlation matrix is equivalent to fitting the 2-parameter normal-ogive IRT model, while the FA modelling of the polychoric matrix is equivalent to fitting the normal-ogive version of Samejima's graded-response model (see e.g., Ferrando & Lorenzo-Seva, 2013, 2014; Mislevy, 1986; or Reckase, 2009).

So, the estimation of the structural FA parameters and subsequent rotation process when the input variables are item scores can be viewed as a calibration process in which a linear or a non-linear IRT model is fitted to the item responses. It was this view that prompted many further developments in FACTOR because we felt that the calibration procedures available in this first period could clearly be improved. For various reasons this point was more evident in the tetrachoric/polychoric modelling. First, the estimation of the tetrachoric/polychoric correlations showed convergence problems and led to unstable estimates in many cases. Second, the smoothing procedures implemented when the matrix was non-Gramian were not satisfactory (Devlin, Gnanadesikan, & Kettenring, 1975, 1981). Finally, and more importantly, PAF, ULS or ML FA estimation based on tetrachoric/polychoric correlations is not a rigorous way of fitting the corresponding IRT model, but only a rough, approximate procedure known as the 'heuristic approach' (Bock & Aitkin, 1981). Overall, we were aware that the EFA estimation based on programs such as NOHARM (for binary items) or POLYFACT were better than those offered by FACTOR.

The second period: 2013 to the present

A series of major modifications that led to version 9 justified a second publication of FACTOR in a program announcement (Lorenzo-Seva & Ferrando, 2013) and this can be regarded as the beginning of the second period. Overall, it is in this period that the main changes mentioned above occurred: (a) implementation of SCFA approaches, (b) developments based on IRT modelling including alternative parameterizations and improved factor scores, and (c) developments based on SEM including robust estimation procedures, scaled fit indices, and Bootstrap-based confidence intervals for virtually all the estimates produced by the program.

Let us start with the SCFA developments. Since about 2010, certain changes started to take place in the EFA-CFA controversy. These included (a) a certain amount of dissatisfaction with the results of using strict CFA approaches, (b) acknowledgment of the problems derived from this use, and (c) 'new' proposals of more flexible forms of modelling (e.g. Marsh, Morin, Parker, & Kaur, 2014). We had already discussed these points in detail more than 10 years previously (Ferrando & Lorenzo-Seva, 2000) but had not been very successful in communicating this discussion.

The new methodological SCFA approach mentioned above was labelled as exploratory SEM (ESEM) and was presented as a new alternative (see e.g., Marsh et al., 2014). And, in some respects, ESEM was indeed new. However, workable and flexible SCFA approaches based on minimal-identification constraints (Howe, 1955) or on target rotations (Browne, 1972) have been available for a long time now. The starting approach we chose for SCFA in FACTOR was Procrustes rotation against a semi-specified target (Browne, 1972). Both orthogonal and oblique rotations can be performed, and this general methodology allows several interesting SCFA solutions to be obtained. The most important are: canonical solutions (useful for investigating the dimensionality of an item set), exploratory bi-factor solutions (e.g., Reise, 2012; compatible with a general factor structure and several group factor structures), and independent cluster-based solutions (McDonald, 2000; the SCFA counterpart of the standard CFA solution). Ferrando and Lorenzo-Seva (2013) provide details on how these solutions can be specified.

We turn now to the IRT-related developments. First, we implemented improved estimation procedures for the tetrachoric/polychoric correlations based on a Bayes approach, and, as we shall describe below, we also implemented new and better general estimation procedures. As a result, item calibration is no longer based on an approximate heuristic approach but on a technically defensible limited-information approach which provides correct standard errors and goodness-of-fit statistics. Second, we provided the alternative IRT parameterization for the uni- and multidimensional two-parameter normal-ogive and graded response models, including the multidimensional difficulty, discrimination and information indices (e.g., Ferrando, 2009a; Reckase, 2009). The main IRT developments in this period, however, have been the improvement in the estimation of factor scores and related person-fit indices. The scores implemented in the first-period version were regression scores preserving inter-factor correlations (Ten Berge, Krijnen, Wansbeek, & Shapiro, 1999), an approach that treated all the responses as if they were continuous. In the second period, however, we implemented a general Bayesian approach in which factor scores are estimated using the same variable treatment as that used in the calibration stage (i.e. binary, graded-ordered or continuous). Furthermore, these estimates use the maximum information available from the calibration stage (including inter-factor correlations). Individual standard errors and reliability estimates, as well as the marginal reliability estimate are provided in the output (see Ferrando & Lorenzo-Seva, 2016, for details).

Factor scores can only be validly interpreted if the response pattern on which they are based is consistent with the FA model. This type of consistency is assessed by person-fit indices, which, when used with reference to critical values, allow the researcher to flag the individuals for whom valid interpretations of their scores is not warranted (see Ferrando, 2015). At present, FACTOR implements only the *lcz* index proposed by Ferrando (2007, 2009b) but we are exploring new possibilities.

Finally, we shall discuss the main SEM-related developments. The first was the implementation of new estimation procedures and robust goodness-of-fit indices. In brief, the following new options are available in FACTOR. For categorical variables: robust ULS estimation, and robust Diagonally Weighted Least Squares (DWLS) estimation with mean-corrected fit indices. For continuous variables, robust ULS, robust DWLS, and robust ML estimation with mean-corrected and mean-and-variance corrected fit indices. With these new options, EFA and SCFA models can now be fitted by FACTOR with the same degree of technical sophistication as any CFA model in a commercial SEM program.

The second main development was the implementation of intensive re-sampling (Bootstrap) procedures for virtually all the estimates computed with the program. Bootstrap re-sampling allows (a) robust estimation to be performed even in small-medium samples by using an empirical estimate of the asymptotic covariance matrix, and (b) obtaining standard errors and confidence intervals for any estimate of interest (e.g. loadings, factor scores, reliability estimates, goodness-of-fit indices, etc.). The third main development, finally, was to properly handle missing responses using Hot-Deck Multiple Imputation (see Lorenzo-Seva & Van Ginkel, 2016, for details).

Criticisms and future directions

We have tried to provide a coherent framework for discussing the origins and evolution of FACTOR. However, many features implemented mainly reflect the somewhat disparate interests of one or other of the authors in different periods. The result is that FACTOR is not completely the coherent and balanced program that we have suggested it is above. Sometimes these points cause confusion among the users of FACTOR, as we are aware by the considerable number of queries that we receive via email.

Apart from the general criticism above, there are many points that can (and should) be improved. One of these points is the estimation of the FA model based on covariance or moment matrices. In this respect, FACTOR is a traditional EFA program that focuses on the analysis of correlation matrices. However, the linear calibration of item scores treated as continuous variables is often better when based on means and covariances than on correlations (Ferrando, 2009a). Furthermore, the feasibility of fitting covariance matrices with EFA will contribute to fill the (false) gap between CFA and EFA. At present, FACTOR can to some extent fit covariances using ULS estimation but there is considerable room for improvement. Another point that needs to be improved concerns the smoothing procedures needed for non-Gramian correlation matrices, which are still not satisfactory: when the negative eigenvalues are large, the smoothing procedure destroys most of the information in the correlation matrix, and the EFA is virtually meaningless. Finally, a third point is the organization of the reliability estimates provided by the program, which should be improved. Many users are confused about which estimates correspond to the sum or raw scores (alpha, omega, GLB) and which to the factor scores (PSD-based and marginal estimates).

We turn finally to some new procedures and techniques that we (hopefully) plan to develop and include in the near future. At the calibration level we would like to implement (a) Multiple-group exploratory and SC factor analysis procedures, and (b) estimation procedures based on minimal identification constraints as a second approach for fitting SCFA solutions. Overall, we feel that the feasibility of (a) fitting FA models based on covariance matrices, (b) performing multiple-group analysis, and (c) specifying direct solutions with no need of rotation, will (almost) close the false gap between EFA and CFA discussed in the article. It will also give the researcher the flexibility and technical sophistication to perform item analysis, something which is not given by any current FA program. At the scoring level, we aim to implement other scoring schemas apart from Bayes EAP, and, in particular, robust estimation techniques such as the Biweight (Mislevy & Bock, 1982). We also plan to determine which person-fit indices are the simplest and perform the best, and implement them in the program so that they can be used as a routine check to identify inconsistent respondents.

As a final reflexion, the main challenge we encounter regarding the issues discussed in this section is how to incorporate the modifications, improvements and new advances, and, at the same time, maintain the general scope, clarity and simplicity of FACTOR, especially for those who use the program for non-psychometric purposes.

Discussion

As discussed in the article, FACTOR was created and developed mainly in response to a state of affairs that the authors perceived to be wrong and misleading. However, there have also been more objective aims that have guided the development of the program. Above all, we wanted the program to be stand-alone and free. Furthermore, we needed to achieve (a) flexibility and versatility to suit the varied interests of all potential users, (b) calculation power so that FACTOR could run relatively fast even under heavy computing demands (specially when Bootstrap re-sampling was used), and (c) a degree of correctness and accuracy comparable to that of commercial programs. We are far from being fully satisfied with FACTOR as it is now, but we do feel that we are gradually achieving our objectives.

The practical relevance of all the developments we are implementing in FACTOR will mainly depend on the general evolution of EFA and SCFA in the future. At present, the leading journals (and their reviewers) still request CFA solutions in applied studies, and only reluctantly accept EFA and SCFA solutions as auxiliary devices if they are very well justified. However, there are also some signs of change. At the applied level, practitioners are becoming increasingly aware of the limitations and rigidity of pure CFA solutions. And, on the methodological field, there is a rediscovery of 'old' FA solutions such as the Bi-Factor, the Multiple Group, and the Independent-Cluster basis (see e.g., Ferrando & Lorenzo-Seva, 2012), all of which are SC. Furthermore, the new versions of these solutions benefit from the advances that have taken place in the CFA domain, mainly: better estimation procedures and improvements in model-data fit assessment. To sum up, we do not believe that EFA and SCFA will become again the dominant paradigm in FA, but we feel that they will regain terrain, and so, in our opinion, both are alive and worth of future methodological developments, especially SCFA. In this respect, there is still ample room for technical improvements, as well as extensions, mainly to the multiple-group data, longitudinal data, and individual-level assessment (e.g., Cudeck & MacCallum, 2007).

In finishing, this article might lead to the (erroneous) impression that FACTOR is the work of just two authors, who have also had to go against the flow. This is not so, and we like to think that FACTOR is more a community of users than simply a free program. For example, the existing manuals and video tutorials have been developed by users. We also receive plenty of suggestions for future improvements, many of which have already been implemented (we would like to be able to incorporate all the suggestions, but there is only so much we can do). We are infinitely grateful to the faithful researchers who put their trust in our software, and the methodological improvements implemented (many of which reflect our positions and/or were proposed by us). We are glad to know that we are helping other researchers to carry out their own research and, to be honest, the periodical report on FACTOR references simply makes our day.

Acknowledgements

The research was supported by a grant from the Catalan Ministry of Universities, Research and the Information Society (2014 SGR 73) and by a grant from the Spanish Ministry of Economy and Competitivity (PSI2014-52884-P).

References

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of the EM algorithm. *Psychometrika,* 46(4), 443-459.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*(1), 605-634.

Browne, M. (1972). Oblique rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology*, *25*(2), 207-212.

Cattell, R. B. (1986). The psychometric properties of tests: consistency, validity and efficiency. In R. B. Cattell & R. C. Johnson (Eds.), *Functional Psychological Testing* (pp. 54-78). New York: Brunner/Mazel.

Cudeck, R., & MacCallum, R. J. (2007). *Factor analysis at 100: Historical developments and future directions*. London: LEA.

Devlin, S. J., Gnanadesikan, R., & Kettenring, J. R. (1975). Robust estimation and outlier detection with correlation coefficients. *Biometrika*, *62*(3), 531-545.

Devlin, S. J., Gnanadesikan, R., & Kettenring, J. R. (1981). Robust estimation of dispersion matrices and principal components. *Journal of the American Statistical Association*, *76*(374), 354-362.

Fernández, L. M. L., Cueto, E. G., Vázquez, M. M., & González, L. L. (2012). Desarrollo y validación del Inventario de Perfeccionismo Infantil (IPI). *Psicothema*, *24*(1), 149-155.

Ferrando, P. J. (2007). Factor-analytic procedures for assessing response pattern scalability. *Multivariate Behavioral Research*, *42*(3), 481-507.

Ferrando, P. J. (2009a). Difficulty, discrimination, and information indices in the linear factor analysis model for continuous item responses. *Applied Psychological Measurement*, *33*(1), 9-24.

Ferrando, P. J. (2009b). Multidimensional factor-analysis-based procedures for assessing scalability in personality measurement. *Structural Equation Modeling*, *16*, 109-133.

Ferrando, P. J. (2015). Assessing person fit in typical-response measures. In S.P. Reise & D.A. Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (pp. 128-155). New York: Routledge/Taylor & Francis Group.

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.

Ferrando, P. J., & Lorenzo-Seva, U. (2013). *Unrestricted item factor analysis and some relations with item response theory*. *Technical Report*. Department of Psychology, Universitat Rovira i Virgili, Tarragona. Retrieved from *http://psico. fcep. urv. es/utilitats/factor*

Ferrando, P. J., & Lorenzo-Seva, U. (2014). El análisis factorial exploratorio de los ítems: algunas consideraciones adicionales. *Anales de Psicología*, *30*(3), 1170-1175.

Ferrando, P. J., & Lorenzo-Seva, U. (2016). A note on improving EAP trait estimation in oblique factor-analytic and item response theory models. *Psicologica*, *37*(2), 235-247.

Howe, W. G. (1955). *Some contributions to factor analysis* (No. ORNL-1919). Oak Ridge National Lab., Tenn..

Izquierdo, I., Olea, J., & Abad, F. J. (2014). El análisis factorial exploratorio en estudios de validación: usos y recomendaciones. *Psicothema, 26*(3), 395-400.

Kelley, T. L. (1935). *Essential Traits of Mental Life, Harvard Studies in Education*, vol. 26. Harvard University Press, Cambridge.

Leveau, L. M. (2013). Bird traits in urban-rural gradients: How many functional groups are there? *Journal of Ornithology*, *154*(3), 655-662.

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). Exploratory Item Factor Analysis: A practical guide revised and up-dated. *Anales de Psicología*, *30*(3), 1151-1169.

Lorenzo-Seva, U. (1999). Promin: A method for oblique factor rotation. *Multivariate Behavioral Research*, *34*(3), 347-356.

Lorenzo-Seva, U. (2003). A factor simplicity index. *Psychometrika*, *68*(1), 49-60.

Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, *38*(1), 88-91.

Lorenzo-Seva, U., & Ferrando, P. J. (2013). FACTOR 9.2: A Comprehensive Program for Fitting Exploratory and Semiconfirmatory Factor Analysis and IRT Models. *Applied Psychological Measurement*, *37*(6), 497-498.

Lorenzo-Seva, U., & Van Ginkel, J. R. (2016). Multiple Imputation of missing values in exploratory factor analysis of multidimensional scales: Estimating latent trait scores. *Anales de Psicología*, *32*(2), 596-608.

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H.A.L. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research*, *46*(2), 340-364.

Marsh, H. W., Morin, A. J., 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*, *10*, 85-110.

McDonald, R. P., & Ho, R. M. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*(1), 64-82.

McDonald, R. P. (2000). A basis for multidimensional item response theory. *Applied Psychological Measurement, 24*(2), 99-114.

Menéndez, A. G., García, P. F., & Viejo, I. T. (2010). Aceptación del dolor crónico en pacientes con bromialgia: adaptación del Chronic Pain Acceptance Questionnaire (CPAQ) a una muestra española. *Psicothema, 22*(4), 997-1003.

Mislevy, R. J. (1986). Recent developments in the factor analysis of categorical variables. *Journal of Educational Statistics, 11*(1), 331.

Mislevy, R. J., & Bock, R. D. (1982). Biweight estimates of latent ability. *Educational and Psychological Measurement*, *42*(3), 725-737.

Parés Casanova, P. M. P., Sinfreu Blasi, I., & Villalba Mata, D. (2012). Principal component analysis of cephalic morphology to classify some Pyrenean cattle. *Animal Genetic Resources*, *50*, 59-64.

Peña-Suárez, E., Muñiz, J., Fonseca-Pedrero, E., & García-Cueto, E. (2013). Assessing organizational climate: Psychometric properties of the CLIOR Scale. *Psicothema*, *25*(1), 137-144.

Reckase, M. (2009). *Multidimensional Item Response Theory*. New York: Springer.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667-696.

Ten Berge, J. M. F., Krijnen, W., Wansbeek, T., & Shapiro, A. (1999). Some new results on correlation-preserving factor scores prediction methods. *Linear Algebra and its Applications*, *289*(1), 311-318.

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, *16*(2), 209-220.

Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, *34*(1), 23-25.