Revista Internacional de Sociología

Revista Internacional de Sociología RIS
vol. 77 (2), e130, abril-junio, 2019, ISSN-L:0034-9712
https://doi.org/10.3989/ris.2019.77.2.17.088

SCALE RECODING IN SOCIOLOGICAL RESEARCH: A NEW VALIDATION METHODOLOGY. AN APPLICATION TO A POLITICAL SURVEY

RECODIFICACIÓN DE ESCALAS EN INVESTIGACIONES SOCIOLÓGICAS: UNA NUEVA METODOLOGÍA DE VALIDACIÓN. APLICACIÓN A UNA ENCUESTA POLÍTICA

Elena Abascal
Universidad Pública de Navarra
eabascal@unavarra.es
ORCID iD: https://orcid.org/0000-0001-5661-5483

Vidal Díaz de Rada
Universidad Pública de Navarra
vidal@unavarra.es
ORCID iD: https://orcid.org/0000-0002-9638-3741

Ignacio García Lautre
Universidad Pública de Navarra
nacho@unavarra.es
ORCID iD: https://orcid.org/0000-0002-4151-8477

M. Isabel Landaluce
Universidad de Burgos
iland@ubu.es
ORCID iD: https://orcid.org/0000-0002-3963-7485

Cómo citar este artículo / Citation: Abascal, E., V. Díaz de Rada, I. García Lautre y M. I. Landaluce. 2019. “Scale recoding in sociological research: a new validation methodology. An application to a political survey”. *Revista Internacional de Sociología* 77(2):e130. https://doi.org/10.3989/ris.2019.77.2.17.088

Received: 08/06/2017. **Accepted:** 10/10/2018
Published online: 04/06/2019

Abstract

The recoding of scale variables is a common step in the analysis of survey data. It is not immune, however, to certain pitfalls, such as the introduction of biases, or potential data distortion. This paper presents a methodological proposal for the validation of any recoding process, whether it involves metric- or categorical-scale variables. The aim of the proposed methodology is to verify the adequacy of the re-codification by indicating how close in structure the re-coded data are to the original data. The basis of the methodology is a factorial analysis technique, Multiple Factor Analysis (MFA), which is performed on a global data table juxtaposing the original-scale and recoded-scale data. The procedure is tested on real-world data drawn from a public opinion poll on perceptions of leading politicians in the Spanish Parliament.

Keywords

Simple and Multiple Factor Analysis; Social Survey Measurement Scales.

Resumen

La recodificación de variables es una etapa habitual en el análisis de datos de encuesta. No está exenta de riesgos, como pueden ser la introducción de sesgos o las posibles deformaciones de los datos originales. En este trabajo se propone una metodología de validación de cualquier recodificación, tanto de variables de escala métrica, como de escala categórica para reducir el número de categorías. El objetivo de la metodología propuesta es comprobar la adecuación de la recodificación, indicando en qué medida los datos recodificados mantienen la misma estructura de los datos originales. La metodología se basa en el uso de un método factorial, Multiple Factor Analysis (MFA), aplicado sobre una tabla global obtenida como yuxtaposición de los datos medidos con la escala original y los mismos datos medidos con la escala recodificada. El procedimiento ha sido testado a partir de un conjunto de datos reales extraídos de una encuesta de opinión pública sobre los principales políticos del Parlamento Español.

Palabras Clave

Análisis Factoriales Simples y Múltiples; Escalas de Medida en Investigaciones Sociales.

Introduction and research aims

The questionnaire is one of the most widely-used sociological data gathering tools. After establishing the number of questions to be included, it is very important to select an appropriate question format. For surveys involving complex issues, one of the most frequent options is to use the so-called battery question format, where the same set of response options is assigned to the whole set of questions (Revilla et al., 2017: 30). A more precise definition is given in Couper et al. (2013: 322), where it is stated that "in grid questions, a series of items is presented (usually in rows), sharing a common set of response options (usually in columns), asking one or more questions about each item". For Roßmann, Gummer and Silber (2017: 2) grid (or matrix) questions, "present respondents a set of items that they are asked to answer using a common response scale, usually a Likert-type or rating scale".

The battery question requires less response time than other formats and, according to Bell et al. (2001), Couper et al. (2001) and Thorndike et al. (2009), generates lower partial non-response rates. There is a lack of consensus on this issue, however, since other studies (Iglesias et al., 2001; Smyth and Olson, 2016; Toepoel, Das and van Soest, 2009; Tourangeau et al., 2004) observe the opposite effect, as well as less differentiation in responses and correlation between item responses. Mention should also be made of the more compact layout of battery questions, where the scale is shown only once; in web surveys, this reduces the need for scrolling or changing webpage (Revilla et al., 2017).

After deciding on the type of question format, the next crucial step is to select the response scale. This decision influences the quality of the responses and determines the statistical method to be used for the survey data analysis (Conrad and Kreuter, 2015). Many researchers decide to use a long scale with a wide array of scores so as to keep to the initial survey objectives. Thus, once collected, the data can be altered (recoded) to find the most appropriate scale for the response analysis. It can also be useful to recode metric variables, such as consumption levels or age, as categorical variables. A few recommendations in this context would be to respect natural boundaries, examine the quantiles, and keep clear of any outlying categories. Finally, when working with categorical variables, an excessive number of categories can distort the analysis of survey data, so grouping to reduce the number could be a good option.

The *Encyclopaedia of Survey Research Methods* uses the term Recoding to "describe the process of making changes to the values of a variable" (Tien, 2008). Data transformation (Bourque and Clark, 1994) or the modification of variables (Weisberg et al. 1996) prior to considering the available statistical methods is a topic that few papers have addressed. Furthermore, the vast majority of these papers explain the recoding process, but hardly any discuss its effect on the data. Even compendiums such as that of Miller (1991) ignore this issue.

As a rule of thumb, the scale recoding should not generate any loss of key information that could seriously compromise the analysis of the original data. Consequently, the researcher must verify or validate the recoded scale to ensure that it does not lead to a biased interpretation of the original data. In consequence, we propose a methodology for finding an objective criterion to help the researcher in this choice, that is, a methodology for scale recoding validation.

The factorial method Multiple Factor Analysis (MFA) is the core of the methodology for scale recoding validation. MFA provides numerical and graphic indicators to show how closely the recoded data resemble the original data. Furthermore, it is a factorial method designed to handle mixed data, that is, to analyse metric and categorical tables simultaneously. This is the case studied later, where the original variables, which are ratings of 0 or 1 (lowest) to 10 (highest) usually considered as metric variables, are recoded as categorical ones. In general, the MFA scale validation methodology is very flexible since it enables the comparison of different types of original scales (categorical or metric) with different types of recoded scales (usually categorical), according to the research objectives.

Abascal and Díaz de Rada (2014) present another case where the MFA scale validation methodology could be very useful. The authors compare the pros and cons of analyzing the same data by metric or categorical scale using factorial methods. They state a preference for the categorical scale, despite the undeniable drawbacks that arise when the original variables have a large number of categories and are analysed by Multiple Correspondence Analysis (MCA), the correct factorial method for this type of data. In this case, some of the categories could be so similar, so very close, that they fail to contribute differentiated information. A large number of MCA factors would have to be selected to explain the relevant information content of the original data and the solution would become unwieldy, with too many factorial planes to interpret. One way to overcome this problem is to group the categories by recoding the original scale in order to reduce the dimensions of the original data without loss of relevant information. This makes it easier to identify the most characteristic features of the data and thus benefit from the acknowledged advantages of the categorical scale over the metric scale.

It follows from the above that it is common practice to recode the original scale in order to obtain a categorical scale with a small number of categories. It goes without saying that, when the recoding process involves merging consecutive categories, it needs to be carried out very carefully, only after a prior analysis and never as a matter of course. The problem is even greater when categorizing a metric variable. In this case, Abascal et al. (2006b) recommend using both scales simultaneously (the original metric scale and the recoded categorical scale), applying a bi-criteria analysis.

This paper is structured as follows. The next section presents data obtained from a survey of public perceptions regarding leading members of the Spanish Parliament. After this, there comes a section explaining each stage of the scale recoding validation MFA methodology, followed by another showing the main results of the empirical check of the factor stability between the original and recoded scales in order to illustrate the appropriateness of the MFA methodology. The main conclusions of the MFA methodology, its limitations and proposals for ways to address them, are presented in the final section.

Data source

The data used in this paper is drawn from a Spanish household survey conducted by the Sociological Research Centre, Centro de Investigaciones Sociológicas (CIS, 2014). The survey in question (Number 3033[1]), which was the latest available at the commencement of this study, was carried out in July 2014. The survey universe was all Spanish citizens (except residents of Ceuta and Melilla) aged 18 years and over from which a sample of 2,500 individuals was selected, in accordance with CIS sample quality standards. In order to present the MFA recoding scale validation methodology more clearly, we have selected a subsample comprising all those respondents who answered the selected questions. The size of the subsample is more than sufficient to apply the methodology, although it precludes the possibility of drawing inferences for the population as a whole.

The survey question selected to illustrate the MFA recoding scale validation methodology reads as follows:

> Please indicate whether you know each of the following political leaders [SHOW CARD]. How do you rate the political performance of NAME OF POLITICAL LEADER KNOWN on a scale of 0 to 10, where 0 means “very poor” and 10 means “very good”?

With the aim of illustrating the MFA recoding validation methodology, we use the ratings for the leaders of the four main parties as given by those respondents who recognized them. The same criterion is used in Abascal and Díaz de Rada (2014): “*A survey covering national, regional and local politician’s produces a considerably lower response rate on some politicians… it was decided to reduce the scope of the survey to politicians from parties of nationwide relevance*…”

The final sample comprises 1,704 respondents who gave their opinions of the four leaders at the time of the survey. Mariano Rajoy, the Spanish president and president of the Popular Party, PP, right wing. Alfredo Pérez Rubalcaba, president of the main opposition party, the Spanish socialist party, PSOE, a moderate left-winger. Cayo Lara, president of the second most important opposition party, United Left, IU/ICU, a left- winger. Finally, Rosa Díez, ex-socialist, co-founder of the UPyD –i.e., the Progressive and Democratic Union, Member of the European Parliament, and a regular media figure. Figure 1 shows the frequency distributions of the ratings for each politician.

The four bar charts in Figure 1 show similar irregular distributions, the most noteworthy features of which are discussed briefly below. It is also immediately apparent that the value with the highest frequency is zero. Furthermore, there is a greater difference of frequency between some consecutive values like 0 and 1 or 1 and 2. Finally, there is a high frequency of DK/NA responses in two of the bar charts. All these features indicate that the variables should be handled as categorical variables.

MFA scale recoding validation methodology

On the whole, the scale recoding validation methodology must enable the comparison of several tables of data to determine whether they have a similar structure. There are several factorial methods for this purpose, as can be seen in Dazy and Le Barzic (1996), Abascal et al. (2001) and Greenacre and Blasius (2014). The proposed methodology, based on MFA, works by comparing the stability of the factors obtained from the original and the recoded data tables and by checking for differences in the internal structure of the two data tables.

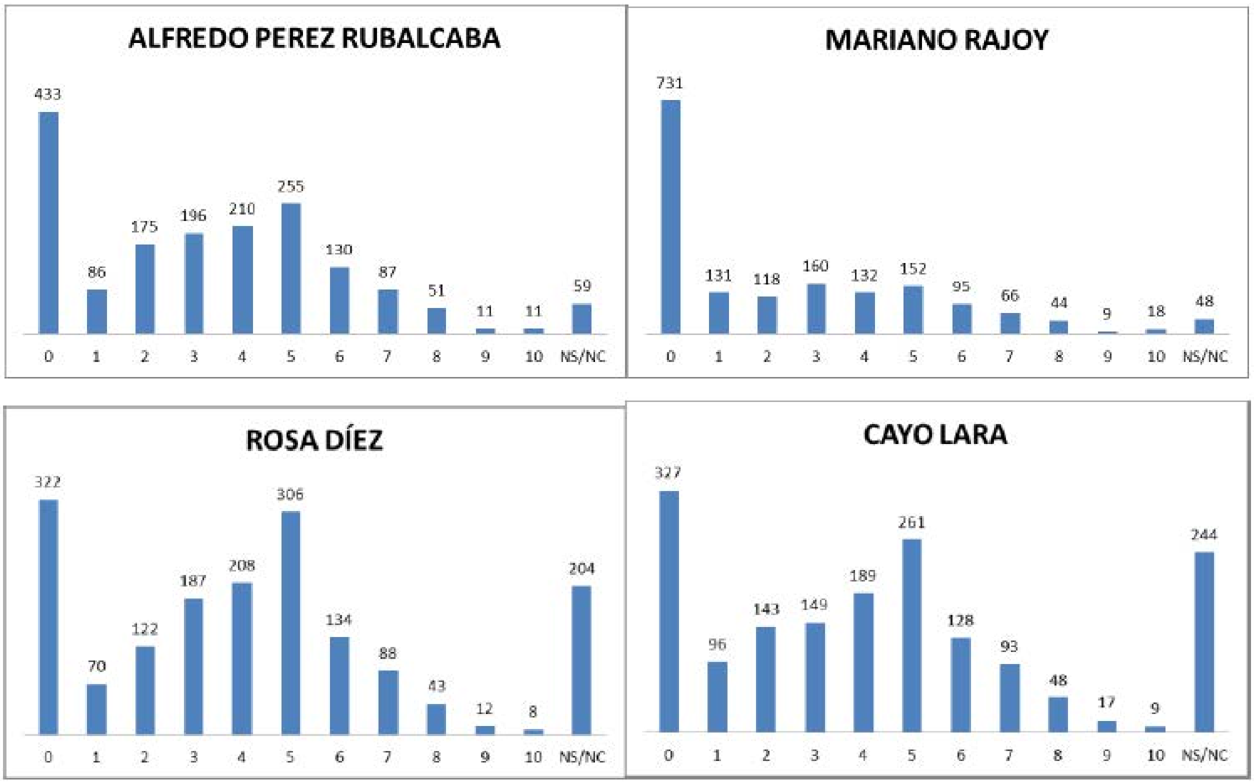
Rationale for selecting MFA scale recoding validation methodology

The MFA, proposed by Escofier and Pagés (1982, 1990, 2016), at the French School of Data Analysis, is a method of factorial analysis adapted for dealing with data tables in which a set of cases is described by means of several groups of variables**.** The aim is not merely to obtain a typology of the cases described by a whole set of variables, but also to discover any possible relationships between patterns emerging within each of the groups.

There are three main reasons for using MFA as the scale validation method:

- MFA can be used simultaneously with groups of metric and categorical variables, the only proviso being that all the variables within each group must be of the same type. The validation of a scale recoding requires an MFA of two groups

Figure 1
Bar charts of the frequency distributions of the politician ratings



of variables drawn from the same data source: one group having been measured on the original scale (all variables within the group will be coded either as metric or as categorical) and another group having been measured on the recoded scale.

- MFA balances out the influence of the two groups (original and recoded variables). The main factors of the MFA reveal the variability between cases, such that the influence of each group is comparable. The result of applying MFA is to balance the influence of the original data against that of the recoded data, such that neither group biases the analysis.
- MFA provides a rich source of numerical and graphical indicators. As well as the usual indicators used in factorial analysis methods, MFA provides measures of relationships between groups and different factor planes, thereby facilitating interpretation of the data, as will be illustrated in the section that follows.

Brief description of MFA

Essentially, MFA is a weighted Principal Component Analysis (PCA) (Figure 2), for which there follows a brief description contextualized to the objectives of this paper, extracted from Abascal et al. (2006):

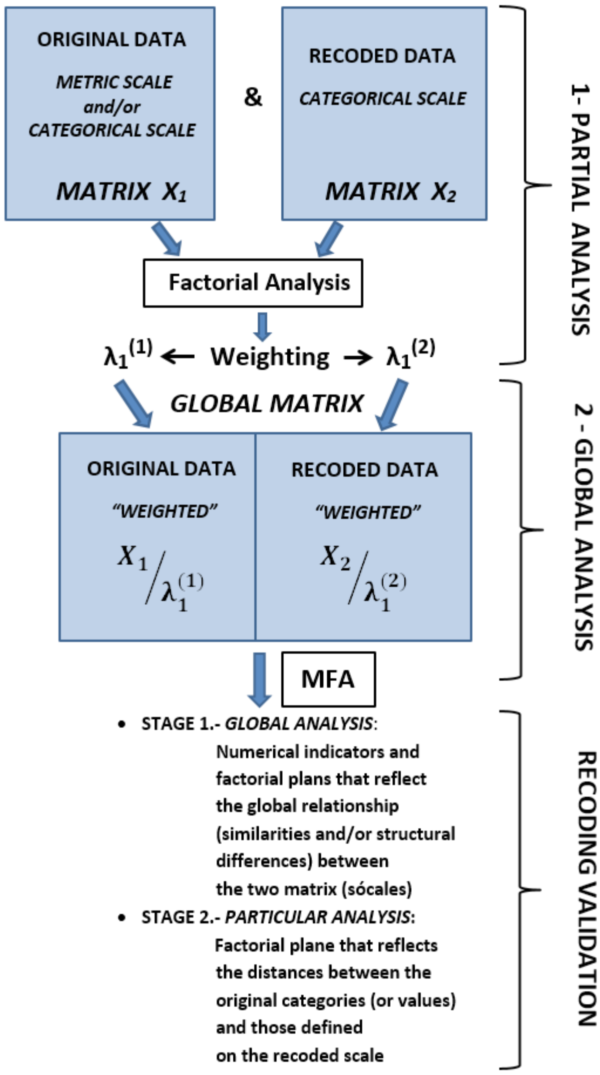
1. *Partial analysis*. Each group of variables is linked to a cloud of cases, known as *partial cloud*, and represented by a rectangular table of data, X_1 and X_2 (the cases are the rows and the variables are the columns). In this stage, a factorial analysis is performed on each of the tables and the first eigenvalue, $\lambda_1^{(1)}$ and $\lambda_1^{(2)}$, in each analysis is retained. The factors obtained in each separate analysis are called *partial factors*.
2. *Global analysis*. A weighted PCA is performed on the global table formed by juxtaposing the *2* tables **[X_1 |X_2]**. Each table is weighted by the inverse of the first eigenvalue retained in the partial analysis, $1/\lambda_1^{(1)}$ **and** $1/\lambda_1^{(2)}$. The factors obtained in the global analysis are called *global factors*. Two main properties are associated with the use of this weight:
   - The structure of each table, X_1 and X_2, is maintained by giving the same weighting to all the variables which form one group.
   - The influence of the groups is balanced because the maximum inertia of any of the *partial clouds* defined by the different groups is 1 in any direction. In addition, the influence of every group in the first global factor is balanced. This can be appreciated in the following expression which gives the distance between any two cases (i, i') in the global cloud,

$$d^2(i,i') = \frac{1}{\lambda_1^{(1)}} d^2\left(i_{(1)}, i'_{(1)}\right) + \frac{1}{\lambda_1^{(2)}} d^2\left(i_{(2)}, i'_{(2)}\right)$$

Where $i_{(1)}i_{(1)}$ and $i'_{(1)}i'_{(1)}$ are two individuals characterized only by the variables of group *1* (original scale) while $i_{(2)}i_{(2)}$ and $i'_{(2)}i'_{(2)}$ are the same individuals characterised only by the variables of group *2* (recoded scale).

Figure 2

Diagram of the scale recoding validation methodology



One noteworthy characteristic of MFA is that it can be used to analyze tables of different types of variables (quantitative and categorical), known as mixed tables. The results from quantitative tables are equivalent to those obtained via PCA, while the results from categorical tables are equivalent to those obtained by means of Multiple Correspondence Analysis (MCA). This greatly facilitates their comparison.

As well as providing the classical results of other factor analysis techniques (PCA or MCA), MFA has the distinctive feature of treating the group of variables as one more element of analysis. In the case at hand, there are two groups: that of the variables measured on the original scale and that of the variables measured on the recoded scale. This yields a large number of specific results. Thus, in MFA, the terms "partial" and "global" take on special importance. In terms of cases (observations), we speak of "partial cases" when the variables of each group are characterized separately. In the context that concerns us, each case gives two partial cases per respondent. One of these partial cases represents the respondent's responses on the original scale (the score of 0 to 10 assigned to the politician, plus the DK/NA option). The other partial case represents the same respondent's replies on the recoded scale (an evaluation based on one of four categories). We speak of "global cases" when the characterization is based on all the information from all the groups of variables. In our context, there is one global case per respondent. This global case captures all of a respondent's answers, taking into account both scales (original and recoded). Global cases can be interpreted as the midpoint of their respective partial cases. In our context, the global cases are the synthesis obtained by combining the two measuring scales (original and recoded).

For a deeper explanation of MFA, the interested reader is referred to Lebart, Morineau and Piron (2000) and Escofier and Pagès (2016) and, for an applied perspective, to García Lautre, I. and Abascal, E. (2003 and 2004), Landaluce, Mª I. (1995), García Lautre, I. (2001) and Dazy and Le Barzic (1996).

Stages of the MFA scale recoding validation methodology

The methodology for the validation of a recoding process consists of an MFA of the global data table obtained from the juxtaposition of the original with the recoded tables (Figure 2).

The results yielded by the MFA are analyzed in the pre-established order, which can be observed in the validation methodology. The focus in each stage is on the issues that are relevant for the objective of the validation process:

Stage 1: Global analysis. The first stage is the interpretation of the graphical and numerical results of the MFA, which are of a global nature, since they apply to the tables (or scales) globally. The main results:

- *The matrix of correlations between the principal factors of the two partial analyses.* If there is stability between the two scales, the matrix should display high correlations between factors of the same order and low correlations between factors of a different order.

- *The matrix of correlations between partial factors and global factors.* To detect the presence (or absence) of common factors in the two coding scenarios, one needs to calculate the coefficient of correlation between the global factor and the corresponding factor in each of the groups under analysis. When the correlation is high (close to 1) the global factor is apparent in all the tables, in other words, it is a common factor. If this occurs on the first axes, it will show that the groups have the same structure and that the recoding is therefore correct. Otherwise, (if there is low correlation between the global and partial first-order axes) it will be a significant sign that the recoding has distorted the initial structure of the data.

- *The factors of the separate analyses, known as partial factors*. In this graph, the focus is on the proximity between the partial factors and the global factors of the MFA. Thus, those partial factors that are highly inter-correlated will be also highly correlated with the global factors. If they are not close, it has to be assumed that the recoding has distorted the data structure.
- *Plane of the groups of variables, in which each group is represented by a point.* In the case at hand, if the two groups are close together, this means that they are depicting the same reality and therefore the recoding is correct.

At this point, the researcher has sufficient objective criteria to decide whether to proceed with recoding. If the results indicate weak factor stability between the two scales, an alternative form of recoding must be found and subjected to the validation procedure. Otherwise, the researcher can proceed to stage 2.

Stage 2: Particular analysis. This stage consists of the analysis of precisely those categories that show the least stability after recoding. It will inform the researcher as to whether the recoding needs to be adjusted, and in what way, or whether it can be considered completely valid. In this procedure, the original and the recoded categories are projected onto the principal factor-planes (known as MFA graphs of the variables) in order to obtain a single mapping of all categories (both original and recoded), and thus fitting the two scales into a single framework. Analysis of the distances between the original and the recoded categories will reveal which are least stable and indicate what adjustments are required. In other words, any recoded categories that are too far away from their respective original categories must be examined closely, so that the recoding can be modified accordingly. Briefly, this stage consists of a visual comparison of the original scale with the recoded scale, the outcome of which depends crucially on the researcher's experience in the interpretation of such graphs.

Once the above two stages are complete, the researcher is equipped to make the final decision regarding the validity of the proposed recoding, and, if a replacement is deemed necessary, the information obtained in the previous stages will indicate the direction the adjustment must take.

Scale recoding validation for the Spanish politician ratings survey

Prior/Preliminary analysis: the decision to recode

In the survey chosen for the purposes of this study, the CIS uses a scale of 0 to 10 for the politician ratings plus one more category for the "don't know /no answer" response option. This is usually treated as a metric scale, that is, by analyzing the average rating for each politician and adding the percentage of non-response, (don't know/no answer). Such data structures are frequently analyzed by PCA, the drawbacks of which are noted and duly explained in Abascal and Díaz de Rada, 2014. In this paper, therefore, the scale is directly considered categorical and subjected to MCA (Abascal et al., 2017).

Having justified this first decision, we need to note that the resulting number of categories is 12 (11 from the scale of 0-10, plus the non-response category, "don't know/no answer") for each rating of each of the four politicians. The total dimension of the data table, therefore, is 1,704 rows (cases) by 48 columns (categories). Under these conditions, the excessive number of categories in the MCA can give rise to negative consequences which can be summed up as follows (Lebart et al., 2000):

- The first factors may be capturing the influence of "rare categories", that is, those selected in only a few cases, and may therefore be eclipsing more common response behaviour or relegating it to a higher order.
- The factors, overall, may be formed by very few categories, which would suggest that they are hardly worth taking into consideration since they do not indicate a general trend and can be difficult to interpret.
- There could be too many factors to examine simultaneously.

The MCA of the data considered in this analysis reveals that the first factors, in fact, present the characteristics mentioned above. Thus, the principal axis, or first factor, is determined mainly by the "non-response" category", in other words, it highlights those respondents who select "don't know/no answer" (for almost all their politician ratings). Close examination of the plane formed by factors 2 and 3 (Figure 3), clearly illustrates the need to recode the original scale for the following reasons:

- There are too many response categories: contiguous scores are very close, and almost overlapping (for example, 2 and 3 in quadrant 4; or 5, 6 and 7 in quadrant1).
- Likewise, the vertical axis (factor 3) shows that ratings lower than 5 (low on the negative side)

are the opposite of the ratings higher than 5 (high on the positive side).

- The zero rating (category 11) virtually alone determines the horizontal axis (factor 2).
- The response categories ordered from 0 to 10 depict a sort of parabola or horseshoe shape, indicating the so-called Guttman effect (Lebart et al., 2000), which is very common in this type of analysis. This effect indicates that the response categories selected by the respondents to rate the politicians are not linearly related (one of the points which confirm the inappropriateness of treating the scale as metric). Respondents selecting one of the extreme response categories (0 or 10) for one politician tend to select the same one for others, but also frequently select the extreme opposite (10 or 0, respectively).

The shortcomings of these results reveal the need to recode the original scale. By reducing the number of categories, it will be possible not only to mitigate their negative effects when MCA is used, but also to retain the aforementioned advantages of using a categorical scale.

In view of the above results, the scale will be recoded as follows:

- Keep the category “don’t know/no answer” (label: -N)
- Keep the no rating category, that is, “zero ratings” (label: -Z), since it behaves very differently from a score of 1, and it appears with very high frequency (Figure 1).
- Define a new category to group ratings of less than 5 (excluding zero) and label it “fail mark” (label:-F)
- Define a new category to group ratings of 5 or more, to be labelled “pass mark”[2] (label:-P)

This considerably reduces the size of the data table, from 12 categories per initial question to 4. It does, however, result in some loss of information and also carries the risk of introducing biases or distortions of the original data. It is for this reason that a recoding validation stage is considered necessary before proceeding with the data analysis. The kernel of this paper, therefore, is a proposal of methodology for achieving this objective. Its underlying philosophy and adequacy are described in the following sections.

Validation of a “correct” recoding

This section illustrates how the proposed methodology permits examination of the stability of the selected recoding described in the preceding subsection. The procedure consists of a MFA of the duplicated data table, as initially coded into 12 categories (group 1) and as recoded into 4 categories (group 2). Both groups are described on a categorical scale. The results obtained in both stages confirm the appropriateness of the recoding, which considerably reduces the size of the problem while maintaining the initial structure of the data. Some of the results are presented below.

Stage 1: Global analysis

The first factors of each group show structures common to both. This is clear from the high correlation between factors of the same order, for the different groups (Table1) and from the virtually null correlation between factors of different order.

Figure 3

Plane of factors 2 and 3 from the MCA, showing the 12 response categories for each of the 4 politicians

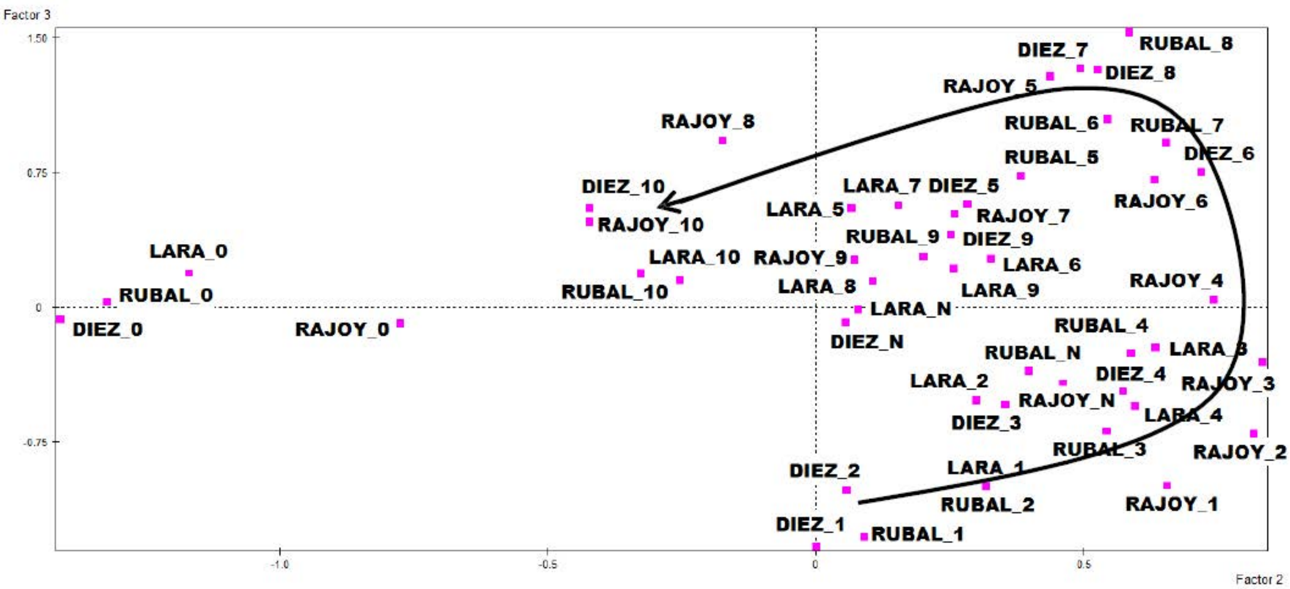


Table 1
Correlations among factors of the partial analyses

	G.2 (recoded)				
	F1	F2	F3	F4	F5
G.1(original)					
F1	**-1.00**	0.01	0.00	0.00	0.00
F2	-0.01	**-0.98**	0.01	0.01	0.00
F3	0.00	0.00	**-0.91**	-0.02	-0.05
F4	0.02	-0.06	-0.06	0.49	0.01
F5	0.00	0.14	0.18	0.35	-0,10

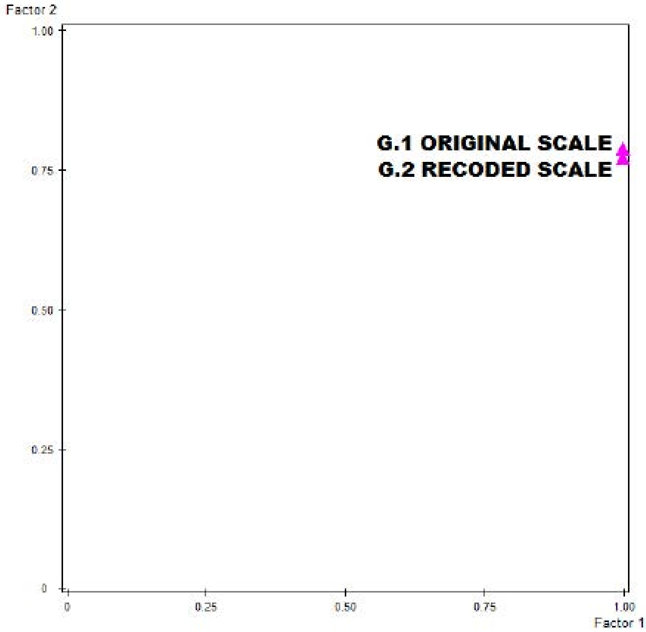
The correlations between the global factors and the factors for the separate groups analyzed (Table2) confirm the stability between the structure of the initial data and that of the recoded data. There is a strong similarity between the results of the partial and global analysis, as can be seen from the correlations of 1 or close to 1.

Table 2
Correlations among the factors of the global analysis and the factors of the partial analyses

Factor	Correlations				
	F1	F2	F3	F4	F5
G.1 O G.1 Original scale	1.00	1.00	1.00	0.99	0.99
G.2 R G.2 Recoded scale	1.00	1.00	1.00	0.99	0.99

The numerical data interpreted so far are illustrated and complemented with graphs. Thus, the very similar overall pattern (internal structure) of the two tables analyzed is confirmed in Table 3 and Figure 4.The contributions of the two groups in the formation of the factors are well balanced (i.e., similar). Each scale (the original one and the recoded one) has its respective point on the plane and the points almost overlap. The coordinates of the groups measure the cumulative inertia contribution of each group to the corresponding axes of the MFA, while also indicating which groups contribute most significantly to the global factors. The weights ensure a maximum axial inertia of 1 in each direction. In this case, the high, and almost coinciding, coordinates enable us to identify two common factors on the scales under comparison. All this allows us to conclude that the information contained in the original data remains unaltered after this recoding.

Figure 4
Principal plane from the MCA, showing the groups



Stage 2: Particular analysis

In this stage, all the ratings of the politicians using the four recoded and twelve original response categories are projected onto the principal factor planes and the distances between them are analyzed.

For the purposes of this example, we take the factor plane formed by the second and third factors (plane 2, 3) because the first factor still captures the “don’t know/no answer” category, as commented on in the preceding subsection. The next step is the interpretation of the distances between the ratings for a single politician, Rajoy (President of Spain at the time of the survey), which is enough to show how the procedure works. Note that the categories corresponding to fail marks are denoted on the original scale by the values 1, 2, 3 and 4 and the recoded fail mark category is denoted by Rajoy _F. Figure 5 zooms in on the first quadrant in plane (2,3) to enable closer inspection of the differences. It is possible to observe the proximity between categories (quantified, generally, by means of the global indicators described and interpreted earlier), and what stands out is the central position occupied by the recoded category (Rajoy _F).

Table 3
Coordinates and contributions of the groups in the global analysis

	Coordinates					Contributions				
	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5
G.1 Original scale	1.00	0.79	0.60	0.50	0.38	**50.0**	50.5	50.9	52.0	52.0
G.2 Recoded scale	1.00	0.77	0.58	0.46	0.35	**50.0**	49.5	49.1	48.0	48.0

The procedure could then continue with the ratings of the other three politicians, but, when both scales show high factor stability, as in this case, it is not strictly necessary. In other words, the measures of global similarity obtained in stage 1 and the particular analysis of Rajoy's ratings, performed in this second stage, enable the validation of the recoding of the original categorical scale. Therefore, despite substantially reducing the dimensions of the studied phenomenon, the recoding causes no relevant loss of original information.

The usefulness of the second stage becomes apparent when the global indicators are found to reveal low factor stability between the two scales, indicating that the recoding has distorted the structure of the initial data. A situation of this nature raises the need to analyze the distances between the variables or categories on the original scale and those on the recoded scale. The purpose of such an analysis is to identify the precise differences between the two scales, show which categories are the least stable, and thus indicate the changes required in the recoding.

Validation of an "incorrect" recoding

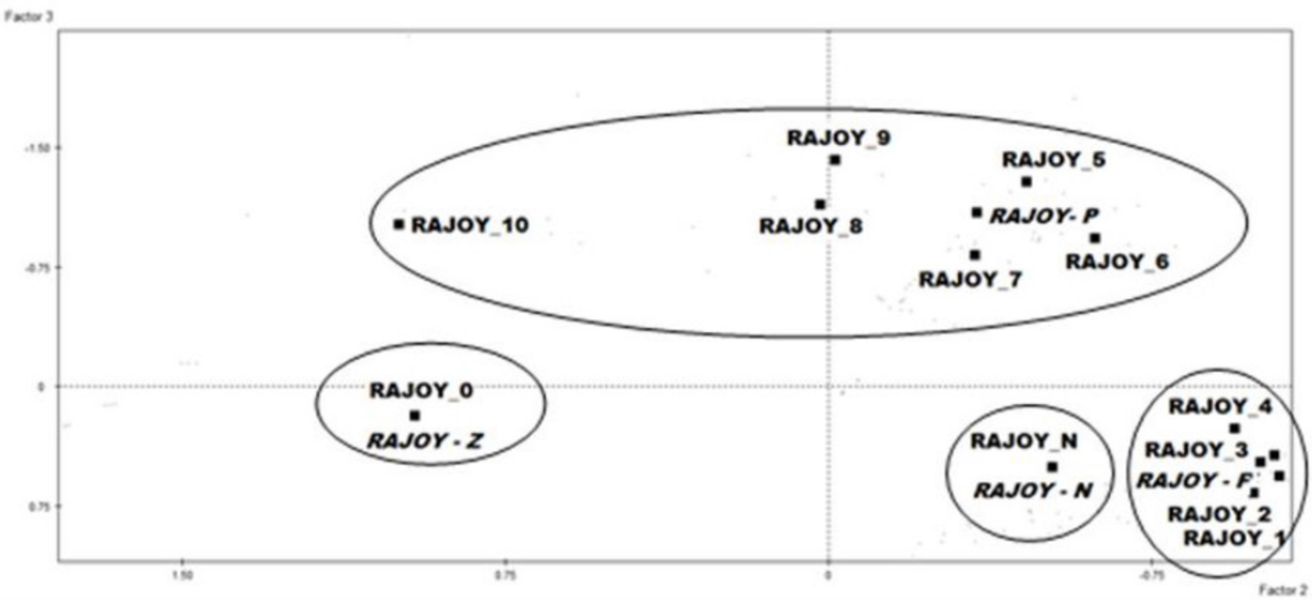
The proposed methodology is able to detect when a recoding is inappropriate and liable to lead to major information loss or biases. In this subsection, the aim is to illustrate the capability of the methodology by analyzing an inappropriate recoding. In the new scale, the null (zero) rating is added to the fail category, leaving only 3 categories: "don't know/no answer", "pass" and "fail". It is known, *a priori*, that this new recoding is inappropriate, as the influence of the zero rating in the second MCA factor will now be diluted within the fail category.

The aim here is to show briefly how the proposed methodology is able to detect this serious recoding error. Thus, the original scale variables form group 1 and the recoded scale variables form group 2. We have included as many indicators from the first stage as are required to illustrate the distortion of the original information due to inappropriate recoding.

The correlations between the partial factors from the analysis of the original scale and those from the new one are very weak, except for the first factor. This is to be expected, given that the recoding maintains the "don't know/no answer" category, which, as seen in the preceding subsection, defines the first factor. The correlation between the second partial factors, however, is very weak (0.36), and actually drops to 0.0 between the third partial factors and subsequent ones. In consequence, the correlation between the partial factors of the same order is weak; so the structure of the original data and the recoded data has changed. In addition, there is strikingly high correlation (0.85) between partial factor 3 in the original data and partial factor 2 in the recoded data. This indicates that the new scale has "deleted" a piece of relevant information, captured by the second principal dimension of the MCA of the original scale. By merging the zero rating with its adjacent categories, the recoding conceals a key piece of information and the proposed methodology detects this easily.

Figure 5

Plane of factors 2 and 3 showing the original categories (RAJOY_number) and the new recoded categories (RAJOY - category)



https://doi.org/10.3989/ris.2019.77.2.17.088

Table 4
Correlations among the factors of the partial analyses

	G.2 (E. recoded)				
G.1(E. original)	F1	F2	F3	F4	F5
F1	**-1.00**	-0.08	-0.01	0.02	0.02
F2	-0.01	**0.36**	**-0.85**	-0.02	0.14
F3	0.00	0.00	-0.09	0.49	0.31
F4	0.01	-0.07	0.02	-0.03	0.09
F5	0.01	-0.03	0.00	0.04	0.04

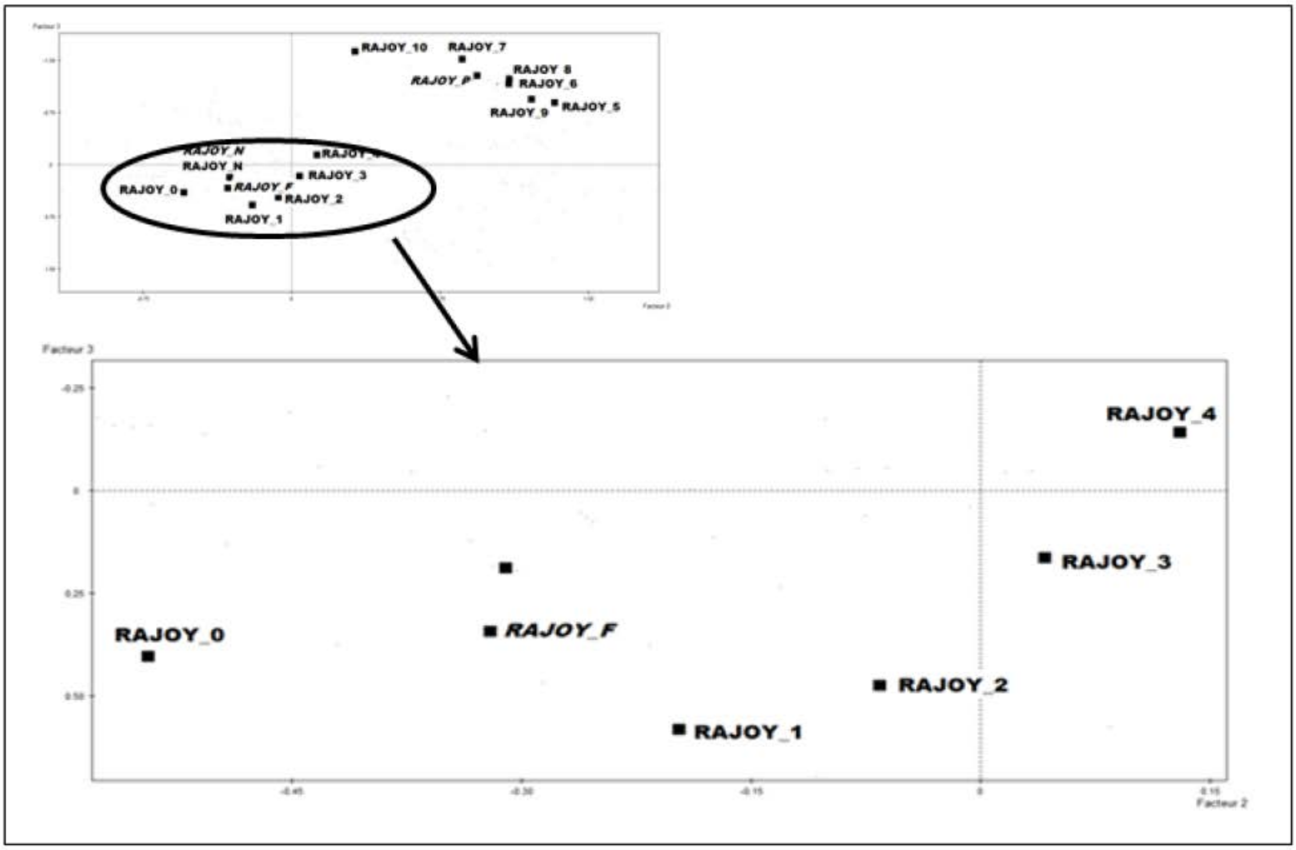
A quick comparison of Figures 5 and 6 reveals that the new recoding modifies the original data structure by merging two different categories (RAJOY_Z and RAJOY_F in the first recoding) into one (*RA JOY_F* in the new recoding). The opposed positions of RAJOY_Z and RAJOY_F with respect to the factor 2 (Figure 5) show their different typology. Both categories are an important part of the information content of the original data and any alteration would constitute a serious distortion. A similar effect is observed for the remaining politicians, but the case of Rajoy is the most illustrative.

Conclusions and future lines of research

It is common practice for researchers to recode the original scale chosen to measure a phenomenon of interest. There are various reasons for recoding, two of the main ones being to reduce the dimension of the data (reduce the number of categories in a categorical scale) or to change the nature of the data to fit more appropriate methodologies (transform metric scales into categorical scales). It is important, however, to ensure that the new scale does not result in any transformation or loss of relevant parts of the original information, since this could substantially alter the results. Validation of the recoding is therefore one more stage in the research process.

The results in the two preceding sections provide clear proof of the adequacy of MFA for finding the answer when faced with the decision of whether to recode the original scale. If the researcher recodes the original scale correctly, the methodology provides various graphic and numerical signals which capture and measure the main similarities between the recoded scale and the original one. Thanks to this methodology, an incorrect recoding can be quickly discarded based on a simple analysis of the correlation structure of the partial factors.

Figure 6
Plane of factors 2 and 3 showing an example of erroneous recoding



Thus, it is worth emphasizing to potential users the major advantage of this validation methodology - its simplicity. The user does not need to be an expert either in factorial analysis or in multi-table analysis. All the results required can be obtained automatically using existing statistical software, some of which is freely available. The decision to recode can be adopted or abandoned after simply following the guidelines indicated by the authors in the example (and counter-example). The proposal for this methodology involves a multi-stage protocol in which the results are interrelated and must therefore be analyzed jointly and in an ordered manner before validating or rejecting the proposed recoding.

One of the most relevant contributions of this paper is that it does not propose a specific recoding criterion, but provides researchers, instead, with a tool that will, ideally, lead them towards a validated recoding. The underlying idea is to determine how far the initial relationship structure among the variables remains unaltered after recoding. The methodology is able to handle both metric and categorical variables, variables with high heterogeneity, variables with different ranges and even variables expressed in different units. The core concept behind the proposed method is the study of factor stability to ensure that the new scale has not produced distortions in the original data structure. We therefore consider it coherent with the philosophy underlying the factorial analysis of multiple tables. From the family of methods developed in recent years for the simultaneous study of multiple data matrices, the choice was MFA. Use of this method is not constrained by the nature of the original data, because it can be used for the analysis of so-called mixed data tables. Thus, one of the tables might contain what was originally either metric or categorical data and the other might be the result of recoding into fewer categories in order to reduce the dimension of the problem.

A practical case based on the opinion poll ratings assigned to four leading Spanish politicians provides a framework in which to demonstrate the suitability of the MFA recoding validation methodology. The usefulness of the procedure is then demonstrated by comparing the structures of two groups of variables, one for the ratings of the politicians on a twelve-category nominal scale (the original scale), and another for their ratings on the recoded four-category nominal scale. Recoding greatly reduces the dimension of the problem, which is an advantage in itself, without provoking any significant change in the initial structure of the data under analysis. A second, intentionally “erroneous”, recoding is carried out to demonstrate that the method has the further capabilities of detecting when a recoding is inappropriate and likely to lead to significant loss of information or biases, and of indicating the direction in which it needs to be modified.

Looking ahead, we have proposals for four future lines of research. The first and most obvious is to apply the validation methodology to quantitative variables, which was not possible in this paper due to lack of space. The context for this is the recoding of different types of household consumptions (in Euros), following an idea proposed in Abascal et al. (2006). The second, on a more theoretical level, is to carry out a deep comparative analysis of multi-table analysis methods suitable for recoding validation. These would include MFA and others that have gained some acceptance among the scientific community, such as STATIS (Lebart et al, 2006). The third involves a plan for a comparative analysis of MFA versus Bootstrap techniques, the latter currently being put forward as an alternative for the analysis of the stability of results between different codings (Lebart, 2007). Finally, we wish to compare the advantages of applying MFA or Non-linear Principal Components Analysis as the factorial method for recoding validation (Meulman et al., 2004).

Notes

[1] Recovered from http://www.cis.es/cis/opencm/ES/2_bancodatos/estudios/ver.jsp?estudio=14099.

[2] Note that this type of marking system is deeply rooted in all stages of education in Spain.

References

Abascal, E., K. Fernández, M. I. Landaluce and J. Modroño. 2001. “Diferentes aplicaciones de las técnicas factoriales de análisis de tablas múltiples en las investigaciones mediante encuestas”. *Metodología de Encuestas* 3: 251-280.

Abascal, E., I. García-Lautre and M. I. Landaluce. 2006. “Multiple factor analysis of mixed tables of metric and categorical data”. Pp. 351-367 In *Multiple Correspondence Analysis and related Methods*. London: Chapman and Hall.

Abascal, E., I. García-Lautre and F. Mallor. 2006b. “Data mining in a bicriteria clustering problem”. *European Journal of Operational Research* 173: 705-716. https://doi.org/10.1016/j.ejor.2005.10.006

Abascal, E. and V. Díaz de Rada. 2014. “Analysis of 0 to 10-point response scales using factorial methods: a new perspective”. *International Journal of Social Research Methodology* 17: 569-584. https://doi.org/10.1080/13645579.2013.799736

Abascal, E., V. Díaz de Rada, I. García-Lautre and M. I. Landaluce. 2017. “Analysis of the response structure to a set of questions with large number of scale points: a new combined metric and categorical approach”. *International Journal of Social Research Methodology* 21(4): 395-407. https://doi.org/10.1080/13645579.2017.1399620

Bell, D. S., C.M. Mangione and C.E. Kahn. 2001. “Randomized testing of alternative survey formats using anonymous volunteers on the world wide web", *Journal of the American Medical Informatics Association* nº 8: 616–620.

Bourque, L.B. and V. A. Clark. 1994. *Processing Data The Survey Example*. Newbury Park, CA: Sage.

CIS (2014) Centro de Investigaciones Sociológicas. Barómetro de julio 2014, estudio 3033.

Conrad, F. and F. Kreuter. 2015. *Questionnaire design for social surveys*, course by Coursera. https://es.coursera.org/learn/questionnaire-design

Couper, M. P., R. Tourangeau, F. G. Conrad and C. Zang. 2013. "The design of grids in web surveys", *Social Science Computer Review*, vol. 31 (3), pp. 322-341.

Couper, M. P., M. Traugott and M. Lamias. 2001. "Web survey design and administration", *Public Opinion Quarterly*, nº 65, pp. 230–253.

Dazy, F. and J. F. Le Barzic. 1996. *L´Analyse des Données Evolutives*. Paris: Technip

Escofier, B. and J. Pagès. 1982. "Comparaison de groupes de variables définies sur le même ensemble d´individus", Working paper 165, Institut National de Recherche en Informatique et en Automatique.

Escofier, B. and J. Pagès. 1990. *Analysis Factorielles Simples et Multiples*. Paris: Dunod.

Escofier, B. and J. Pagès. 2016. *Analysis Factorielles Simples et Multiples. Cours et éstudes de cas*.5e edition. Paris: Dunod.

García-Lautre, I. 2001. Medición y Análisis de las Infraestructuras: Una Nueva Metodología basada en el Análisis Factorial Múltiple. Unpublised Ph.D, thesis. Pamplona. Public University of Navarra.

García-Lautre, I. and E. Abascal. 2003. "Una metodología para el estudio de la evolución de variables latentes. Análisis de las infraestructuras de carreteras de las comunidades autónomas (1975-2000)". *Estadística Española*, Vol. 45, Nº 153, pp. 193-210.

García-Lautre, I. and E. Abascal. 2004. "A methodology measuring latent variables based on multiple analysis" *Computational statistics and data analysis* 45(3): 505-517

Greenacre, M. and J. Blasius. 2014. Visualization and Verbalization of Data. *Computer Science and Data Analysis Series*. London: Chapman and Hall/Taylor and Francis Group/CRC.

Iglesias, C.P., Y. F. Birks and D.J. Torgerson. 2001. "Improving the measurement of quality of life in older people: The york SF-12", *Quarterly Journal of Medicine* nº 94: 695–698.

Landaluce, M.I. 1995. "Estudio de la estructura de gasto medio de las Comunidades Autónomas españolas. Una aplicación del Análisis Factorial Multiple". Unpublised Ph.D, thesis. Bilbao. Public University of Basque Country.

Lebart, L., A. Morineau and M. Piron. 2000. *Statistique exploratoire multidimensionnelle*. Paris: Dunod.

Lebart, L. 2006. "Validation Techniques in Multiple Correspondence Analysis". Pp. 179-195 in *Multiple Correspondence Analysis and related Methods*, edited by M. Greenacre and J. Blasius. London: Chapman and Hall.

Lebart, L., M. Piron and A. Morineau. 2006. *Statistique exploratoire multidimensionnelle. Visualisation et inference en fouilles de données*. 4 ed. Paris: Dunod.

Lebart, L. 2007. "Which Bootstrap for Principal Axes Methods?". Pp. 581-588 i*n Selected Contributions in Data Analysis and Classifications. Studies in Classifications, Data Analysis, and Knowledge Organization,* edited by P. Birto, G. Cucumel, P. Bertrand and F. de Carvalho. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-73560-1_55

Meulman, J.J., A.J. Van Der Kooij and W.J. Heiser. 2004. "Principal Components Analysis with nonlinear optimal scaling transformations for ordinal and nominal data". Pp. 49-70 in *The sage handbook of quantitative methodology for the social sciences,* edited by D. Kaplan, Thousand Oaks, CA: Sage.

Miller, D.C. 1991. *Handbook of Research Design and Social Measurement*. London: Sage.

Revilla, M., D. Toninelli and C. Ochoa. 2017. "An experiment comparing grids and item-by-item formats in web surveys completed through PCs and smartphones", *Telematics and Informatics*, nº 34: 30-42.

Roßmann, J., T. Gummer and H. Silber. 2017. "Mitigating satisficing in cognitively demanding grid questions: evidence from two web-based experiments", *Journal of Survey Statistics and Methodology* 6(3): 376-400. https://doi.org/10.1093/jssam/smx020

Smyth, Jolene D. and K. Olson. 2016. "A Comparison of Fully Labeled and Top-Labeled Grid Question Formats". Paper presented at the International Conference on Questionnaire Design, Development, Evaluation, and Testing, Miami, FL, November 9-13, 2016.

SPAD7.3 (Software used). Data management. Analyse des Données. Data Mining Coheris-Spad.

Tien, C. 2008. "Recoded Variable". Pp. 696-697 in *Encyclopedia of Survey Research Methods* (vol. 1).

Thorndike, F. P., P. Carlbring, F.L. Smyth, J.C. Magee, L. Gonder-Frederick, L.G. Ost and L.M. Ritterband. 2009. "Web-based measurement: effect of completing single or multiple items per webpage", *Computers in Human Behavior*, nº 25: 393–401.

Toepoel, V., M. Das and A. van Soest. 2009. "Design of web questionnaires: The effects of the number of items per screen", *Field Methods*, nº 21: 200–213.

Tourangeau, R., M.P. Couper and F.G. Conrad. 2004. "Spacing, position, and order: Interpretive heuristics for visual features of survey questions", *Public Opinion Quarterly*, nº 68: 368–393.

Weisberg, F., J.A. Krosnic and B.D. Bowen. 1996. *An Introduction to Survey Research, Polling and Data Analysis,* 3 edition. California, USA: Sage.

ELENA ABASCAL is a PhD in Economics and Business. Author of several books among which stand out, *Multivariate methods for commercial research and Analysis of surveys*. She is the author of scientific publications and contributions in congresses, being the main research line the multivariate statistical methods applied to social, economic and marketing studies and the design and analysis of surveys. She is a professor in the Department of Statistics and Operational Research at the Public University of Navarra.

VIDAL DÍAZ DE RADA is a Graduate and PhD in Sociology and a Graduate in Research and Market Techniques. He has published more than a hundred papers, mainly on methodology and sociology of consumption, highlighting among his latest publications *Comparison of survey data collection methods* (REIS, 2017), *Face-to-face surveys using paper or computer. Difference in attitudes* (Papers, 2017), *Main reasons of food-away from home: A typology of spanish society* (RES, 2017) and *Fragmentación en los consumos de ocio en un contexto de crisis económica* (ICARIA, 2017). He is a professor in the Department of Sociology at the Public University of Navarra and member of*I-COMMUNITAS-Institute for Advanced Social Research*.

IGNACIO GARCÍA LAUTRE is a PhD in Business Administration and Management. He is the author of scientific publications and contributions in congresses being the main research line multivariate statistical methods applied to social studies, economic and analysis of surveys. He is a professor in the Department of Statistics and Operational Research at the Public University of Navarra.

M. ISABEL LANDALUCE is a PhD in Economics. She is the author of scientific publications and contributions in congresses, both nationally and internationally, being the main research line multivariate statistical methods applied to social studies, economic and analysis of surveys. She is a professor in the Department of Applied Economics of the University of Burgos, in the area of Quantitative Methods for Economics and Enterprise.