Normal science and its tools: Reviewing the effects of exploratory factor analysis in management

Luciano Rossoni

Universidade do Grande Rio – Rio de Janeiro/RJ, Brasil

Ricardo Engelbert

ISE Business School – São Paulo/SP, Brasil

Ney Luiz Bellegard

Universidade Positivo – Curitiba/PR, Brasil

RESUMO

Ciência normal e suas ferramentas: Revisando os efeitos dos métodos de análise fatorial exploratória em Administração

O objetivo neste estudo foi investigar como diferentes métodos de extração, de definição de fatores e de rotação da análise fatorial exploratória afetam o ajuste de escalas de mensuração. Para tanto, foi feita uma meta-análise de 23 estudos, na qual os resultados apontaram que o método de Componentes Principais proporciona maior variância explicada, enquanto o método de Máxima Verossimilhança aumenta a confiabilidade. Entre as rotações, a Varimax fornece maior confiabilidade e o Quartimax a menor correlação entre fatores. Na conclusão do trabalho são destacadas as implicações para a pesquisa quantitativa, assim como sugeridos novos estudos.

Palavras-chave: análise fatorial exploratória, confiabilidade, métodos quantitativos, *survey*.

Luciano Rossoni, Professor do Programa de Pós-Graduação em Administração da Universidade Grande Rio (CEP 20021-180 – Rio de Janeiro/RJ, Brasil).
E-mail: lrossoni@gmail.com
Address:
Universidade Grande Rio
Rua da Lapa, 86
20021-180 – Rio de Janeiro, RJ – Brasil

Ricardo Engelbert, Professor do ISE Business School (CEP 01321-001 – São Paulo/SP, Brasil).
E-mail: ricardo@engelbert.com.br

Ney Luiz Bellegard, Doutor em Administração pela Universidade Positivo (CEP 80250 060 – Curitiba/PR, Brasil), Consultor de Empresas.
E-mail: neybellegard@gvmail.br

1. INTRODUCTION

However heterogeneous, eclectic, and diverse the paradigms in the social sciences, it cannot be denied that some tend to be more visible in the academic community. The field of administration is no different: some themes and perspectives are more easily accepted, some theories are considered to be legitimate, and some rules and methodological procedures are recognized as being valid. One in particular, however much criticism it receives and however many limitations it has, is dominant in organizational analysis: the quantitative

research paradigm. It is not difficult to confirm its ubiquity in the main international management journals, and likewise in the most prominent periodicals in the administration field in Brazil, since most theoretical-empirical studies are of a quantitative nature.

Since it is a recurring paradigm, a good part of its procedures and techniques are clearly defined and its rules are relatively well-accepted by those researchers who have mastered it, which makes it similar to what (Kuhn, 2009) called "normal science". Since its earliest days (Cronbach & Meehl, 1955; Loevinger, 1957), the central undertaking of this research tradition has been to develop valid and reliable scales based on measuring phenomena that are inherently subjective or social. To do so researchers use data collection instruments, typically questionnaires, in which questions or statements are scored with the aim of quantitatively assessing a particular phenomenon, which is measured in different degrees on a scale (e.g. the Likert scale). In the jargon of quantitative research, these phenomena are called constructs, and one of their fundamental assumptions is their latent nature, which leads to the need to use various indicators or items for indirectly accessing them (Netemeyer, Bearden, & Sharma, 2003).

Problems inherent in the operationalization of variables by way of multiple items arise, on the one hand, when a social phenomenon is reduced to numerical scales; and on the other because of the distance between the theoretical concept and the empirical evaluation of the phenomenon (Cronbach & Meehl, 1955; Loevinger, 1957; Netemeyer *et al.*, 2003). But despite being extremely important, these substantive questions are not the only ones to affect the quality of the measuring instruments: questions of a technical nature related to the statistical procedures used for dealing with scales are also fundamental, so much so that texts like the one by (Churchill, 1979) are wholly dedicated to deliberating about a measures development paradigm.

Despite some of the problems and controversies that exist in the procedures deliberated upon by authors like Churchill [*vide* Smith's (1999) criticism], their importance when it comes to consolidating systematic procedures in the construction of scales must be recognized, especially because such "manuals" were fundamental for spreading the use of the statistical technique for constructing scales that is generically labeled "exploratory factor analysis". So if today we can understand quantitative analysis in social sciences to be a paradigm, we can see exploratory factor analysis as one of its research tools. Just like a telescope, these tools need adjustments and improvements, such as specifications and tests that set the limits of their use.

In view of the fact that the current rule is to assess any phenomenon by way of various items, exploratory factor analysis helps the researcher identify first of all how many dimensions a construct has, and secondly to fit each one of the items into the dimension most directly related to them.

Therefore, after assessing the dimensionality of the construct, an attempt is made to check the extent to which these dimensions are internally consistent, or reliable. The problem is that there are various options for adjusting this technique (e.g. various extraction and rotation methods), which end up raising doubts in researchers' minds as to which is most appropriate for their use. Despite some interesting studies that discuss these questions (Aranha & Zambaldi, 2008; Guadagnoli & Velicer, 1988; Stevens, 2009), the objective of our study is to contribute towards improving the use of these tools based on a meta-analysis of twenty-three articles that were published in the administration area in Brazil, as we seek to understand how different extraction, factor definition, and rotation methods affect the fit of exploratory factor analysis.

Following this introduction there are four parts to the structure of this article: in the first we review the exploratory factor analysis procedure and its stages, where we highlight some of the issues and questions that have guided this empirical research. We then present the methodological procedures, by providing details of the data collection work, the creation of indicators, and the statistical analysis method we used. Next, we give the results of the research, by answering the questions raised in the theoretical framework. Finally, we discuss the implications of our findings for developing and refining scales in practice.

2. NOTES ON EXPLORATORY FACTOR ANALYSIS

Exploratory factor analysis (EFA) is a multivariate interdependence technique that is widely used in research in the field of administration, especially research of the survey type, and which has two primary purposes. The first is to obtain a minimum number of factors that contain the maximum possible amount of information contained in the original variables used in the model, and with the greatest possible reliability (Hair, Black, Babin, Anderson, & Tatham, 2009; Johnson & Wichern, 2007; Netemeyer *et al.*, 2003). This reduction in the number of variables is desirable when it is intended to submit the data to other multivariate analysis techniques, in which there can be no strong correlations between the independent variables, as is the case with regression techniques, thus generating a more parsimonious model (Hair *et al.*, 2009; Johnson & Wichern, 2007). Although there may be a correlation between these factors, factor analysis guarantees a concentration of the information from the original variables (Aranha & Zambaldi, 2008; Hair *et al.*, 2009). The second purpose, which is related to the first, is to identify how indicators used empirically are configured in factors that are not directly observed, representing the facets or dimensions of the phenomenon being investigated (Johnson & Wichern, 2007). In short, an attempt is made to identify how many dimensions a construct has (Netemeyer *et al.*, 2003), which is the most relevant decision a researcher has to make when carrying out factor analysis (Johnson & Wichern, 2007). It is important to emphasize that factor analysis cannot

perform miracles when it creates factors or dimensions. The data from the observable variables must truly represent the phenomenon being investigated; if not, the factors obtained will be statistically consistent, but irrelevant in relation to the study object (Hair *et al.*, 2009).

Another type of factor analysis is confirmatory analysis, which aims to assess the degree to which the data satisfy a particular conceptual structure extracted from a theory (Hair *et al.*, 2009; Rencher & Christensen, 2012; Tabachnick & Fidell, 2013). While in exploratory factor analysis we statistically explore the number of factors that best fit the data, in confirmatory factor analysis the factors and their respective indicators are defined *a priori* according to the theoretical model (Hair *et al.*, 2009). Confirmatory analysis is also characterized as being an interdependence technique, because we do not define any type of dependence relationship between the variables used and the resulting factors (Bezerra, 2009; Field, 2009; Netemeyer *et al.*, 2003). To meet the objectives of this study, however, we shall be focusing solely on exploratory factor analysis.

3. STAGES OF EXPLORATORY FACTOR ANALYSIS

Rather than a mere technique, we can visualize exploratory factor analysis as a procedure, which is contained in the scalar development paradigm (Churchill, 1979) and limited by a series of stages, each of which has a specific set of rules. The first involves an analysis of the requirements in terms of sample conception and the characteristics of the variables. The second specifies the adjustment tests. The third refers to the analysis of the number of dimensions, in which there are different methods for extracting the dimensions, such as different factor definition methods. The fourth stage involves the rotation of the variables in different factors, with the objective of obtaining a better fit or adjustment. Finally, in the last stage, dimension reliability is assessed.

Requirements. Reviewing the requirements for designing research that uses EFA, as cited by important authors in the area (Field, 2009; Hair *et al.*, 2009), we have: (a) the use of metric variables (because their correlations need to be calculated) with a normal distribution: a certain collinearity is expected because of the correlation between variables; (b) there must be at least five variables for representing each factor; (c) the sample must be homogenous and have between at least ten and fifteen times more observations than the number of variables; (d) this sample must have no fewer than fifty observations, and preferably more than one hundred cases; and obviously, (e) there must be an underlying structure in the set of variables analyzed. In cases in which factor analysis is developed using categorical variables, which is not the case with this study's analysis, a further two demands are added: (f) the categorical variables must be transformed into dichotomous variables (dummies); and (g) these dichotomous variables must not be present in any large number.

One aspect to be highlighted in the first requirement is the assumption of univariate and multivariate normality, which also refers to assumptions about the independence of the observations and the randomness of the sample (Bentler & Chou, 1987; Damásio, 2012; Hair *et al.*, 2009). It is common, however, for factors that are extracted by means of Likert-type scales not to have a normal distribution (Curran, West, & Finch, 1996), so much so that some authors even suggest checking whether the indicators have asymmetrical and unbalanced distributions (Laros, 2011; Lopes, 2005), because normality tests, like the Shapiro-Wilk test, may perform weakly in the multivariate normality assessment (Cantelmo & Ferreira, 2007). In the case of samples with multivariate normality problems, Curran *et al.* (1996), Damásio (2012), and Laros (2011) point out that some factor extraction methods, especially the *maximum likelihood* method, may give results that are compromised. Hypothetically this would not prevent the use of this extraction method when the assumptions of randomness and independence are maintained (Bentler & Chou, 1987; Curran *et al.*, 1996), despite several authors agreeing that *principal components* analysis is the best alternative when there is no multivariate normality (Curran *et al.*, 1996; Damásio, 2012; Hair *et al.*, 2009; Laros, 2011; Lopes, 2005).

Factor adjustment. Checking whether the data are sufficient for a stable factors' solution is normally carried out by way of the **Kaiser-Meyer-Olkin (KMO)** measure of sampling adequacy, which considers the variance proportion of the indicators that can be explained by a latent variable (Lorenzo-Seva, Timmerman, & Kiers, 2011), and "represents the ratio of the squared correlation between variables to the squared partial correlation between variables" (Field, 2009, p. 571). As the values of the test vary from 0 to 1, values above 0.7 are recommended as being desirable for applying EFA (Damásio, 2012; Hair *et al.*, 2009). Another widely used test is **Bartlett's sphericity test**, which examines the whole correlation matrix to determine the adequacy of factor analysis based on identifying the correlation between variables. "It supplies the statistical significance that the correlation matrix has significant correlations between at least some of the variables" (Hair *et al.*, 2009, p. 110), and becomes more sensitive as the size of the sample increases. A statistically significant Bartlett test ($p < 0.05$) indicates that sufficient correlations exist between the variables to continue with the analysis. As the number of cases decreases (Johnson & Wichern, 2007), this test becomes more robust than the KMO test. There are also other forms used less frequently in administration studies for evaluating the adequacy of factor analysis, which involve the measure of sampling adequacy (MSA), or the simple inspection of the correlation matrix in search of a substantial number of correlations above 0.3. The fundamental issue is that, even knowing the importance of fulfilling the requirements of the previous stage, there has been little empirical investigation using real data as to how each of the requirements affects EFA fit in terms of sphericity and the KMO. An exception to this

is the study by Guadagnoli and Velicer (1988) who, by way of the Monte Carlo simulation, showed that the ratio between sample size and the number of variables, unlike common sense, did not affect factor stability. On the other hand, the sample size itself had some effect. But the authors used simulated, not real, data. Given this absence of evaluations using empirical data our question is: **1)** Do the number of cases, the number of variables, the number of cases per variable, and the scalar amplitude interfere in the EFA adjustment?

Dimensionality. The mathematical and statistical techniques used in factor analysis seek to maximize the explanation of the factors identified and to identify the number of dimensions (Netemeyer *et al*., 2003). These techniques involve both methods for extracting these factors, and the rotation methods used for obtaining factors that have a greater degree of reliability (Aranha & Zambaldi, 2008), which we shall discuss in the following sections. For **factor extraction** a method needs to be chosen that takes into account the objectives of the analysis and the researcher's previous knowledge of variance. Some of these methods were developed for applicability to the whole population and not just with a sample, which means the data cannot be extrapolated beyond their limits. Generalization is only possible if the analyses are repeated using different samples. That is why the choice of extraction method must mainly take into account the interest in generalizing the results of the factor analysis (Field, 2009).

One of the most widely-used extraction techniques is **Principal Components Analysis** (PCA), which is particularly indicated when the objective is to summarize most of the information into a minimum number of factors, by concentrating the explanatory power on the first factor, or on the initial factors. Due to its algebraic extraction method some people do not even consider PCA to be a factor analysis method and there is much argument against its use, as can be seen in Field (2009). Another method used in factor extraction is **Principal Axis Factoring** (PA) or simply Factor Analysis, which is used more often when the objective is to identify latent dimensions that reflect what the variables have in common (Hair *et al*., 2009). The results of this technique are also limited to the cases analyzed and it is not indicated for generalization. Finally, the **Maximum-Likelihood Estimator (MLE)** method is statistical and not algebraic as is PCA, instead using iterative processes for extracting factors. It seeks to estimate factor loadings for the population "which makes the sample correlation matrix the most likely to be observed" (Aranha & Zambaldi, 2008, p. 86). It considers that cases were obtained by random selection and that the measured variables constitute the population of variables of interest to the study (Field, 2009).

Defining the number of factors starts with the dilemma of obtaining the maximum explanation of the variance with the minimum number of factors, which makes interpretation simpler (Johnson & Wichern, 2007). The conceptual basis is one of the criteria that must be considered for this definition,

and in this case the researcher defines the number of factors to be extracted before beginning to collect the data. Empirical evidence also serves as an argument for the choice of the number of factors, with the criteria for this definition – which are generally used in a concurrent way – involving (Hair *et al*., 2009): (a) eigenvalues, or latent roots larger than one, which is more reliable when the number of variables is between 20 and 50; (b) the percentage of accumulated variance, which aims for a certain degree of explanation guaranteed by the factors; (c) the Scree Plot test, which identifies "the optimum number of factors that can be extracted before the amount of unique variance begins to dominate the structure of the common variance" (Hair *et al*., 2009, p. 114); and (d) the presence of variables that lead to the heterogeneity of the cases, which generate factors with less capacity for explaining the total variance. Even knowing the assumptions of the different extraction methods, such as the different methods used for defining the number of factors, some questions persist regarding their use: **2)** Does the number of factors differ when we use different extraction methods? **3)** Which extraction methods have the greatest total explained variance? **4)** Do the construct variables converge on the same factor when different extraction methods are used?

Fitting the variables based on factor rotation. Factor rotation is carried out with the objective of obtaining better and easier intepretations of the results by concentrating the variable loadings on a particular factor. The results obtained have the same degree of total variance, but facilitate the analysis process by more directly associating the variable with a single factor and increasing its explanatory power (Bezerra, 2009). Rotation methods may be orthogonal when looking for factors that are not inter-correlated, and that is why there is a ninety degree, or oblique, angle between them, although a certain degree of correlation between factors is allowed, as represented by the different ninety degree angle. The recommendation is to carry out both types of rotation and check the correlation between the factors obtained. If the correlation is minimal when carrying out an oblique rotation, then an orthogonal rotation can be used (Field, 2009).

Of all the orthogonal rotation methods used, the most usual is **Varimax**, the principal characteristic of which is the attempt to concentrate the factor loading of a variable in a single factor, thus maximizing loading dispersion between factors and preventing this variable from giving high factor loadings for various factors (Bezerra, 2009). Varimax therefore offers a clearer separation of the factors, by concentrating on simplifying the columns of the rotated matrix (Hair *et al*., 2009). Another method of orthogonal rotation is the **Quartimax** method, which tries to maximize the loading dispersion of a variable throughout all the factors, which results in a factor with many variables with high loadings (Field, 2009). So Quartimax concentrates on simplifying the lines of the rotated matrix. This concentration, obtained around a single factor, may create difficulties when it comes to interpreting the structure of factors

resulting from factor analysis. One method that tries to combine the characteristics of these two previous methods is called **Equamax**, which seeks a joint simplification of the columns and lines of the rotated matrix, but as (Bezerra, 2009, p. 90) finds, "it is not commonly used", and according to (Field, 2009, p. 568), "its behavior is very erratic". The best known oblique rotaion method is the **Direct Oblimin**, which generates correlated factors with high but very complex eigenvalues, which makes analysis difficult. In some statistical packages (as is the case with SPSS), the allowed degree of correlation between the resulting factors is defined beforehand by the researcher, which may generate interpretation distortions. When the problem deals with a large number of data **Promax** rotation can be used, which is faster to calculate than Oblimin (Bezerra, 2009).

According to (Hair *et al*., 2009, p. 119), "no specific rule has been developed to guide the researcher in the selection of an orthogonal or oblique rotational technique in particular", but Field (2009) argues that in the case of research carried out with data obtained from human perceptions, such as research that employs a survey for investigating administration phenomena, orthogonal rotations should never be used, because in this particular line of research the factors obtained should not be correlated, and this hardly ever occurs. Even though there are differences between the assumptions of each rotation method, the question remains: **5)** Does the use of different rotation methods really define the variables in different dimensions?

Reliability and the correlation between factors. Analysis of the reliability of a factor obtained from EFA provides information about the relationship (internal correlations) between the individual items used in composing that factor, and by how much they result in a consistent measurement scale. The oldest and most widely used reliability measure is **Cronbach's alpha coefficient**, which represents the internal consistency of the factor based on the average correlation between the items that go to form the factor; it also considers the number of incorporated variables. This measure varies between 0 and 1, in which the higher the value the more reliable a dimension is. Despite its wide use, some authors question it, because such a reliability measure gives too much weight to the number of variables incorporated in the factor, which may lead to it being considered reliable (with an *alpha* greater than 0.7), without it necessarily being entirely consistent. For these reasons, some researchers have been using the **Intraclass Correlation Coefficient** (ICC) because they consider it to be more useful and accurate than Cronbach's *alpha* (Clark & Watson, 1995). This measure is more consistent because it takes into account only the correlations between the variables contained in a particular factor; it is not inflated by the number of variables of the construct (McGraw & Wong, 1996).

Knowing that we can partly evaluate the quality of a scale based on the reliability of a construct enables us to raise further questions: **6)** Does the choice of one extraction method over another affect the values of Cronbach's alpha and the ICC, generating more or less reliable factors? **7)** Does the use of different rotation methods also not have an influence on the reliability measures? Furthermore, the use of extraction and rotation methods in factor analysis also aims to obtain factors that are less inter-related. This is important because one of the fundamental assumptions is that the dimensions are empirically different. Therefore: **8)** Do different extraction and rotation methods affect factor discrimination? Finally, we also question the point up to which the adjustment indicators are related to the reliability and discrimination of the dimensions. This leads to our final research question: **9)** Do the number of cases, the number of variables, the number of cases per variable, and scalar amplitude affect the reliability and correlation of the factors? It is this question and to the others previously presented that we seek to answer in this study.

4. METHODOLOGICAL PROCEDURES

4.1. Data collection

Our study included a search for empirical articles that used the exploratory factor analysis method in the management area, published in Brazil between 2006 and 2010. To do so we intentionally selected six of the main journals from the area, in view of their evaluation in Qualis/CAPES (B1 and A2 classification) and because they publish a large number of quantitative studies. They are: Brazilian Administration Review (BAR); *Revista de Administração Contemporânea* (*RAC*) [Contemporary Administration Review]; *Revista de Administração Contemporânea Eletrônica* (*RAC*-E) [Electronic Contemporary Administration Review]; *Revista de Administração de Empresas* (*RAE*) [Business Administration Review]; *RAE-Eletrônica* [the *RAE's* e-version]; and *Revista de Administração* (*RAUSP*) [Administration Review]. Having selected the journals, we downloaded all the academic articles published in the period being analyzed, a total of 719 documents. We then read the abstract, the methodology, and the data analysis of the articles, which enabled us to identify 374 studies of a quantitative nature (52% of the total), and 179 studies that collected data by way of a survey (47.9% of the quantitative articles), of which 107 used the exploratory factor analysis method (59.8% of the survey-type studies). We recorded identifying data for each one of these articles, in particular the name, title, and e-mail address of the author, with the objective of sending a request by email to all the authors, asking for the database from which the article originated and the data collection instrument they used. Of the 107 articles identified, we failed to receive a reply to 42; in 17 cases the authors had lost the data, and in 21 there was some impediment, or they simply refused to send the information. As a result we received 27 databases. Four of them had to be removed because of problems with the file or due to an inconsistency in the analysis, which left us with a sample of 23 studies, which are

identified in the references with an asterisk (*). In the majority of these studies (20 cases), the authors explicitly mentioned the scalar validation process. In the other three cases we identified no problems with the indicators used in the constructs, nor in the factors that resulted from the factor analysis.

We believe that the non-randomness of the studies analyzed did not cause any type of selection bias. On the contrary, the proportional distribution of the studies that adhered to research by journal and the variety of themes involving the constructs analyzed point to the sample being very representative. Another concern we had, since this is a meta-analysis, was to see whether there was any bias caused by the publications selected. As Borenstein, Hedges, Higgins, and Rothstein (2009) and Sutton (2009) point out, publication bias occurs when the selection of studies does not represent the knowledge of the field. As we selected the articles from periodicals that are both generalist and, at the same time, the most relevant for quantitative research in the administration area, we believe that this bias was mitigated. We also believe that the publication bias caused by the tendency of the journals to select articles with significant results (Borenstein *et al.*, 2009) was not sufficient to distort the sample. This is a common problem in meta-analyses that assess the causality relationship between variables, which is not the case with this research.

4.2. Indicators

With the databases, the collection instruments, and their respective articles to hand, we built a tabulation matrix, identifying first of all the authorship and name of the study, the number of *cases*, the name given to the *construct* analyzed, the number of *variables* (items or indicators) and the amplitude of the *scales*. We also created an indicator by dividing the number of cases by the number of variables (*cases* per *variable*). We then reproduced the exploratory factor analysis process for each of the studies using the SPSS software, with three analysis protocols as the parameter. In the first we used three different extraction methods: Principal Components, Principal Axis Factoring, and Maximum Likelihood, which were chosen due to how extensively they are used. We took note of the Kaiser-Meyer--Olkin (KMO) test value of each study, such as the significance of Bartlett's sphericity test. For each of the extraction methods, we identified the *number of factors* generated, using both the Eigenvalue and Scree Plot methods. We subsequently recorded the *total explained variance percentage* for each of the extraction methods relative to the two factor definition methods.

In the second protocol, for each of the factors identified in the three extraction methods, we sought to assess the average reliability of each construct using Cronbach's *alpha* and Intraclass Correlation Coefficient (ICC, by the Two-Way Mixed model) method, and the *average correlation* between them by way of Pearson's "*r*" test. But we only analyzed the factors identified by the Scree Plot method, with the aim of

simplifying the analysis. Before this we defined some criteria for choosing the variables that would be part of each factor: we used the rotated matrix generated by the Varimax method; we ignored factor loadings less than 0.4; in cases in which the variables were present in two factors, we considered the one in which the absolute value in module of the factor loading was bigger; we inverted the variable if it was negatively related to the factor; and we created the factor based on the average of the indicators. After taking these measures, we had to create two more tables to record the results. In the first we identified the values for Cronbach's *alpha* and the ICC of each of the factors in each study, by way of three extraction methods (Principal Components, Principal Axis Factoring, and Maximum Likelihood). In the second, and ignoring the sign, we assessed Pearson's correlation between all the factors identified for each of the three extraction methods. To add these results to our analysis matrix, we calculated the average values for Cronbach's *alpha*, the ICC, and Pearson's "*r*" for each study.

Finally, in the third protocol we assessed Cronbach's *alpha*, ICC, and Pearson's "*r*", as in the previous protocol, but comparing three of the most widely used rotation methods: Varimax, Quartimax, and Direct Oblimin (Oblique). The limitation criteria were the same as for the previous protocol, although instead of limiting our analysis to a single rotation method, we limited it to a single extraction method, *principal components*. In the case of oblique rotation (Direct Oblimin), on the other hand, we defined the Delta value as being zero. We also created two additional tables, one for noting Cronbach's *alpha* and the ICC for each of the factors, and another for recording Pearson's correlation. We then obtained the average values for Cronbach's *alpha*, ICC, and Pearson's "*r*" for each study, but now for the different rotation methods that we added.

In addition to the three protocols, we assessed to what extent the variables fitted the same factors, by contrasting different definition methods (Eigenvalue and Scree Plot), extraction methods (Principal Components, Principal Axis Factoring, and Maximum Likelihood), and rotation methods (Varimax, Quartimax, and Direct Oblimin). For this we created nine indicators (3 extraction methods x 2 definition methods + 3 rotation methods) to assess the degree of convergence in percentage terms, dividing the number of variables that fitted the same factor by the total number of variables of the analyzed construct. So the bigger the convergence between the two different extraction or rotation methods, the greater the percentage of variables that fitted the same factor.

4.3. Method

This study is a meta-analysis: a type of investigation that combines quantitative results that originate from multiple studies, which by way of statistical methods produces a general summary of these results (Littel, Corcoran, & Pillai, 2008). In our study in particular these results refer to the parameters

generated by the application of exploratory factor analysis. As our sample was composed of 23 studies, in order to carry out the meta-analysis we had to use Spearman's non-parametric correlation instead of linear regression models for comparing the variables in a linear fashion. These results are shown in Tables 2 and 4. In any event, we compared the results of Spearman's correlation with Pearson's correlation, the results of which converged in all the analyses. In those cases in which it was necessary to compare the average values of the explained variance between the different extraction and rotation methods (Table 1); the average convergence of the variables between the factors based on the different extraction and rotation methods used (Table 2); and comparing the average values of Cronbach's alpha, ICC, and Pearson's "*r*" (Table 3, second column), we had to use a model that allowed us to evaluate the same cases under different conditions. The most appropriate for this purpose is the General Linear Model (GLM) of Repeated Measures, which reduces non-systematic variability and allows for a smaller number of cases to be worked with. However, to satisfy the adjustment parameters of the model, we first tested the sphericity hypothesis using Mauchly's test, in which the null hypothesis ($p > 0.05$) indicates the similarity between the experimental conditions. In those cases in which the null hypothesis is not true, we assessed whether the difference between the averages was significant, by way of Greenhouse-Geisser's correction of the F- statistic value. As there is evidence of conservatism in this correction under circumstances in which sphericity is greater than 0.9 and Greenhouse-Geisser's test is greater than 0.75 (Field, 2009),

we ascertain the significance under these conditions by way of Huynh-Feldt's correction. We also additionally compare the average results using Pillai's multivariate trace test. Finally, to compare each of the conditions individually, we compare the averages using the Post Hoc test with Sidak's correction, which lessens the effects when there is sphericity, without the loss associated with corrections such as Bonferroni's.

5. RESULTS

To present the results we followed the same sequence as the theoretical framework, so we will first discuss fulfilling the requirements for using exploratory factor analysis, such as the adjustment statistics, and thereby answer the first two questions. In our sample, we found that most of the variables were metric, thus fulfilling one of the basic conditions. However, there were not always five or more variables per factor, thus going against one of the requirements of Hair *et al.* (2009). This assumption can be questioned, because the suggestion of using at least five indicators is related to the development of scales (e.g. Churchill, 1979) and not necessarily their treatment. As far as concerns the number of cases, only one of them had a value less than 50 (see minimum value in Table 1), with all the others having more than 100 observations. However, when we divided the number of cases per variable we find that 10 of the 23 studies (43%) did not have 10 observations per variable. Most of the variables did not have a normal distribution in the Kolmogorov-Smirnov and Shapiro-Wilk tests, but had a symmetry less than or close to 1.5, and kurtosis less than 3.

Table 1

Adjustment and Explained Variance Statistics

	Minimum	Maximum	Average
Cases	22	1025	306
Variables	3	69	28
Cases per variable	2	27	13
Scalar amplitude	4	11	6
Kaiser-Meyer-Olkin (KMO)	0.566	0.971	0.804
Total Explained Variance (%)			
Principal Components, Eigenvalue*	47.3	73.8	62.4
Principal Axis Factoring, Eigenvalue	35.1	63.0	50.7
Maximum Likelihood, Eigenvalue	35.6	63.1	50.8
Principal Components, Scree Plot*	30.6	67.8	48.2
Principal Axis Factoring. Scree Plot	26.6	56.6	40.5
Maximum Likelihood, Scree Plot	26.5	56.4	40.5

* Significant difference when compared to the other methods ($p < 0.001$ in Sidak's *post hoc* test).

Generally speaking the studies fulfilled the requirements, but we still need to check the point up to which they interfere in the factor analysis adjustment. But to do this we must describe them beforehand. All Bartlett's sphericity tests were significant ($p < 0.05$), showing that the matrices were duly correlated, with this indicator being extremely stable. However, the Kaiser-Meyer-Olkin (KMO) test had values less than 0.7 in four cases, which hypothetically would compromise a stable solution. To find an answer to the **first question,** we assessed the possible sources of variation of the KMO. Our data indicated that it is very sensitive to sample size (Spearman's rho = *0.733, p<0.001*), and to a lesser degree related to the number of variables (*rho = 0.418, p = 0.024*) and to the number of cases per variable (*rho = 0.410, p = 0.026*). But there was no correlation with scalar amplitude (*rho = 0.030, p = 0.447*).

In seeking to answer the **second question**, we found that the number of factors identified in EFA is totally independent of the extraction method used (Principal Components, Principal Axis Factoring, and Maximum Likelihood), since they were all identical, both for the Eigenvalue definition method and for the Scree Plot method. The only noticeable and currently known difference is that the number of factors identified by the Eigenvalue method (average of 7 and maximum 18) is greater than with the Scree Plot method (average of 3 and maximum 5).

Answering the **third question,** the results indicate that there are differences in terms of explanation potential. As can be seen in Table 1, the average variance explained by the *principal components* extraction method is greater than for the others (Greenhouse-Geisser, Sidak, and Pillai trace tests, with *p < 0.001*). When the factors are defined by the Eigenvalue method the average of the *principal components* was 62.4%, which is significantly greater than the others, which were between 50.7% and 50.8%. This also happened with the Scree Plot method, with an average of 48.2% for *principal components* and 40.5% for the others. We can state, therefore, that this method absorbs greater total explained variance than the other methods.

If there is any difference in the degree of explained variance because of the different extraction methods used, it is obvious that some of the variables that go to make up each one of the constructs are positioned in different factors when we alter the method. In Table 2, and specifically in the second column, we show the degree of average convergence of the variables depending on the extraction method used.

When the Eigenvalue factor definition method is used, we observed that the convergence varied from 73% to 76%, such differences not being significant when we crosschecked different extraction methods ($p = 0.577$, Greenhouse-Geisser test). When the same extraction methods were compared, but now defining the factors using the Scree Plot method, we found that the convergence between the *principal components* and *principal axis factoring* methods, which was 79%, is greater than with the other methods, the values of which were around 65% ($p < 0.05$ in Sidak's Post Hoc test). So in answer to the **fourth question,**

we can state that there are differences in the fit of the variables in the factors when we use different extraction methods, which can be seen by the degree of average convergence. In fact the degree of convergence tends to be constant, with the exception of the combination between the *principal components* and *principal axis factoring* methods, based on the Scree Plot definition, where there was greater convergence.

The answer to the **fifth question** is also found in Table 2. The data indicated that the convergence of the variables between the different rotation methods, on average, was greater than with the extraction methods. While convergence varied from 65% to 79% for the extraction methods, with the rotation methods this variation was between 81% and 84%.

It has to be emphasized that even with the change from the extraction to the rotation method most of the variables still fit in the same factor. But which variables tend to converge on the same factors and which on different factors? Individually observing the variables of each study analyzed, those that had the lowest factor loading (between 0.4 and 0.59) tended to be more volatile and fluctuated more frequently between the different factors depending on the method used. In cases in which the factor loading was high (over 0.7), there was practically no fluctuation, which is in line with what was found by Stevens (2009). This leads us to conclude that the variables of greatest statistical importance are the most stable, which would not compromise the accuracy of the factors so much if a method was inappropriately chosen.

In addition, in trying to identify some probable causes for the convergence, we sought to correlate it with some of the indicators, which can be seen in Table 2. We found no evidence that the number of cases, scalar amplitude, and the KMO value affected convergence. In those cases where we used the Scree Plot method, which reduces the number of factors to a greater degree, the more variables that were involved in the factor analysis, the smaller the convergence, this relationship being significantly strong when we contrasted the Varimax and Quartimax methods (*rho = -0.798, p < 0.01*). When using the Scree Plot method, convergence also proved to be sensitive to the number of cases per variable, but its effect is positive: the more cases per variable, the greater the convergence. We also find a relationship that is somewhat curious: the number of factors identified by the Scree Plot method bears no relation to the convergence between factors, although the number of factors identified by the Eigenvalue method has a negative influence on the convergence between factors generated using the Scree Plot method. This leads us to conclude that the greater the reduction generated by the Scree Plot method relative to the Eigenvalue method, the less convergent are the different extraction and rotation methods.

To answer Questions 6, 7, and 8 we assessed the average reliability by way of Cronbach's *alpha* and the intraclass correlation (ICC), and we also calculated the correlation between the factors resulting from factor analysis for each of

Table 2

Convergence of the Variables Between the Extraction and Rotation Methods

Comparison of the Methods	Average	Correlation (Spearman's rho)						
		Cases	Variables	Cases per Variable	Scale	KMO	Eigenvalue Factors	Scree Factors
Extraction by Eigenvalue								
Principal Components vs. Principal Axis Factoring	76%	0.126	-0.180	0.329	0.145	0.357	-0.304	-0.266
Principal Components vs. Maximum Likelihood	73%	0.044	-0.271	0.282	0.150	0.235	-0.390	-0.388
Principal Axis Factoring vs. Maximum Likelihood	74%	0.035	-0.386	0.292	0.081	0.085	-0.381	-0.068
Extraction by Scree Plot								
Principal Components vs. Principal Axis Factoring	79%	0.334	-0.445[*]	0.707[**]	-0.201	0.302	-0.676[**]	-0.092
Principal Components vs. Maximum Likelihood	65%	0.120	-0.485[*]	0.589[**]	-0.051	0.332	-0.720[**]	-0.240
Principal Axis Factoring vs. Maximum Likelihood	65%	0.196	-0.428[*]	0.593[**]	-0.063	0.392	-0.665[**]	-0.150
Rotation by Scree Plot								
Varimax vs. Quartimax	83%	0.044	-0.798[**]	0.610[**]	-0.170	-0.105	-0.798[**]	-0.198
Varimax vs. Direct Oblimin	81%	0.098	-0.569[**]	0.469[*]	-0.279	-0.050	-0.504[*]	-0.252
Quartimax vs. Direct Oblimin	84%	0.386	-0.337	0.639[**]	-0.248	0.170	-0.429[*]	-0.067

[*]. Correlation is significant at the 0.05 level (1-tailed).

[**]. Correlation is significant at the 0.01 level (1-tailed).

the extraction and rotation methods. The values obtained are shown in Table 3. Among the different extraction methods (Items 1, 2, and 3 of the Table), *maximum likelihood* had the biggest Cronbach's alpha (0.803), but Sidak's test indicated no significant differences between it and the other extraction methods ($p = 0.192$ and 0.093). By the ICC evaluation, the most reliable method was *principal components* ($F = 17.143$, $p < 0.001$, *Sidak's* $p < 0.001$), while there were no significant differences for *maximum likelihood* and *principal axis factoring* (Sidak's $p = 0.281$). Answering the **sixth question**, we found no strong evidence that the extraction method exercises any great influence on the average reliability of the constructs, since there is no significant difference from Cronbach's alpha. Despite the ICC having indicated that the factors extracted by way of the *principal components* method were less reliable, between the other two there was no difference.

Answering the **seventh question**, there was also no difference in reliability between the rotation methods (Items 3, 4, and 5 in Table 3) relative to the values of Cronbach's *alpha* (Sidak's $p > 0.05$), even though the absolute values of the Quartimax rotation suggested they are less reliable. For the ICC, Varimax rotation

proved more reliable than Direct Oblimin (Sidak's $p = 0.008$), but not more reliable than Quartimax (Sidak's $p = 0.825$), which in turn is not significantly more reliable than Direct Oblimin (Sidak's $p = 0.279$). Given these results, we find that the choice of rotation method has little influence on reliability.

With regard to the **eighth question**, we found no significant discrimination differences in the factors between the different extraction methods used ($F = 0.535$, $p = 0.590$), as can be seen in Items 1, 2, and 3 of Table 3. On the other hand, of the rotation methods, Quartimax had the lowest correlation between the factors ($F = 8.188$, $p = 0.001$, Sidak's $p < 0.05$), while there was no difference between Varimax and Direct Oblimin (Sidak's $p = 0.489$). It was expected that Varimax would also show a low correlation between factors because it is an orthogonal rotation method, like Quartimax, but the results did not indicate this.

Finally, with the objective of understanding which adjustment indicators are related to more or less reliability and factor discrimination, we compared them with the number of cases, the number of variables, the number of cases per variable, scalar amplitude, and the value of the KMO (Table 4). This enabled us to answer the **ninth question**.

Table 3

Average Reliability (Alpha and ICC) and Correlation Between Factors

Extraction, Rotation	Cronbach (*Alpha*)	ICC	Pearson's *r*
1. Principal Axis Factoring, Varimax	0.795	0.412	0.440
2. Maximum Likelihood, Varimax	0.803	0.435	0.445
3. Principal Components, Varimax	0.780	0.367	0.430
4. Principal Components, Quartimax	0.758	0.363	0.374
5. Principal Components, Direct Oblimin	0.774	0.355	0.429

Table 4

Correlation Between Adjustment Measures With Cronbach's Alpha, ICC, and Pearson's r

Extraction and Rotaion Methods	Cases	Variables	Cases per Variable	Scale	KMO
Cronbach's *Alpha*					
1. Principal Axis Factoring, Varimax	0.322	0.433*	0.072	0.098	0.750**
2. Maximum Likelihood, Varimax	0.270	0.525**	-0.050	0.198	0.683**
3. Principal Components, Varimax	0.247	0.489*	-0.029	0.186	0.697**
4. Principal Components, Quartimax	0.430*	0.409*	0.177	0.222	0.747**
5. Principal Components, Direct Oblimin	0.282	0.505**	-0.002	0.166	0.704**
Intraclass Correlation (ICC)					
1. Principal Axis Factoring, Varimax	0.093	-0.629**	0.574**	0.036	0.223
2. Maximum Likelihood, Varimax	0.164	-0.341	0.365*	0.213	0.312
3. Principal Components, Varimax	0.238	-0.268	0.484*	0.069	0.563**
4. Principal Components, Quartimax	0.300	-0.319	0.524**	0.157	0.561**
5. Principal Components, Direct Oblimin	0.184	-0.287	0.457*	0.146	0.577**
Pearson's Correlation Between Factors					
1. Principal Axis Factoring, Varimax	0.188	0.222	0.135	-0.044	0.663**
2. Maximum Likelihood, Varimax	0.244	0.195	0.194	-0.028	0.744**
3. Principal Components, Varimax	0.247	0.224	0.198	-0.077	0.748**
4. Principal Components, Quartimax	0.151	0.075	0.214	-0.109	0.706**
5. Principal Components, Direct Oblimin	0.166	0.219	0.036	-0.149	0.743**

*. Correlation is significant at the 0.05 level (1-tailed); Spearman's rho.

**. Correlation is significant at the 0.01 level (1-tailed).

As we can see in Table 4, the value of Cronbach's *alpha* is strongly correlated with the adjustment values of the KMO test (*rho = 0.683* to *0.750*), more so than with the number of variables (*rho = 0.409* to *0.425*), which is one of the indicators that go to make up the *alpha* measure (Netemeyer *et al.*, 2003). Scalar amplitude, the number of cases, and the cases per variable did not prove to be related to Cronbach's *alpha*, with the exception of the Quartimax rotation, which increases significantly with the number of cases (*rho = 0.430*).

Reliability as assessed by the ICC is positively influenced by the proportion of cases per variable (*rho = 0.365* to *0.574*) and shows no relationship with either the number of cases or with scalar amplitude. It is interesting to note also that the ICC generated by the *principal components* extraction method is positively related to the KMO (*rho = 0.561* to *0.577*), which does not occur with the other extraction methods. We also observed that the ICC is negatively sensitive to the number of variables only for the *principal axis factoring* extraction method, which is

plausible, since the increase in the number of variables per factor tends to reduce the ICC value, especially when using extraction methods that do not minimize the correlation.

With regard to factor discrimination, we found that there is a strong correlation between the value of the KMO test and the average correlation between factors, which was expected, because this indicator takes into consideration the relationship between correlation and partial correlation. As the KMO reflects the general correlation based on the indicators, it is obvious that this would also reflect in a greater correlation of the factors. The other indicators did not prove to be significant.

6. DISCUSSION AND CONCLUSIONS

Starting with the objective of this study, which was to improve the use of exploratory factor analysis in survey-type research based on a comparison of different extraction, factor definition, and rotation methods, we reach some conclusions with regard to how they affect the fit of the dimensions. To do

so, we shall look again at the answers obtained for each one of the questions, which are shown in Table 5.

First, we found that the significance of Bartlett's test did not vary in relation to the parameters of the samples, but that the KMO test was highly sensitive to the number of cases, and to a lesser extent to the number of cases per variable. Scalar amplitude, on the other hand, showed no relationship whatsoever with KMO or Bartlett. Secondly, the data indicated that there is no significant relationship between the number of factors identified using the rotation method, since they were all identical. This result is important because, as Johnson & Wichern (2007) pointed out, definition of the number of dimensions is the most important decision with regard to the use of factor analysis. Third, we found that of the extraction methods, *principal components* analysis was the one that gave the greatest percentage of explained variance, regardless of the factor definition method chosen. In line with what was indicated in the literature (Hair *et al.*, 2009; Johnson & Wichern, 2007), if the objective of any researcher is to maximize explained

Table 5

Summary of the Results

Research Question	Results
1) Do the number of cases, the number of variables, the number of cases per variable, and the scalar amplitude interfere in the EFA adjustment?	a) The significance of Bartlett's sphericity test did not vary. b) The Kaiser-Meyer--Olkin *(KMO)* test was very sensitive to sample size and moderately related to the number of variables and the proportion of cases per variable. There is no influence from scalar amplitude.
2) Does the number of factors vary when we use different extraction methods?	No. Regardless of the extraction method, the number of factors is the same, using both the Eigenvalue and Scree Plot methods.
3) Which extraction methods have the greatest total explained variance?	The extraction method by *principal components* has a greater total explained variance than the other methods.
4) Do the construct variables converge on the same factor when different extraction methods are used?	No. While most of the variables converged on the same factor, some converged on different factors. However, those that tend to converge less are the ones with a smaller factor loading (between 0.4 and 0.59).
5) Does the use of different rotation methods really define the variables in different dimensions?	Yes. Despite the vast majority of the variables fitting the same factor, a small part converged on other factors. This small part is exactly that formed by variables with a low factor loading.
6) Does the choice of one extraction method over another affect the values of Cronbach's alpha and the ICC, generating more or less reliable factors?	We found no strong evidence that any great influence is exercised by the extraction method on reliability, despite the fact that ICC indicated that the *principal components* method was less reliable.
7) Does the use of different rotation methods have an influence on the reliability measures?	No. The rotation method has little influence on reliability, despite the fact that the Varimax method showed itself to be a little more reliable by way of the ICC.
8) Do different extraction and rotation methods affect factor discrimination?	The use of different extraction methods does not affect discrimination, but of the rotation methods, Quartimax was the one that discriminated most.
9) Does the number of cases, the number of variables, the number of cases per variable, and scalar amplitude affect the reliability and correlation of the factors?	The value of Cronbach's *alpha* is associated with the value of the Kaiser-Meyer--Olkin *(KMO)*, even more than the number of variables. The value of the *intraclass correlation* (ICC) tends to rise with the increase in the proportion of cases per variable; specifically, when the *principal components* extraction method is used there is influence from the KMO on the ICC. Discrimination falls with the increase in the KMO, since there is an increase in the correlation between factors.

variance based on a reduced number of factors, hypothetically the most appropriate method would be *principal components*. Fourth, we identified that definition of the variables in the factors extracted requires care, because the results indicate that the different extraction and rotation methods position part of them in different factors. If in doubt, researchers can eliminate those variables with factor loadings less than 0.6, which would almost eliminate the differences of the position of the variables between the factors. Another measure we identified for reducing this variability is to increase the number of cases per variable, which relates to the need to have a larger sample. But this would only be significant when the Scree Plot method is used. Fifth, as occurred with the use of different extraction methods, the choice of rotation method affects the allocation of the variables, as once again the ones that vary most are those with lower loading factors. Sixth, we assessed the influence of the different extraction methods on reliability (Cronbach's *alpha* and the ICC). We found no significant differences in reliability when using Cronbach's *alpha*. However, with regard to the reliability assessed by way of the ICC, the *principal components* extraction method proved to be less reliable. Seventh, the results indicate that rotation method has little influence on reliability, despite the Varimax method proving to be significantly more reliable with the ICC. Eighth, there is no evidence that different extraction methods affect factor discrimination. Of the rotation methods, on the other hand, the Quartimax method significantly reduced the correlation between factors. Ninth and finally, the data indicated that an increase in the KMO value increases the reliability of Cronbach's *alpha* as well as the ICC, except that the latter is only sensitive when the *principal components* extraction method is used. On the one hand, the increase in the KMO is positive, while on the other it is negative since the results indicated that its increase is associated with a drop in discrimination. We must also add that the number of cases per variable is associated with the greater reliability evaluated by the ICC, as well as the number of variables being associated with a larger Cronbach's *alpha*. This was nothing more than

we expected, since the value of the *alpha* also considers the number of variables.

In short, our results indicate that if researchers want the maximum explained variance, they must choose the *principal components* extraction method. If they want greater reliability using the ICC, they must choose the *maximum likelihood* extraction method. Of the rotation methods, Varimax supplies greater reliability with the ICC, and Quartimax the least correlation between factors. Furthermore, as made clear in the analyses and listed among the factor analysis assumptions, it is important to have a sample that allows a good number of cases per variable, since the greater this proportion, the better the adjustment, which solves many of the problems and incongruences in factor analysis.

In practical terms, we believe that our study contributes to the field of research in administration in particular, and to the social sciences in general, because it empirically finds how the choice of factor analysis method has an impact on the quality of the scales in explanation, reliability, and discrimination percentage terms. If doubts existed as to which adjustment method to choose, we believe that some of these have been resolved, thus improving the use of this research tool. As Johnson & Wichern (2007) emphasized, in the current stage the use of factor analysis can still be seen as an art, in which there is no "best way" of proceeding. That is why, when in doubt, the suggestion is always to compare the results of different extraction and rotation methods.

Despite its contributions, this study has its limitations, which suggest future studies. The first is that we crosschecked different extraction methods with different rotation methods to assess the reliability differences and the correlation between factors. The second is the lack of convergent and discriminant validity tests between factors. Future studies could incorporate these elements and work with confirmatory factor analysis using structural equation models. Furthermore, the sample could be expanded, allowing the use of other meta-analysis statistical techniques.◆

REFERENCES

Note: References marked with an asterisk indicate studies included in the meta-analysis.

*Aguiar, A. C. (2007). Percepções de pesquisadores sobre atividades cooperativas de P&D: Uma análise com base na teoria institucional. *Revista de Administração Contemporânea*, *11*(4), 59-86.

Aranha, F., & Zambaldi, F. (2008). *Análise fatorial em administração* (1ª ed.). São Paulo: Cengage Learning.

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, *16*(1), 78-117.

Bezerra, F. A. (2009). Análise fatorial. In *Análise multivariada para os cursos de administração, ciências contábeis e economia* (pp. 73-130). São Paulo: Atlas.

*Boehe, D. M. (2007). Desenvolvimento de produtos em subsidiárias de empresas multinacionais no Brasil. *Revista de Administração de Empresas*, *47*(1), 33-45.

*Boehe, D. M., Milan, G. S., & Toni, D. (2009). Desempenho do processo de desenvolvimento de novos produtos: O peso relativo de fatores organizacionais, mercadológicos e operacionais. *Revista de Administração* [RAUSP], *44*(3), 250-264.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

Cantelmo, N. F., & Ferreira, D. F. (2007). Desempenho de testes de normalidade multivariados avaliado por

REFERENCES

simulação Monte Carlo. *Ciência e Agrotecnologia*, *31*(6), 1630-1636.

Churchill, R. (1979). A paradigm for developing better measures for marketing constructs. *Journal of Marketing Research*, *16*(1), 53-64.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309-319.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16-29.

Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, *11*(2), 213-228.

Field, A. (2009). *Descobrindo a estatística usando o SPSS* (2ª ed.). Porto Alegre: Artmed.

*França, L. H. F. P. (2009). Influências sociais nas atitudes dos "top" executivos em face da aposentadoria: Um estudo transcultural. *Revista de Administração Contemporânea*, *13*(1), 17-35.

*Gimenez, F. A. P., & Inácio Junior, E. (2006). Validação do instrumento Team Factors Inventory em empresas brasileiras. *Revista de Administração Contemporânea*, *10*(edição especial), 137-158.

*Godoy, A. S., Antonello, C. S., Souza-Bido, D., & Silva, D. (2009). O desenvolvimento das competências de alunos formados do curso de administração: Um estudo de modelagem de equações estruturais. *Revista de Administração* [RAUSP], *44*(3), 265-278.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103*(2), 265-275.

Hair, J., Black, W. C., Babin, B. B., Anderson, R. E., & Tatham, R. L. (2009). *Análise multivariada de dados* (6a ed.). Porto Alegre: Bookman.

Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6th ed.). Upper Saddle River: Pearson.

Kuhn, T. S. (2009). *A estrutura das revoluções científicas* (9a ed.). São Paulo: Perspectiva.

*Larentis, F., & Slongo, L. A. (2008). Relacionamento em canais de marketing como fonte de vantagem competitiva sustentável: Um estudo com fabricantes de móveis e lojas exclusivas. *Revista de Administração* [RAUSP], *43*(3), 209-223.

Laros, J. A. (2011). O uso da análise fatorial: algumas diretrizes para pesquisadores. In *Análise fatorial para pesquisadores* (pp. 141-160). Brasília: LabPAM Saber e Tecnologia.

Littel, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis*. New York: Oxford University Press.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, *3*(3), 635-694.

*Lombardi, M. F. S., & Brito, E. P. Z. (2010). Incerteza subjetiva no processo de decisão estratégica: Uma proposta de mensuração. *Revista de Administração Contemporânea*, *14*(6), 990-1010.

Lopes, H. E. G. (2005). Abrindo a caixa preta: Considerações sobre a utilização da análise fatorial confirmatória nas pesquisas em administração. *Revista Economia & Gestão*, *5*(11), 97-116.

*Lopes, E. L., Hernandez, J. M. C., & Nohara, J. J. (2009). Escalas concorrentes para a mensuração da qualidade percebida: Uma comparação entre a Servqual e a RSQ. *Revista de Administração de Empresas*, *49*(4), 401-416.

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

*Maciel, C. O., Hocayen-da-Silva, A. J., & Castro, M. (2008). O ideário de escola na ótica dos docentes: Pura subjetividade ou padrões estruturados de cognição nos cursos de administração? *Revista de Administração Contemporânea*, *12*(3), 659-688.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30-46.

*Medeiros, O. R., & Quinteiro, L. G. L. (2008). Financial disclosure and international capital mobility in Latin America. *Brazilian Administration Review*, *5*(2), 160-176.

*Moraes, E. A., Graeml, A. R., Sanchez, O. P., & Mesquita, F. S. B. (2006). Seleção de cursos de educação continuada por executivos de empresas. *RAE-Eletrônica*, *5*(2), 1-22.

Netemeyer, R. G., Bearden, W. O. E., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks: Sage.

*Oliveira Jr., M. M., Borini, F. M., & Guevara, A. J. de H. (2009). The strategic relevance of the Brazilian subsidiaries of multinational corporations. *Revista de Administração* [RAUSP], *44*(4), 285-298.

*Paiva, J. C. N., Barbosa, F. V., & Ribeiro, A. H. P. (2009). Proposta de escala para mensurar o valor percebido no varejo bancário brasileiro. *Revista de Administração Contemporânea*, *13*(2), 310-327.

*Pérez-Nebra, A. R., & Torres, C. V. (2010). Medindo a imagem do destino turístico: Uma pesquisa baseada na teoria de resposta ao item. *Revista de Administração Contemporânea*, *14*(1), 80-89.

*Petroll, M. N., Damacena, C., & Zanluchi, E. (2007). Apelos sexuais em propagandas: Um experimento no contexto universitário brasileiro. *RAC-Eletrônica*, *1*(1), 67-83.

REFERENCES

*Pires, P. J., & Costa Filho, B. A. (2008). Fatores do índice de prontidão à tecnologia (TRI) como elementos diferenciadores entre usuários e não usuários de internet banking e como antecedentes do modelo de aceitação de tecnologia (TAM). *Revista de Administração Contemporânea*, *12*(2), 429-456.

*Regos, A., Cunha, M. P., & Souto, S. (2007). Espiritualidade nas organizações e comprometimento organizacional. *RAE-Eletrônica*, *6*(2), 1-27.

Rencher, A. C., & Christensen, W. F. (2012). *Methods of multivariate analysis* (3rd ed.). Hoboken: Wiley.

Smith, A. M. (1999). Some problems when adopting Churchill's paradigm for the development of service quality measurement scales. *Journal of Business Research*, *46*(1), 109-120.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences*. New York: Routledge.

Sutton, A. J. (2009). Publication bias. In *The handbook of research synthesis and meta-analysis* (pp. 435-452). New York: Russell Sage Foundation.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson.

*Thomaz, J. C., & Brito, E. P. Z. (2010). Reputação corporativa: Construtos formativos e implicações para a gestão. *Revista de Administração Contemporânea*, *14*(2), 229-250.

*Trad, S., & Maximiano, A. C. A. (2009). Seis sigma: Fatores críticos de sucesso para sua implantação. *Revista de Administração Contemporânea*, *13*(4), 647-662.

*Vieira, V. A. (2010). Mensuração da qualidade de serviço no varejo eletrônico e seu impacto sobre as intenções comportamentais. *Revista de Administração de Empresas*, *50*(2), 199-214.

*Vieira, V. A., & Appio, J. (2010). O impacto da conexidade no comportamento do consumidor em relação aos programas de televisão. *Revista de Administração Contemporânea*, *14*(4), 703-721.

*Zacharias, M. L. B., Figueiredo, K. F., & Almeida, V. M. C. (2008). Determinantes da satisfação dos clientes com serviços bancários. *RAE-Eletrônica*, *7*(2), 1-23.

ABSTRACT

Normal science and its tools: Reviewing the effects of factor analysis in management

The aim of this study is to investigate how different methods of extraction, factor definition, and rotation of exploratory factor analysis affect the fit of measurement scales. For this purpose, we undertook a meta-analysis of 23 studies. Our results indicate that the Principal Components method provides greater explained variance, while the Maximum Likelihood method increases reliability. Of the rotations methods, Varimax provides greater reliability while Quartimax provides lower correlation between factors. In conclusion, this study highlights implications for quantitative research and suggests potential new studies.

Keywords: exploratory factor analysis, reliability, quantitative methods, survey.

RESUMEN

Ciencia normal y sus herramientas: Revisando los efectos de los métodos de análisis factorial exploratorio en Administración

El objetivo de este estudio fue investigar cómo los diferentes métodos de extracción, la definición de los factores y la rotación del análisis factorial exploratorio afectan el ajuste de las escalas de medición. Para ello, se realizó un meta-análisis de 23 estudios. Nuestros resultados indican que el método de componentes principales proporciona una mayor varianza explicada, mientras que el método de máxima verosimilitud aumenta la fiabilidad. Entre los métodos de rotaciones Varimax proporciona una mayor fiabilidad y Quartimax proporciona menor correlación entre los factores. En conclusión, este estudio pone de relieve las implicaciones para la investigación cuantitativa y sugiere nuevos estudios.

Palabras clave: análisis factorial exploratorio, fiabilidad, métodos cuantitativos, encuesta.