International Journal of Social Research Methodology, 2014
Vol. 17, No. 5, 569–584, http://dx.doi.org/10.1080/13645579.2013.799736

Analysis of 0 to 10-point response scales using factorial methods: a new perspective

Elena Abascal[a]* and Vidal Díaz de Rada[b]

[a]*Departament of Statistics, Public University of Navarre, Pamplona, Navarra, Spain;* [b]*Department of Sociology, Public University of Navarre, Pamplona, Navarra, Spain*

(*Received 29 January 2013; final version received 23 April 2013*)

This paper analyses the implications of using survey response scales with 10 or more categories which are often analysed using factorial techniques such as principal components analysis. Although it is the type of metric that determines the choice of this method of analysis, these scales, despite being considered metric, lack some metric properties, such that a score of 10 in a satisfaction survey, for example, does not represent twice the satisfaction of a score of 5. Their ordinal nature also means they cannot be treated as purely categorical. Thus, having both metric and categorical properties, they allow the use of different factor analysis techniques. The second objective is to apply different factorial techniques to a given test item in order to detect potential differences in the analytical information obtained by each method. Finally, we will discuss their relative advantages and disadvantages for use with this type of survey question.

Keywords: principal components analysis; multiple correspondence analysis; grid question; social survey measurement scales

1. Introduction and research aims

Whether the survey is aimed at gathering factual data or measuring subjective states, one of the most important questionnaire design issues is deciding the response range (Saris & Gallhofer, 2007). Thus, when surveying an issue such as frequency of attendance at entertainment events, there are three possibilities: the dichotomous response format (does attend/does not attend); the ordinal response format based on a reference time frame (less than once a month, 1–4 times a month, 5–8 times a month, once or twice a week, more than 9 times a month, more than twice a week); and, thirdly, the exact response format (number of times in the past week/month/year). While the choice of format involves only slight variations in the wording of the question,[1] it has a *transcendental* impact on the accuracy of the data. The third of the above formats provides much more accurate information than either of the others. It also allows the interviewer to return to the other formats by, say, categorizing the scores or even by classing respondents reporting less than a certain number of attendances as non-attenders.

This is one of the most widely documented areas of survey-based research (DeVellis, 1991; Krosnick & Fabrigar, 1997; Schwarz, Hippler, Deutsch, & Strack,

*Corresponding author. Email: eabascal@unavarra.es

1985) and, interestingly, it continues to draw much attention, mainly with respect to the use of the gridded response format (Krosnick & Presser, 2010; Maitland, 2009; Revilla, Saris, & Krosnick, 2009; Saris, Revilla, Krosnick, & Schaeffer, 2010, among others). In fact, a brief look at the contents page of any survey methodology journal is enough to show that–even today – there is little consensus regarding issues such as: the number of scoring items to include in the response grid (Maitland, 2009); the wording of extreme alternatives (Garg, 1996; Lyons, 1998; Pearcy et al., 2001; Saris et al., 2010); the appropriateness of labels for numerical categories (Dillman, Smyth, & Christian, 2009); the inclusion (or not) of a middle alternative (Moors, 2008; Bishop, 1987, 1990); the choice between unipolar- or bipolar-coding (Schwarz et al., 1985; Schwarz, Knäuper, Hippler, Noelle-Newman, & Clark, 1991); the choice between a vertical or horizontal scale format (Christian & Dillman, 2004); the order of presentation in an agree/disagree format; and whether to use top- or bottom-coding (Hartley & Betts, 2010), etc.

This paper contributes to the research on this subject by studying the implications of the response range in scales with 10 or more categories. Attention is focused on value scales of 0–10 and 1–10 points, and the type of coding system used when analysing the data with factorial techniques such as principal components analysis (PCA). The second objective is to test for variance in the results for a given item (grid question) attributable to the type of factorial technique used to analyse the data; specifically, PCA or multiple correspondence analysis (MCA). This paper discusses the advantages and disadvantages of these two techniques for the analysis of this type of question in a politician rating survey taken by a representative sample of the Spanish population.

The paper is structured as follows. It begins with a brief overview of the state of the research on response scales and continues with a presentation of the basics of PCA and MCA. Section 4 provides the details of the survey on which the two analysis techniques are tested followed by a comparative analysis of two factorial procedures. The main conclusions from the analysis are given in Section 4.

2. Response scales

The study of measuring scales for survey questionnaires can be approached from one of two perspectives, depending on whether the focus is on the *measuring of* phenomena or the *numerical properties* of the responses (and the alternatives for analysing the obtained data). The following is a detailed examination of each one.

2.1. Measurement issues for consideration

If survey instruments (questionnaire items) are to provide valid, reliable measures, attention must be paid to issues as varied as the response range; the inclusion, or otherwise, of a middle alternative; the wording and terminology of the end-point labels the choice between top and bottom coding or labelling; the use of a horizontal/vertical scale format; varying or equal distances between options; the inclusion of one or several ideas within the scale; and, in recent years, alternatives for the use of graphics in online surveys (Couper, 2009; Tourangeau, Couper, & Conrad, 2004, 2007). Due to the characteristics of the question that is analysed in this paper, attention must be paid to the first four of the above issues.

2.1.1. Scale range, the number of alternatives

One of the issues that has received most attention in the existing literature is the number of response categories (Krosnick & Fabrigar, 1997; Krosnick & Presser, 2010) to include in the scale, both with a view to measurement accuracy (reducing measurement error by respondents) and to the reduction of questionnaire completion time. It is believed that an increase in the number of categories reduces reliability, although some researchers have reported high levels of reliability in dichotomous scales (Alwin, 2007). One of the most frequently cited works, Cox (1980), recommends the use of 7-point scales because they offer high reliability, a low percentage of undecided responses, and ease of discrimination (for respondents) between the scale values. In a comparative analysis of 5-point and 11-point scales, Dawes concludes that the two produce similar averages; although the latter offers more variance. Similar findings were obtained in another study comparing 5, 7 and 10-point scales, where the first two were found equally reliable and the 10-point scale slightly less so (Dawes, 2008). A more recent study by Saris and Gallhofer (2007, p. 147) recommends the use of scales with 11 categories. While this paper uses a 0–10-point scale, it should be noted that a comparison of 0–10- and 1–10-point scales by Schwarz et al. (1991) showed that the former generated higher scoring patterns.

2.1.2. Inclusion or otherwise of a middle alternative

A second key issue in the research on this subject has to do with whether or not the scale has a mid-point. A mid-point provides an alternative for people with no knowledge of or defined attitude towards the issue under discussion, although it may also be selected by respondents who, despite having a clearly defined attitude, cannot be bothered to think, are trying to finish sooner or, simply, wish to avoid imagined confrontations with the interviewer (survey satisficing). The 11-point scale (0–10) used in this paper has a middle point of 5 between the two extreme alternatives.

Other important issues are effects resulting from the *wording of the extreme alternatives* (Schwarz et al., 1991, p. 571) or the *coding format (polar/bipolar)*. Several studies have also shown that the wording of verbal end-point labels affects response distribution, leading to the conclusion that respondents pay far more attention to the meaning of the terms employed than might be expected (Schwarz et al., 1991, p. 571). These issues will not be addressed here because, in the survey question used for this study, the respondent is instructed to select 0 to express a very poor rating of the politician, and 10 to express a very high rating. It is not the main objective of the paper.

2.2. Analysis techniques and response categories

Our focus will be on the measurement level of each variable, that is, whether it is quantitative (metric) or qualitative (non-metric). The term quantitative refers to interval and ratio variables; the term qualitative refers to nominal and ordinal variables (Stevens, 1946, 1951). It is nevertheless worth recalling at this point that measurement by intervals or ratios – which the majority of statistical tests allow – is not widely used in social research, where quantifiable factors are few and far between. Due to this and the increasing use of factorial techniques such as PCA, a large number of researchers use questions requiring respondents to give ratings on scales of 1–7, 1–10, or 0–10, since it is considered that distributions of more than five categories can be treated as metric (O'Brien, 1979).

Metric treatment of this type of scales, nevertheless, has its drawbacks. One of the metric properties of a rating scale is that it should be possible to find the numerical distance between observations. Is it right to assume, therefore, that a respondent who gives a satisfaction rating of 8 is twice as satisfied as one who gives a rating of 4? If two respondents give the same rating, does this mean that they feel exactly the same level of satisfaction? The generally observed tendency is for survey respondents to give relatively high scores.

The aim of this paper is to analyse the extent to which the demands placed on the data collection method by the choice of data processing technique lead to major loss of information quality. Questionnaires containing multiple-category items require a degree of effort that some respondents find impossible in the circumstances in which they are asked to complete the survey. As an alternative to using PCA, which requires a scale of at least 1–7, it is possible to use scales with fewer categories, which are easier to answer and can be analysed using a qualitative variable analysis method such as MCA. This type of procedure can also increase data collection capacity since, by treating the variables as non-metric, it is able to detect both linear and non-linear relationships.

3. The characteristics of PCA and MCA

PCA and MCA are multivariate descriptive statistics techniques that can be used to reduce large tables of data into two-dimensional graphs typically featuring numerical labels to aid interpretation. The two procedures share a common principle: factor analysis of a set of variables. In PCA, all the variables are metric, while in MCA they are all categorical.

These techniques improve upon variable correlation analysis by revealing the actual structure of the relationship. The aim in both cases is to find sets of synthetic variables or factors that are mutually orthogonal and retain the maximum amount of information from the original dataset. In addition, they graphically portray the correlation between the variables analysed and the observations, thus providing a global view of the phenomenon under study.

Once plotted, the graphs can be used to display other data (Lebart, Morineau, & Piron, 2000). These additional variables are called supplementary or illustrative variables as opposed to the active variables of the analysis, which determine the solution space. Supplementary variables have no influence on the geometric orientation of the axes; rather, they support and complement the interpretation of the configuration of active variable categories (Greenacre & Blasius, 2006).

Both PCA and MCA provide graphic displays of the observations, variables and categories in a survey, from which it possible to examine the relationships that exist between them and visualize the most relevant information contained in the data (Lebart et al., 2000; Abascal & Grande, 2005). The aim is to obtain an overall picture, a global representation of the response patterns.

3.1. Characteristics of normed PCA

- It treats the variables as metric, thus assuming equidistant intervals between category values, such that equal distances between response categories imply equally spaced degrees of the characteristic under analysis. Correspondingly, zero is taken to stand for a total lack of the said characteristic.

- It analyses the linear correlation between variables, while the non-linear relationships so frequently found in the rating questions that are typical of sociological research remain undetected.
- It excludes the no-answer and do not know response categories and thus does not capture all the available information from the survey since it misses some respondent data.
- It computes the Euclidean distance between two respondents, such that an equal difference between their responses indicates equal distance between their characteristics. There may still be a very great distance between two respondents that differ in only one response, if it is one in which their scores differ widely. The distance in this case could be the same as between two respondents showing a *slight* difference in all their responses.[2]
- It interprets distances between respondents in terms of the similarities in their responses over the entire set of variables.

3.2. Characteristics of MCA

- The method serves to capture both linear and non-linear associations between the variables (Greenacre & Blasius, 2006).
- There are two ways to consider (interpret) similarity between categories:
    - Similarity between categories increases with their simultaneous presence (or absence) in a large number of respondents (regardless of other categories); and
    - Similarity between categories increases as their association with the same categories of other variables increases.
- Distances between respondents increase with the number of categories in which they differ.
- Each response category is assigned a weight according to the number of respondents who select it, that is, its frequency. Thus, when repeatedly selected by the same small group of respondents, low-frequency categories can contribute too much to the first factor and thus make it less meaningful than the second.
- The more response categories there are for a question, the more it contributes to the inertia and the larger the number of axes on which it may load. In the case in hand, the number of response categories is the same for all the active variables. The dimensions of the problem are greater because every question involves as many variables as it does categories.
- Do not know and no-response are considered as two different categories and these receive the same treatment as all the rest. They are not treated as missing values as they are in PCA. Besides, the non-response patterns can be explored by focusing on their relationships with other categories. (For more information on the method see Grange & Lebart, 1993; Greenacre & Blasius, 2006).

4. Case study: centro de investigaciones sociológicas (sociological research centre) February barometer, case number 2859

The survey universe was all Spanish citizens (except residents of Ceuta and Melilla) aged 18 years and over. In-home interviews were conducted using face-to-face questionnaires. During the second week (7th–16th) of January 2011, 2478 interviews took place, the respondents having been selected by multi-stage sampling, in which the primary units (municipalities[3]) and secondary units (sections) were proportionally randomly selected and individuals were then selected by performing a random walk through each section with quota on age and sex. A multistage sample, stratified by conglomerates, in which the strata were obtained by crossing province with town size divided into five categories. The questionnaire took an average of 18.5 min to complete.

The main part of the questionnaire is made up of questions relating to political attitudes, including one in which respondents were asked whether they had heard of and how they rated certain political leaders. They had to base their responses on a value scale of 0–10 where 0 meant very poor and 10 meant very good (the wording of the question is shown in Box 1). Based on the considerations discussed in Section 2, respondents were instructed to give their scores on an 11-category response scale, with the mid-point (5), such that the ratings ranged from very poor to very good, using a unipolar-coding scheme of 0–10. It is worth pointing out that since this was the first question of its type to appear in the questionnaire, respondents may initially have been somewhat disconcerted by it.

Square 1. Question in the questionnaire.

Please indicate whether you know each of the following political leaders and then rate their political performance on a scale of 0 to 10, where a score of 0 means 'very poor' and a score of 10 means 'very good'.

Source: Centro de Investigaciones Sociológicas (2011: Survey number 2859).

A survey covering national, regional and local politicians produces a considerably lower response rate on some politicians (Table 1). After a preliminary

Table 1. Showing the number and percentage of respondents claiming to know each politician, together with the means and standard deviations of all the results.

	Known by (number of cases)	Known by (%)	Mean score	Standard deviation
J.L. Rodríguez Zapatero	2478	100	3.30	2.659
Mariano Rajoy	2476	99.9	3.25	2.519
Cayo Lara	1139	46.0	3.04	2.255
Josep A. Durán i Lleida	1585	64.0	4.40	2.462
Íñigo Urkullu	1014	40.9	2.96	2.290
Rosa Díez	1761	71.1	3.75	2.341
Joan Puigcercós	1154	46.6	2.69	2.203
Guillermo Vázquez	473	19.1	2.71	2.149
Paulino Rivero	652	26.3	3.38	2.326
Uxue Barkos	549	22.2	3.11	2.465
Yolanda Barcina	488	19.7	2.92	2.405

analysis, therefore, it was decided to reduce the scope of the survey to politicians from parties of nationwide relevance: José Luis Rodríguez Zapatero (soft left wing, the president of the nation and president of the Spanish socialist party, PSOE, at the time of the survey), Mariano Rajoy (right wing, the president of the main opposition party, Popular Party, PP), Cayo Lara (left wing, the president of the second most important opposition party, United Left, IU/ICU) and Rosa Díez (an ex-socialist, a co-founder of the UPyD – named the Progressive and Democratic Union, a Member of the European Parliament, and a regular media figure). By using only respondents who claimed to know all four politicians, we reduced the sample to 1058 individuals. Given that the aim of the paper is not to draw estimates from the population, but to perform a comparative analysis of two data collection modes, we do not consider potential small-sample bias to be an issue in this case. Table 1 reveals that the leaders of the main political parties are well known but poorly rated by the respondents.

Examination of the frequencies of the politician ratings raises the same uncertainties that emerge when these variables are treated metrically. There are a large number of zero ratings, but the rest of the frequency distribution shows an increasing trend up to and including point 5 on the rating scale and a decreasing trend thereafter. The difference between a score of 0 and a score of 1 appears to be greater than between –say– 4 and 5 or 5 and 6.

Observation also shows that most of the distributions are positively skewed, with the peak to the left of centre, and that overall scoring is lower than in previous similar surveys (Mata López, Luque Castillo, & Ortega Ruiz, 2010). We suspect that this may be partly due to the state of the Spanish economy.

In January 2011, when the survey took place, the country had an unemployment rate of 21.29% (4.910.000 unemployed); and only 12 months previously, the government had admitted that the country was undergoing one of its worst ever economic crises. In fact, only four months earlier (August 2010), the population had suffered some of the most severe cutbacks in social spending in the country's recent memory, together with a reduction in public workers' salaries. This had eroded public confidence in political leaders and, when this survey asked respondents (in an open-ended question) to name the country's three most serious problems, 84% of the respondents cited unemployment, 54% the economy and 21% politicians and political parties.[4]

5. Results obtained from the factor analyses

Survey data based on rating scales with scores ranging from 0 to 10 are suitable not only for PCA (which treats the information as metric data) but also for MCA (because the scores can be treated as categories). Comparison of the results produced by each of these procedures will reveal their respective advantages.

The analyses were performed on the ratings of four politicians using the SPAD program. A set of variables including socio-demographic characteristics (gender and ideology) and other data (voting recall, personal financial situation and opinion of the country's economic situation) were included to act as supplementary variables. These, while having no influence in the extraction of the factors, when projected onto them, support and complement the interpretation of the configuration variable categories (political evaluation) and their positions.

5.1. *Ratings (of political figures) treated as metric variables (PCA)*

The recommended method for data analysis when treating a scale of 0–10 as metric is normed PCA, which enables simultaneous analysis of the four sets of ratings in order to detect the factors of differentiation between respondents, and identify any response categories that tend to be selected by the same respondents. Under this method, any do not know or no-answer responses given by respondents when asked to rate a politician are treated as missing values and replaced with the scale's mid-point score. Two factors of differentiation between respondents emerge:

> The first is interpreted as a *size factor*. This shows the tendency of respondents to give all high (low) ratings. By separating respondents who give all the politicians an above-average score – the positive side of the scale – from those who give them all a below-average score (on the negative side), it shows that some respondents are generous in their scoring while others are stricter. This factor accounts for 38.58% of the variance in the data.

> The second is an ideology factor. This separates respondents who give higher ratings to right wing politicians (left-hand side of the factor) from those who give higher ratings to left-wing politicians (right-hand side). This factor accounts for 32.36% of the variance in the data.

The similarity in the percentage of variance explained by each of these factors calls for their joint interpretation (as shown in Figure 1), which reveals the lack of any relationship between respondents' ratings of Rajoy and Zapatero (there is a medium distance between them). Factor 2, meanwhile, reveals a positive relationship between their ratings of Rajoy and Díez, who share a conservative ideology, and their ratings of Zapatero and Lara, who are ideologically left wing.

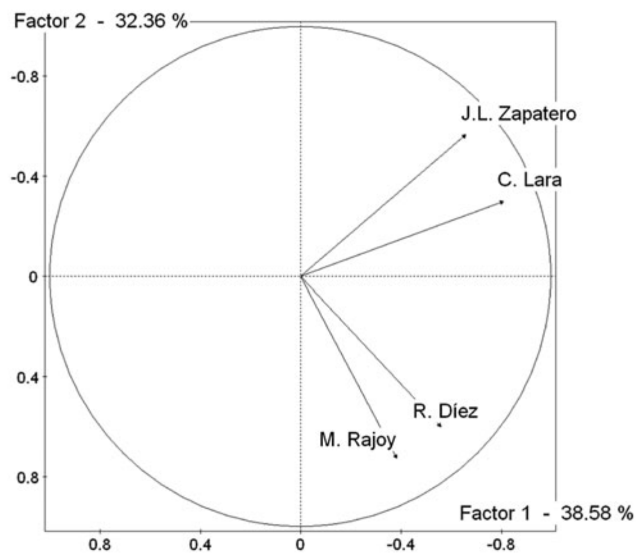


Figure 1. Principal plane provided by PCA: representations of variables.

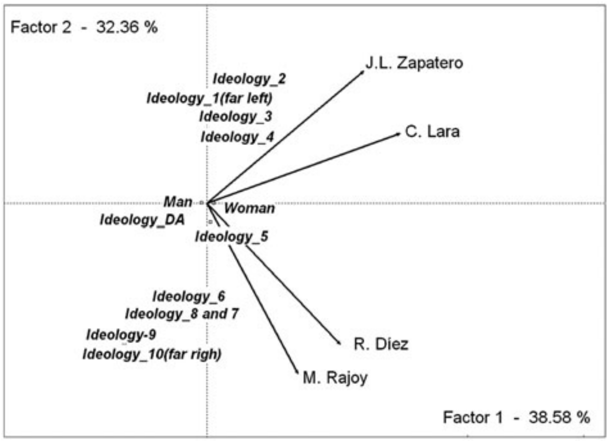


Figure 2. Principal plane provided by PCA: representations of categories for two supplementary questions and the metric variables.

On the same plane, it is possible to represent the categories of the categorical variables, treating them as supplementary variables. The coordinate of a category is its centre of gravity (weighted average/mean) on the factor representing the group of respondents who selected that response category. Each variable defines as many groups of respondents as there are response categories, and the representation of a category is interpreted as the position of the average respondent in that group. Figure 2 displays the respondents' gender and ideological categories. The ideology data were obtained by asking the respondents to describe their perspective on a scale of 1–10,[5] where 1 indicates left wing and 10 indicates right wing.

Figure 2 confirms the above interpretation of factor 2 in that the left-wing leaders are more highly rated by those respondents declaring a left-wing perspective, and the right-wing leaders by those that declaring a right-wing perspective. The central positioning of gender shows that men's ratings do not differ significantly from those given by women.

5.2. Ratings (of political figures) treated as categorical variables (MCA)

In the analysis of the ratings described in this section, scores of 0–10 as 11 are treated as different categories. Replies of do not know and no-answer are combined into a single category (DK/NA), with the same potential role – a priori – as all the rest.[6] The ratings of the four politicians are analysed simultaneously using MCA in order to find the factors of differentiation among the respondents, and identify and classify the response options. This global analysis approach is in fact the main contribution of this paper relative to previous research, which has used partial (variable-by-variable) analysis.

The observed factors of differentiation this time are not the same as those obtained in the PCA described in the previous section

The first is a non-response factor, which separates those who give no answer or say they do not know (left-hand side) from the rest. Figure 3 shows that, overall, it is the same respondents who select the non-response option. At the same time, however, it shows that the non-responses to the invitation to rate Zapatero and Rajoy are further away from the centre, because they are rare (less frequent) categories. This behaviour remains undetected in PCA. This first factor is of little interest in the analysis of the survey structure because what it reveals is the obvious.

The second factor is interpreted as a *reprobation score*. It separates the respondents who give a very poor rating (0) from the rest. The observations of the extreme response category very good (10) have a very low frequency and do not contribute to the factor, but appear in association not only with nearby categories such as 9, but also with those at the opposite extreme, that is, the zero-rating observations.

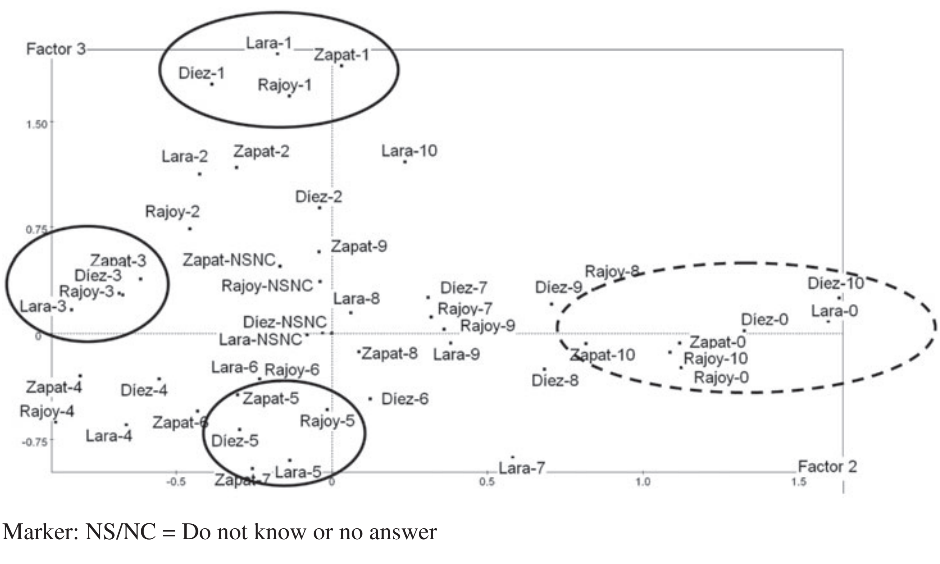
The third is defined as the *score moderation factor* or discretion factor. It distinguishes between those giving scores towards the lower end (1 or 2, but not 0) or the upper end (8 or 9 but not 10) of the scale from those giving intermediate scores.

The number of factors is very high because the 12 response categories for each question are orthogonal. The percentage of inertia is not given here because it is not a good measure of the quality of the analysis.

The representation of factors 2 and 3 in Figure 4 reveals the proximity among categories of the same value, that is, the proximity of all 1 responses or 5 responses irrespective of which politician is being rated (except in the case of Cayo Lara).

Zapat-DKNR Rajoy-DKNR Diez-DKNR Lara-DKNR 0 the rest

Figure 3. Representation of the do not know/no answer categories on the first factor.

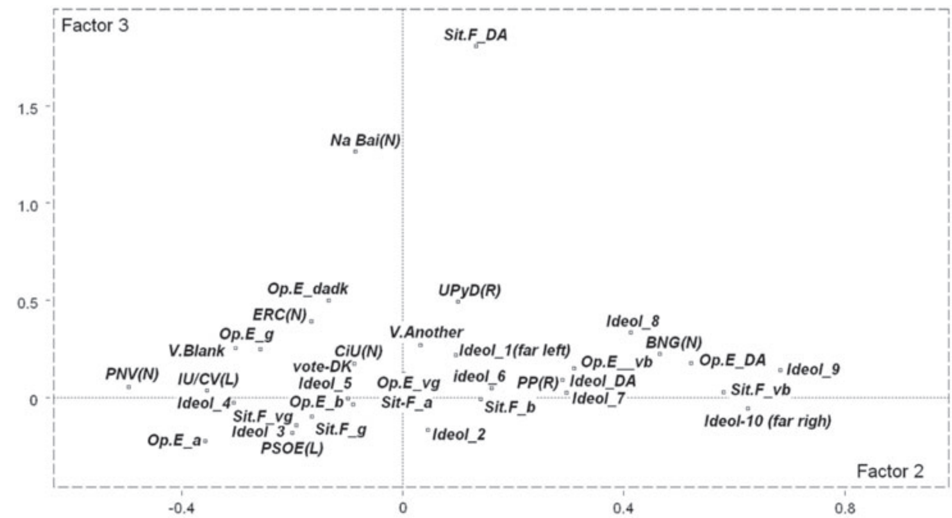


Marker: NS/NC = Do not know or no answer

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

This reveals that respondents have a tendency to repeat the same scores for all the politicians. In other words, having selected a certain response option for the first, they proceed to choose the same one for all the rest (groupings enclosed in circles).

It can be seen that the response categories appear in ascending order from 0 to 10, forming a kind of parabola or horseshoe shape, with the ends almost meeting: that is, they show the typical Guttman effect. This shows that the responses do not exhibit the linear relationship (which is analysed using PCA). That is, respondents who select one of the extreme options (0 or 10) for one politician usually select that same one for the rest, but also frequently select the opposite extreme. Thus, for example, of the 226 respondents who gave Rajoy a score of 0, 86 also gave Zapatero a score of 0 while 7 of them gave Zapatero a score of 10. It would be impossible to capture this information using PCA.



Marker:

Voting recall from last elections:

- Right wing (**R in chart**): PP (Popular Party), UPyD (Progressive and Democratic Union).
- Left wing (L **in chart)**: PSOE (Socialist Party), IU/ICV (United Left)
- Nationalist Party (N**)**: BNG, ERC, Na Bai, PNV and CIU.

Vote-NA (No answer), V. Another (Another party), V. Blank (Blank vote), No vote (Refused to vote) vote-DK (Don´t know).

Opinion of the country's economic situation: Op.E-vg (very good), Op.E -g (good), Op.E -a (average), Op.E -b (bad), Op.E -vb (very bad), Op.E -nadk (No answer,don´t know). Ideology: Ideol (Ideology-i).

Declared personal financial situation: Sit-F-vg (very good), Sit-F-a (average), Sit-F -b (bad), Sit-F -vb (very bad), Sit-F -nadk (No answer,don´t know)

Figure 5. Plane of factors 2 and 3 from the MCA, showing the supplementary categories.

Analysis of the supplementary variables reveals a gender difference in the first factor, in that the frequency of do not know no-answer responses is higher in women than in men. The second factor is more strongly associated with ideology but not with extreme ideological positions.

Figure 5 shows the supplementary response categories on voting recall from last elections, ideological perspective, opinion of the country's economic situation and declared personal financial situation. Analysis of the factorial plane, including the usual analysis quality indicators,[7] provides a socio-political profile of the different opinion groups. Thus, those who describe their personal financial situation as good appear close to the intermediate scores and far away from the extreme scores (0 and 10). The complete opposite is found for those describing their personal financial situation as bad or very bad, who give the politicians extreme scores.

6. Conclusions

The application of two different analysis techniques on the same response grid has enabled us to illustrate the importance of the choice of scale and the decision to analyse the resulting survey data metrically or qualitatively.

6.1. With respect to the choice of scale

Most textbooks on the subject recommend the use of scales with 7 points or less, not only for the sake of reliability and validity (Krosnick & Fabrigar, 1997), but also to make it easier for respondents to discriminate between response categories. The use of scales with more than 10 points – such as that used in this paper, which has a range of 0–10 – may pose problems for some respondents, even for a question as simple as how they rate their country's political leaders. Although this can result in a higher non-response rate on some items, it is more usual for respondents – albeit unconsciously – to select the same category every time. This is one of the manifestations of the *Survey Satisficing* theory (Krosnick, 1991), and is known as *non-differentiation* of response.

Having settled the issue of the number of response categories, the next aspect to consider is whether or not the scale should have a mid-point. Scales with an odd number of points provide a mid-point which is useful for respondents who are reluctant to think about the issue at hand or express their opinion of it in front of the interviewer. In this study, around 15%, on average, of those interviewed selected this option. The percentage was slightly lower in the rating of the president of the government and higher in the rating of Cayo Lara and Rosa Díez (17.5 and 20% respectively). The high frequency of mid-point scores (5) for lesser-known leaders suggests that this response was used mainly by interviewees with no clear opinion.

6.2. On methods of analysis

PCA treats a survey response scale as metric and, since it is based on the identification of linear relationships between variables, it fails to detect other types of (non-linear) relationships that are often present in responses on this kind of issues. Furthermore, PCA does not specifically analyse non-response, or other *imprecise* categories, such as 'do not know'. MCA considers these as response categories in their own right. Therefore, the key differentiating factor between respondents in this case is whether they do or do not manifest an opinion about the politicians (does not know or does not answer).

By treating the responses as category variables, MCA is able to detect any kind of relationship between response categories, a highly desirable feature in cases such as the one in hand where the relationship is approximately linear at the intermediate points of the scale but not at the extremes, where some respondents are found to give one politician a score of 10 while giving another a score of 0, thus breaking the linearity. MCA is also more appropriate in contexts where a group of respondents all select the same response option in every question but one, as in this case, where there is a subgroup of respondents that selects an extreme category (such as 10) and a subgroup that selects the opposite extreme (i.e. 0). This kind of situation remains undetected when using PCA.

Another context in which MCA is the better option is when there are not many categories. The survey used for our analysis, where there are 12 categories per variable and none of the questions allowed two responses, the categories are orthogonal and the dimensions of the problem are large enough to generate numerous factors. When working with large numbers of categories, the first factors can reveal rare respondents (that is, very different from the majority but similar to each other). This detracts greatly from the relevance of these factors for interpretative purposes. It is therefore preferable to use MCA with a shorter ranging scale, such as a 1 to 5-point scale, which generates fewer factors from a larger number of categories, thus enabling an easier and richer interpretation.

6.3. On the results obtained

The first factor in the analysis of components (PCA) is the *size factor*, which refers to the tendency to give high (low) ratings. The results of this analysis reveal that respondents show a general tendency to give similar ratings for all the politicians. This coexists with, as the MCA shows, a minority response pattern of strong support to one politician coupled with the condemnation of the rest.

Furthermore, the number of observations per politician differs because the do not know/no-answer response categories are considered missing data. The data include 1027 observations for Zapatero and only 738 for Cayo Lara. In MCA, however, all politicians have 1058 observations, which include the do not know/no-answer response categories. PCA results in a loss of information that does not occur with MCA, with the result that the first factor obtained in MCA is different.

As a final point, let us mention that it is difficult for respondents to answer surveys using scales of 0–10 because it is unclear how far consecutive scores differ. For instance, the difference between 0 and 1 is much greater than between any other two consecutive scores, because, as well as the quantitative difference, there is also a qualitative difference (whether to score or not). In surveys aimed at identifying behaviour patterns. It is better to use a method that is capable of picking up both quantitative and qualitative variance.

One option is to regroup the categories to reduce the number of dimensions and run an MCA. The recoding should not be automatic (based on the combination of consecutive categories). It needs to take into account the distribution of frequencies because the distance between consecutive categories is not always the same. For example, although there is a big difference between 0 and 1, there may not be much difference between 5 and 7. One alternative is first to run a learning process on the data-set and define the data-recoding method after an initial analysis.

By applying the two analysis techniques to the same response grid, we have been able to show the importance of the choice of response scale and the decision whether to treat the data metrically or qualitatively. Although MCA is the most appropriate technique, it can present problems. We would therefore recommend using it after first reducing the number of categories.

Notes

1. The dichotomous response answers the question Do you attend the theatre?, while the ordinal response refers to the number of attendances during the reference time frame: How many times a month do you attend the theatre: less than once a month, between 1 and 4 times a month, between 5 and 8 times …? The third type requires the respondent to state the number of times he/she has attended the theatre in – say – the last month.
2. Examples of two respondents giving four scores:
   Respondent 1 (4, 5, 6, 7) respondent 2 (4, 5, 6, 5) the squared distance = 4
   Respondent 1 (4, 5, 6, 7) respondent 2 (5, 6, 7, 8) the squared distance = 4
3. The sample covered 237 municipalities in 48 province.
4. These percentages add up to more than 100 because respondents were allowed to name up to three issues.
5. This question was asked at the end of the questionnaire when the respondent was given a card depicting the scale with the first and last points labelled left wing and right wing respectively.
6. Note that it is the same question as analysed in the previous section, but this time these responses are treated as different categories (not as missing values).
7. Specifically, the *p*-value for the tests of significance of the coordinates on each axis and the contributions to inertia.

Notes on contributors

Elena Abascal, PhD in Economics and Business, Tenured lecturer in Statistics and Operations Research. Her research interests lie in multivariate data analysis and survey analysis.

Vidal Díaz de Rada, PhD in Sociology, Bachelor's degree in Marketing Research, lecturer in the Public University of Navarra Department of Sociology. His research interests lie in the methodology and sociology of consumer research and the study of consumer behavior.

References

Abascal, E., & Grande, I. (2005). *Análisis de Encuestas* [Survey data analysis]. Madrid: ESIC.

Alwin, D. F. (2007). *Margins of error: A study of reliability in survey measurement*. New York, NY: Wiley.

Bishop, G. F. (1987). Experiments with the middle response alternative in survey questions. *Public Opinion Quarterly, 51*, 220–232.

Bishop, G. F. (1990). Issue involvement and response effects in public opinion surveys. *Public Opinion Quarterly, 54*, 209–218.

Christian, L. H., & Dillman, D. A. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Quarterly, 68*, 57–60.

Couper, M. P. (2009). *Designing effective web surveys*. Cambridge: Cambridge University Press.

Cox, E. P. (1980). The optimal number of response alternatives for a scale: A review. *Journal of Marketing Research, 12*, 158–167.

Dawes, J. (2008). Do data characteristics change according to the number of points used? An experiment using 5-point, 7-point, and 10-point scales. *International Journal of Market Research, 50*, 61–70.

DeVellis, R. F. (1991). *Scale development. Theory and applications*. London: Sage.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail and mixed-mode surveys: The tailored design method* (3rd ed.). New York, NY: Wiley.

Garg, R. (1996). The influence of positive and negative wording and issue involvement on responses to likert scales in marketing research. *International Journal of Market Research, 38*, 15–25.

Grange, D., & Lebart, L. (1993). *Traitements statistiques des enquêtes* [Treatment survey statistical]. Paris: Dunod.

Greenacre, M., & Blasius, J. (Eds.). (2006). *Multiple correspondance analysis and related methods*. London: Chapman an Hall/Taylor and Francis.

Hartley, J., & Betts, L. R. (2010). Four layouts and a finding: the effects of changes in the order of the verbal labels and numerical values on likert-type scales. *International Journal of Social Research Methodology, 13*, 17–27.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*, 213–236.

Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. In L. E. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 141–164). New York, NY: Wiley.

Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. Marsden & J. Wright (Eds.), *Handbook of survey research* (pp. 263–313). Bingley: Emerald Group.

Lebart, L., Morineau, A., & Piron, M. (2000). *Statistique exploratoire multidimensionnelle* [Exploratory analysis of multivariate data]. Paris: Dunod.

Lyons, W. (1998). Beyond agreement and disagreement: The inappropriate use of Likert items in the applied research culture. *International Journal of Social Research Methodology, 1*, 75–83.

Maitland, A. (2009, June). How many scale points should I include for attitudinal questions? *Survey Practice*, Retrieved May 22, 2012 from http://surveypractice.wordpress.com/2009/06/29/scale-points/

Mata López, T., Luque Castillo, F. J., & Ortega Ruiz, M. (2010). *La percepción de los ministros del gobierno de España (1984–2009)* [The perception of government ministers in Spain (1984–2009)]. Madrid: CIS.

Moors, G. (2008). Exploring the effect of a middle response category on response style in attitude measurement. *Quality and Quantity, 42*, 779–794.

O'Brien, R. M. (1979). The use of Pearson's *r* with ordinal data. *American Sociological Review, 44*, 851–857.

Pearcy, D., & Reinecke, F. L. (2001). Four subtle sins in scale development: some suggestions for strengthening the current paradigm. *International Journal of Market Research, 43*, 12–23.

Revilla, M., Saris, W. E., & Krosnick, J. A. (2009). *Choosing the number of categories in agree-disagree scales. RECSM (Research and Expertise Centre for Survey Methodology)*. Working Papers, no. 5, Retrieved May 15, 2012 from www.upf.edu/survey/

Saris, W., & Gallhofer, I. N. (2007). *Design, evaluation, and analysis of questionnaires for survey research*. New York, NY: Wiley.

Saris, W. E., Revilla, M., Krosnick, J. A., & Schaeffer, E. M. (2010). Comparing questions with agree/disagree response options to questions with item-specific response options. *Survey Research Methods, 4*, 61–79.

Schwarz, N., Hippler, H. J., Deutsch, B., & Strack, F. (1985). Response scales: Effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly, 49*, 388–395.

Schwarz, N., Knäuper, B., Hippler, H. J., Noelle-Newman, E., & Clark, L. (1991). Rating scales: Numeric values may change the meaning of scale labels. *Public Opinion Quarterly, 55*, 570–582.

SPAD 7.3. Coheris Spad. *Data management. Analyse des Données*. Data Mining Coheris-Spad. Retrieved from http://www.spad.eu.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*, 677–680.
Stevens, S. S. (1951). Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 55–68). New York, NY: Wiley.
Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position and order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly, 68*, 368–393.
Tourangeau, R., Couper, M. P., & Conrad, F. (2007). Color, labels, and interpretative heuristic for response scales. *Public Opinion Quarterly, 71*, 91–112.