

Quality of Data from Matrix Questions in Face-to-face Surveys: A Compared Study Using Paper and Electronic Questionnaires

Calidad de los datos de preguntas de batería en encuestas presenciales: una comparación de un estudio con cuestionario en papel y en formato electrónico

Vidal Díaz de Rada

Key words

- Questionnaires
- Computers
- Data Collection
- Response effects

Palabras clave

- Cuestionarios
- Ordenadores
- Recogida de datos
- Efectos de respuesta

Abstract

This study compares the quality of responses to eight matrix questions (36 variables) included in a survey administered on paper and by computer in two equivalent samples. Three criteria have considered: 1) the average of non-response questions for each mode; 2) the repetition (or not) of the same response in different question items (differentiation); and, 3) the choice of extreme and middle response alternatives on the scale (easy responses). The paper questionnaires provided a greater number of non-response questions, with no change in non-response rates when taking into account scale range: 4, 5, 10 and 11 categories. The electronic questionnaire showed slightly greater differentiation and fewer extreme responses.

Resumen

En este trabajo se considera la calidad de respuesta de ocho preguntas de batería (36 variables) incluidas en un cuestionario que es administrado en papel y con ordenador en dos muestras equivalentes. Se comparan tres criterios: 1) promedio de preguntas no respondidas en cada modalidad; 2) repetición –o no– de la misma respuesta en varios ítems de una pregunta (diferenciación); y, 3) elección de extremos y punto central (respuestas fáciles). Los cuestionarios en papel proporcionan un mayor número de no respuestas, falta de respuesta que no varía cuando se considera la amplitud de la escala: 4, 5, 10 y 11 categorías. El cuestionario con ordenador presenta una –ligera– mayor diferenciación y unas respuestas menos extremas.

Citation

Díaz de Rada, Vidal (2015). "Quality of Data from Matrix Questions in Face-to-face Surveys: A Compared Study Using Paper and Electronic Questionnaires". *Revista Española de Investigaciones Sociológicas*, 152: 167-178. (<http://dx.doi.org/10.5477/cis/reis.152.167>)

Vidal Díaz de Rada: Universidad Pública de Navarra | vidal@unavarra.es

INTRODUCTION¹

Face to face surveys dominated information collection means until the late 20th century; however, new innovations have led to a decrease in their current use. The use of increasingly smaller and more powerful computers has allowed for the substitution of paper (for questionnaires) by computers, resulting in notable improvements in the information collection as it permits a greater concentration of the interviewer in the interview situation and a verification of the information when it is collected.

This study compares the *quality* of the information included in the matrix questions, comparing paper questionnaires and computer-based questionnaires in studies 2893 and 2894 of the Centre for Sociological Research (hereinafter, the CIS). The quality measurement was made by considering three indicators: 1) the average of the matrix questions that were not responded to in the questionnaires completed with each method, differentiating between “don’t know” and “no response”, 2) the *differentiation* capacity, that is, that the interviewed subjects do not systematically select the same response category; and 3) the selection of “easy responses”.

FACE TO FACE, COMPUTER-AIDED SURVEY.

Face to face surveys are data collection measures in which an interviewer administers a structured (or partially structured) questionnaire to an interviewee within a limited period of time and in the presence (typically the home) of the interviewee (De Leeuw, 1992). The

application of computer advances to the discipline has led to the development of the *computerized* interview (Keller, 1995) which seeks to “integrate and automatize the survey process to the maximum” (Martin & Manner, 1995). Over the 21st century, advances in hardware have led to the creation of (mini) computers that are very lightweight and offer autonomy, permitting the use of *Computer Assisted Telephone Interviewing*, (hereinafter, CATI) in the face-to-face interview; substituting the traditional pencil and paper questionnaire with an “integrated” questionnaire that permits the verification and taping of the information while the questionnaire is being answered (Martin & Manners, 1995).

CAPI improves data collection, reducing measurement errors (Cea D’Ancona, 2012) as it eliminates the “contamination” of some questions with others and integrates the data processing during the interview. That is, all of the characteristics of the CATI system –related to the interview process– are used by the CAPI system computer equipment (Couper, 1996). However, the novelty of the computer assisted personal interview is that it combines the advantages of the CATI to the personal interview, thus eliminating certain negative elements of the telephone interviews such as the impossibility of showing objects to the interviewee (Gordoni et al., 2012), the large number of interviews that are interrupted, the lack of coverage of specific collectives (Cea D’Ancona, 2012), the use of answering machines to screen calls, etc.

RESEARCH DESIGN: POST-ELECTORAL STUDY OF THE MADRID COMMUNITY ELECTIONS (CIS STUDY 2893 AND 2894).

Objectives and techniques used.

As previously stated, the purpose of this study is to compare the quality of the responses to the same questionnaire administered on

¹ This research paper is part of a study financed by the Ministry of Economy and Competitiveness, ref. CSO2012-34257. I am grateful for the substantial improvements suggested by the three anonymous reviewers and by the Editorial Board of the REIS. Thanks to Isabel do Campo, Araceli Mateos and Berta Álvarez-Miranda (CIS).

paper (hereinafter, PAPI) and via laptop computer (hereinafter, CAPI) to two similar samples that were selected in the same manner in fieldwork carried out simultaneously by the same interviewers (CIS, 2012).

Matrix questions are a set of linked questions that complement each other as they refer to the same topic and use the same response. Azofra (1999) defines them as questions that use the same formulation to ask about different aspects (Azofra, p. 15) and suggests that they provide fluidity to the questionnaire. They are frequently used in questionnaires due to their ability to offer depth in regards to the topics of study, although their abused use may tire the interviewee² and jeopardize the quality of information due to the risk of “automatically” and unconsciously selecting the same response.

To measure the quality of the response to the questionnaire, come contributions from relevant publications³ were used, in order to evaluate the response based on the following indicators:

- a) Number of *no responses* in the questionnaire (either “don’t know” or “no response”), and the noting of variation based on the instrument used in the collection (PAPEL/CAPI) or in the educational level, a result of the satisficing effect.

This is a topic of extreme importance in the field of survey research, not only due to the loss (or lack of collection) of some very valuable and fundamental information (Tabachnick & Fidell, 1989), but also due to the implications that it has on data analysis (De Leeuw, 1992); especially when working with multi-variate analysis techniques. Many of these –ever more frequently used– suggest application pro-

blems when the questionnaire has a large number of questions without responses. Thus, for example, Jaccard and Wan (1996), consider that in a multiple regression with 15 variables, the size of the sample may be reduced by half if each of the variables used has not been responded to by at least 5% of the interviewees.

- b) Repetition of the same response in various items of a matrix question.
- c) Selection of the easiest responses (extreme and central categories) in questions of attitude with a high number of categories (Zigerell, 2011).

These variable indicators with quantitative metrics require a comparison of the mean values between each sample, thus a hypothesis test (or significance tests) shall be used, specifically, the mean differences, to determine if the inequalities are significant, as has been done in similar studies (Buelens & van den Brakel, 2010; Millar & Dillman, 2011; Schräpler, Schupp & Wagner, 2010). At a second time, regression was used to collectively determine the influence of the socio-demographic variables and the method used in the questionnaire administration. That is, having determined that there are differences, for example, in the number of no responses, this shall be considered a dependent term in the regression. The independent terms are the method used to administer the questionnaire, the sociodemographic and socioeconomic characteristics of the interviewees. That is, gender, age and level of studies completed within the former, and the relation with the activity, professional situation, income and socioeconomic status within the socioeconomic characteristics.

Data sources

The study population includes residents of the Community of Madrid between May 27th and July 1st 2011, who have the right to vote

² “It makes a questionnaire endless”, suggests M.J. Azofra (1999).

³ Among others, Heerwegh, 2009; Schaeffer et al., 2010; Stoop et al., 2010).

in the autonomous elections of the community. The 922 paper questionnaires and the 450 electronic questionnaires are applied based on the typical sample procedure of the CIS.

The questionnaire was made up of 73 questions offering 95 response variables. In addition to the 22 questions including socio-demographic information, there are eight matrix questions offering 36 responses; and these are the subject of this study.

RESULTS: QUALITY IN RESPONSE TO THE MATRIX QUESTION.

The matrix questions offer considerable information but, on many occasions, they generate negative sensations in the interviewees due to the large amount of information that is requested. The eight matrix questions from the questionnaire produce 36 responses-- more than one third of the entire questionnaire (not considering the socio-demographic information). They have considerable variation in response formats, as there are questions with four, five, ten and eleven response categories. The broadness of the response is used as a guide, beginning with the questions having the least options and ending with those having the largest number of response categories.

Having determined that the study samples are equivalent⁴, the analysis of the three matrix questions having the fewest response categories is conducted; specifically, that of

four responses, two regarding attitudes (usefulness and level of agreement) and one on behaviors. We shall begin with the latter which suggests the "*Frequency with which they discussed the elections (with family or friends) during the election campaign*", presenting four response options: *regularly, sometimes, rarely and almost never* (Table 1). Only 18 individuals left the item without a response, resulting in an average of non-responded items of 0.018; this average value decreases to 0.0044 in the CAPI survey, and which increases by five times in the PAPI. The absence of "don't know" responses is explained by the increased response offered to the behavior questions.

The second aspect considered that also presents no differences is that of differentiation, an indicator referring to the variability in the response for each of the items making up the question. For this, the differentiation rate has been calculated for the responses P_d (rho) based on the criteria provided by McCarty and Shrum (2000). High values of the P_d rate suggest that the interviewee has used many response options, little repetition of already given responses, indicating a higher degree of differentiation. The absence of a limit value has led to a modification in this calculation, based upon the transformation proposed by Narayan Annand (2006), which simplifies its interpretation by using a minimum value of 0 and a maximum of 1. Thus, it is considered that a value of 0.341 (Chart 1) reveals a low level of differentiation, which does not vary based on the data collection instrument that is used.

The other two questions having four response categories refer to the influence of the electoral campaign and the degree of agreement with four phrases referring to the concerns of the politicians. In the first question, 18 individuals did not respond (11 "don't know" and 7 "no response" answers), resulting in an average partial no response of 0.020 "don't know" and 0.0167 "no response". Once again, the number of no responses

⁴ Only with equivalent samples is it possible to ensure that the obtained differences (between both sub-samples) are caused by the information collection method that is used.

The equivalence of the samples is demonstrated in http://www.cis.es/cis/opencm/ES/1_encuestas/estudios/resBusqueda.jsp?nEstudioMin=2893&nEstudioMax=2894&mesIni=&anioIni=&mesFin=&anioFin=&muestraIni=&muestraFin=&publicado=3&cualitativo=3&ptitulo=&pcontenido=&ambito=0&universo=0&sexo=0&edad=0&bt nBuscarE=Buscar (access 9th of August of 2014).

TABLE 1. *Quality in the response to matrix questions: mean number of no responses and differentiation rate*

Four category response				
Responses: <i>many, sufficient, few and none</i> (question 11).				
Responses: <i>regularly, sometimes, rarely, never and almost never</i> (question 14).				
Responses: <i>strongly agree, agree, disagree, and strongly disagree</i> (question 28).				
	PAPI	CAPI	Student's T PAPI-CAPI	Average (not considering method)
Quest. 14: <i>Frequency of discussion about the elections during the electoral campaign, 2 items.</i>				
No response				
Don't respond	0.025	0.004	2.83**	0.018
Differentiation	0.341	0.324	0.61	0.336
Quest. 11: <i>Usefulness of diverse aspects of the electoral campaign, 4 items.</i>				
No response				
Don't respond	0.025	0.000	2.59*	0.017
Don't know	0.018	0.022	0.23	0.020
Differentiation	0.309	0.307	0.16	0.309
Quest. 28: <i>Agreement with different phrases about the politicians, 4 items.</i>				
No response				
Don't respond	0.093	0.007	5.43**	0.065
Don't know	0.165	0.127	1.57	0.152
Differentiation	0.603	0.610	0.44	0.606

Five category response				
Responses, in questions 5 and 15: <i>very good, good, regular, bad and very bad.</i>				
In question 27: <i>very near, near, neither near nor far, distant, very distant.</i>				
	PAPI	CAPI	Student's T PAPI-CAPI	Average (not considering method)
Quest. 5: <i>Community of Madrid Government work scale, 8 items.</i>				
No response				
Don't respond	0.083	0.007	3.32**	0.058
Don't know	0.458	0.411	0.84	0.442
Differentiation	0.335	0.348	1.49	0.651
Quest. 15: <i>Assessment of the electoral campaign of the political parties in the autonomous elections of the Community of Madrid in 2011, 4 items (referring to the four parties).</i>				
No response				
Don't respond	0.083	0.002	4.54**	0.057
Don't know	0.946	0.950	0.04	0.947
Differentiation	0.336	0.349	0.306	0.345
Quest. 27: <i>Closeness to political parties of the Community of Madrid, 4 items.</i>				
No response				
Don't respond	0.113	0.046	2.329**	0.091
Don't know	0.117	0.132	0.527	0.122
Differentiation	0.575	0.625	2.517*	0.592

(*) Significant difference at 0.05..

(**) Significant difference at 0,01.

Source: Author's creation with CIS studies 2893 and 2894. N = 1,377.

is greater in the PAPI, while the CAPI shows a slight increase in the number of “don’t know” responses, a difference that is not significant. The differentiation value is again low and barely varies for each data collection method. The trend is quite similar in the third question, with the difference that in this case, the differentiation value increases notably, until reaching a value of 0.6, although it does not vary according to the method used. The number of no responses is very low, although always slightly slower in the CAPI, which at the same time, has a slight greater number of “don’t know” responses.

A regression is used to detect the collective influence of the socio-demographic, socio-economic and collection tool features in the response to these questions. The small number of no responses in the questions having four response categories has led us to *dichotomize* the dependent variable based on whether the interviewee responds –or not– to the variables of this question. This situation recommends the use of a logistics regression, considering the same terms as independent variables and considering the act of responding or not responding as the dependent variable (differentiating between “don’t know” and “no response”). No variable influences the number of “don’t know” responses in the four response category questions. As for the lack of a response (“no response” category), the questionnaire method only influenced in question 14 and 28, adding the family unit income in question 14. The coefficient analysis revealed that the computer-based interviewees (and those with high incomes in question 14) had a lower number of no responses, an improvement in the questionnaire completion.

Questions with five response categories include the interviewee’s opinions regarding the work carried out by the Community of Madrid government in the fields of education, healthcare, housing, etc. (*very good, good, regular, bad and very bad*); an assessment of the electoral campaign of the

most important political parties (*very good, good, regular...*) and finally, their sensation regarding their closeness to these parties (*very near, somewhat, neither near nor far, distant, very distant*). The procedure was identical to the previous question, although an aspect was added to allow for the odd number coding. Specifically, the existence of an intermediate point in the responses (the “regular” assessment in questions 5 and 15– and “neither near nor far” in question 27), a manner of proceeding that is rejected by some experts in the creation of questionnaires since they consider that it includes a large number of responses of individuals who do not wish to think or, if they do have an opinion, do not wish to state it (Converse & Presser, 1986). The calculated measure is the number of selections of the central value as compared to the total values selected, indicating the percentage of responses in the central category, some 30% in question 5 (Table 2).

A panoramic view of the second part of Table 1 reveals a “Don’t respond” rate that is significantly higher in the PAPI survey and an average increase in the differentiation indicator, an increase that is higher in the CAPI method (although only in the last question is there a significant difference). Even more surprising is the high number of interviewees (approximately $\frac{1}{4}$ of the sample) that did not provide a definite response (Table 2), selecting the central option (the “regular” option), the magnitude of which only slightly varied depending on the method used. This is a very significant finding, which we believe should be considered when creating questionnaires, always attempting to avoid a central “indefinite” category.

The regression analysis with the socio-demographic, socio-economic and collection method variables (Table 3) reveal that when considering “Don’t respond” as a dependent term, the administration method influences in the three questions. The magnitude of the coefficients indicates that

TABLE 2. *Percentage of interviewees that select the “easy” categories: central and extreme*

Five category response				
Responses, in questions 5 and 15: <i>very good, good, regular, bad and very bad.</i>				
In question 27: <i>very near, near, neither near nor far, distant, very distant.</i>				
	PAPI	CAPI	Student's T PAPI-CAPI	Average (not considering method)
Quest. 5: Community of Madrid Government work scale, 8 items.				
Central selection category	29.7%	30.6%	0.75	30.0%
Quest. 15: Assessment of the electoral campaign of the political parties in the autonomous elections of the Community of Madrid in 2011, 4 items (referring to the four parties).				
Central selection category	25.1%	27.7%	1.39	26.0%
Quest. 27: Closeness to political parties of the Community of Madrid, 4 items.				
Central selection category	25.7%	23.7%	1.130	25.0%
Ten category response				
Responses: <i>scale of 1 to 10 showed on a card: 1 left and 10 right.</i>				
	PAPI	CAPI	Student's T PAPI-CAPI	Average (not considering method)
Quest. 35: Scale of Community of Madrid ideological and political party self-positioning, four items (referring to the four parties).				
Extreme selection (1 and 10)	10.9%	7.7%	3.53**	9.9%
Central category				
selection (5 and 6)	12.6%	14.4%	2.676**	13.2%
Easy selection categorizes	23.5%	22.2%	1.321	23.1%
Eleven category response				
Responses: <i>scale assessment from 0 to 10, 0 very bad assessment and 10 very good assessment.</i>				
	PAPI	CAPI	Student's T PAPI-CAPI	Average (not considering method)
Quest. 26: Knowledge and assessment scale of national and autonomous political leaders, six items (four autonomous leaders and two national leaders).				
Extreme selection (0 and 10)	25.8%	24.9%	0.533	25.5%
Central category				
selection (5)	13.8%	13.1%	0.560	13.6%
Easy selection categories	39.6%	38.1%	0.880	39.1%

(*) Significant difference at 0,05.

(**) Significant difference at 0,01.

Source: Author's creation with CIS studies 2893 and 2894. N = 1,377.

method used is the most influential term. The “Don't know” response varies depending on gender (men provided few “Don't know” responses in questions 5 and 15), age in question 5 (the older the worse the response), completed studies and profes-

sional situation in questions 15 and 27 (greater studies better response and paid workers, better response), income in question 15, and socioeconomic status in question 27. High income and social status are associated with a better response.

TABLE 3. Influence of sociodemographic, socio-economic and method variables in responding /not responding to questions 5, 15 and 24 (logistics regression)

	Question 5				Question 15				Question 27					
	NR		NS		NR		NS		NR		NS			
	Coef.	E.T.	Coef.	E.T.	Coef.	E.T.	Coef.	E.T.	Coef.	E.T.	Coef.	E.T.		
Collection Method														
(Reference CAPI)	-1.819*	0.74	-0.052	0.142	-2.458 *	1.024	-0.116	0.126	-1.43 **	0.537	-0.12	0.22		
Socio-demographic variable														
Gender: men														
(Reference women)	0.275	0.42	0.259	0.143	-0.199	0.448	0.393 *	0.126	0.305	0.37	-0.016	0.232		
Age	-0.011	0.018	-0.03 **	0.006	-0.033	0.019	-0.002	0.005	0.004	0.015	-0.006	0.009		
Studies	0.187	0.194	0.028	0.058	0.016	0.179	0.161 *	0.052	0.189	0.162	0.289 *	0.102		
Socio-economic variables														
Relation with the activity														
(Reference housewife)														
Workers	-0.768	0.896	-0.042	0.256	0.302	0.785	0.07	0.233	0.041	0.685	0.669	0.369		
Unemployed	-0.901	0.983	-0.229	0.293	-0.242	0.893	-0.036	0.265	-0.627	0.728	0.537	0.423		
Retired	-0.443	0.892	0.295	0.263	0.435	0.739	0.164	0.249	0.282	0.762	0.566	0.365		
Students	-0.887	1.471	-0.582	0.461	-0.393	1.416	0.362	0.409	0.494	1.297	-0.191	0.634		
Professional situation	-0.437	0.756	0.189	0.191	-0.658	0.754	0.454 *	0.173	-0.328	0.621	0.568 *	0.269		
Family income	-0.044	0.091	0.053	0.032	-0.185	0.102	0.116 **	0.029	-0.017	0.081	0.089	0.047		
Socio-economic status	-0.22	0.167	-0.012	0.058	0.032	0.178	-0.014	0.051	-0.032	0.147	-0.24 *	0.089		
Constant	6.974*	1.991	2.04 **	0.574	8.978 **	2.156	-1.087 *	0.506	4.205 *	1.576	1.142	0.863		
-2 log de Verosimilitud	236.272		1378.54		100.894		1656.528		304.77		685.049			
Pseudo R ² (Nagelkerke)	0.079		0.066		0.086		0.071		0.059		0.123			
% casos correctos	98		75.8		98.9		63.5		97.3		91.4			

Notes:.

* Significance at 0,05.

** Significance at 0,01.

Source: Author's creation with CIS studies 2893 and 2894. N = 1,377.

TABLE 4. *Quality in the assessment of response to matrix questions: mean number of questions and differentiation*

Ten category response				
Responses: scale of 1 to 10 showed on a card: 1 left and 10 right				
	PAPI	CAPI	Student's T PAPI-CAPI	Average (not considering method)
Quest. 35: Scale of Community of Madrid ideological and political party self-positioning, four items (referring to the four parties).				
No response				
Don't respond	0.259	0.061	5.32**	0.052
Don't know	0.638	0.637	0.00	0.638
Differentiation	0.912	0.890	1.70*	0.905
Eleven category response				
Responses: scale assessment from 0 to 10, 0 very bad assessment and 10 very good assessment.				
	PAPI	CAPI	Student's T PAPI-CAPI	Average (not considering method)
Quest. 26: Knowledge and assessment scale of national and autonomous political leaders, six items (four autonomous leaders—Esperanza Aguirre, Tomás Gómez, Gregorio Gordo and Luis de Velasco***— and two national leaders: Mariano Rajoy and José Luis Rodríguez Zapatero).				
No response				
Don't respond	0.070	0.015	3.13**	0.052
Don't know	0.428	0.393	0.72	0.417
Differentiation (***)	0.701	0.743	2.92**	0.719
(*)	Significant difference at 0.05.			
(**)	Significant difference at 0.01.			
(***)	Only considers two autonomous leaders (Esperanza Aguirre and Tomás Gómez) and two national leaders (Mariano Rajoy and José Luis Rodríguez Zapatero), due to the large number of interviewees that do not respond to the assessment of Gregorio Gordo and Luis de Velasco since they do not know them (44% and 58%, respectively).			

Source: Author's creation with CIS studies 2893 and 2894. N = 1,377.

In Table 4, the two matrix questions with the greatest number of categories are presented, a scale of ideological placement (1-10) of the four main political parties of the Madrid Community, and a scale (0-10) of assessment of the six national and autonomous political leaders; two and four respectively. The interpretation of these assessment questions with numerous categories reveals that the CAPI survey had a lower

number of no responses and a greater variability of the responses in question having eleven response categories. The regression analysis⁵ reveals a similar relationship to that found in the matrix questions having

⁵ Due to space constraints and for easier reading, they were not shown. The similarity with the coefficients shown in the table 3 has also contributed to the decision not to show these coefficients.

five response categories. The lack of a response had a great relationship with the method of administration of the questionnaire and –to a lesser coefficient– with family income and socioeconomic status in question 35, with no significant relation found in question 26. In question 35, those responding with CAPI and the most elevated layers presented lower levels of no responses. When situating “Don’t know” as a dependent term, gender, family income and level of studies had the greatest relationship for question 35, while question 26 only showed a relationship with age. Men, high incomes and high levels of studies resulted in fewer “Don’t know” responses and the older subjects had greater rates of ignorance.

CONCLUSIONS

By understanding the limits and potential of surveys, it is possible to correctly interpret their results; “understanding the method from its sources of errors” (Newman, 1970). This study has compared the responses obtained from a questionnaire that is administered on paper and via computer in two equivalent samples, considering three quality indicators: average of matrix questions that were not responded to in the questionnaires for each method, differentiating between “Don’t know” and “Don’t respond”, variability of the responses of each interviewee on the scales, and selection of central or extreme categories; that is, the “easy” responses.

Results regarding the lack of a response to the matrix question varied only slightly based on the extensiveness of the scale. There was a greater number of “Don’t know” selections and a significant difference found between the PAPI and CAPI in the failures to respond; simply put, there was an increased lack of respect shown in the PAPI surveys. The differentiation indicator measures the

variability in the responses in each of the items making up the question and has very low values, indicating that the interviewee used few response options (high repetition of previous responses) and only in three comparisons—of the 8 that were carried out—did it vary based on the response measure, with the CAPI having a greater differentiation (in 2 out of 3).

The scales with five response categories included an “indifferent” or “neither one nor the other” response that was selected in 25-30% of all responses. This central position was more “diluted” in the questions having 11 categories since here, the central value was only selected in 13.6% of the responses; This value maintained the same in both collection methods. These responses, which some consider to be a “Don’t know” response, conceal (according to, among others, Lynn & Purdon, 1994) a notable increase in the number of “Don’t know” responses. For the variables of greater amplitude, two specific indicators have been created to include the *extreme scores* (percent of selections of the first and last category) and *easy responses* (selection of extremes and central category); indicators that do not offer consistent results as there are high values in the CAPI in one question and no differences in the other.

To summarize, the majority of the non responses in the questions on opinions and attitudes consisted of “Don’t know”, and are responses that do not vary based on the procedure used, but rather, are related to the level of studies completed, family incomes and gender; as revealed in Table 3. The number of “Don’t know” responses decreases in the interviewees having more studies, higher incomes and in men. It is possible that the study topic itself, on political attitudes and opinions, is responsible for this situation, although perhaps the current climate of social rejection towards politics may also shed light on this finding.

As we do not have indicators such as interview duration or number of incidences per interview, this may be considered one of the major limitations of the study; however, the data records did not provide such information. Upon requesting this information from CIS personnel, it was found that on many occasions, the interviewers were forced to restart the equipment in the middle of the interview, thereby extending the duration of the same. The fact that one of the two studies (no. 2894) is a pilot study on the administration tool suggests that caution should be used when generalizing from the results obtained in this study.

BIBLIOGRAPHY

- Azofra, M. J. (1999). *Cuestionarios*. Madrid: CIS.
- Buelens, Bart and van den Brakel, Jan (2010). "On the Necessity to Include Personal Interviewing in Mixed-mode Survey". *Survey Practice*, October (on line). <http://surveypractice.org/index.php/SurveyPractice/article/view/145>, acceso 16 de junio de 2015.
- Cea D'Ancona, M^a Ángeles (2012). *Fundamentos y aplicaciones en metodología cuantitativa*. Madrid: Síntesis.
- Centro de Investigaciones Sociológicas (2012). *Informe breve estudio piloto implantación CAPI. Postelectoral elecciones autonómicas y municipales 2011. Comunidad de Madrid* (on line). http://www.cis.es/cis/opencms/ES/12_NotasInvestigacion/Investigaciones/2012/NotaInvestigacion0007.html (on line), last access June 16, 2015.
- Converse, Jean M. and Presser, Stanley (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Newbury Park, CA: Sage, Quantitative Research Methods Series, 63.
- Couper, Mick P. (1996). "Changes in Interviewing Setting under CAPI". *Journal of Official Statistics*, 12(3): 301-316.
- De Leeuw, Edith D. (1992). *Data Quality in Mail, Telephone and Face to Face Surveys*. Amsterdam: TT-Publikaties.
- Gordoni, Galit M. et al. (2012). "Measurement Invariance across Face-to-Face and Telephone Modes: The Case of Minority-Status Collectivistic-oriented Groups". *International Journal of Public Opinion Research*, 24: 185-207.
- Heerwegh, Dirk (2009). "Mode Differences between Face-to-face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects". *International Journal of Public Opinion Research*, 21: 111-120.
- Jaccard, James J. and Wan, Choi K. (1996). *LISREL Approaches to Interaction Effects in Multiple Regression*. Thousand Oaks: Sage.
- Keller, Wouter J. (1995). "Changes in Statistical Technology". *Journal of Official Statistics*, 1(1): 115-127.
- Lynn, Peter and Purdon, Susan (1994). "Time Series and Lap-tops: The Change to Computer Assisted Interviewing". In: Jowell, R. et al. *The British Social Attitudes*. Aldershot: Dartmouth.
- Martin, Jean; O'Muircheartaigh, Colm and Curtice, John (1993). "The Use of CAPI for Attitude Surveys: An Experimental Comparison with Traditional Methods". *Journal of Official Statistics*, 9: 641-661.
- Martin, Jean and Manners, Tony (1995). "Computer Assisted Personal Interviewing in Survey Research". In: Lee, R. M. (ed.). *Information Technology for the Social Sciences*. London: UCL Press.
- McCarty, John A. and Shrum, Larry J. (2000). "The Measurement of Personal Values in Survey Research: A Test of Alternative Rating Procedures". *Public Opinion Quarterly*, 64: 271-298.
- Millar, Morgan M. and Dillman, Don A. (2011). *Do Mail and Web Produce Different Answers? Mode Differences in Question Responses and Item Non-Response Rates*. Paper presented at the 66th annual conference of the AAPOR.
- Narayan Anand, Sowmya et al. (2006). *Non-differentiation in Ratings: Tests of Satisficing Theory Predictions*. Ohio: Department of Psychology, Ohio State University,
- Noelle, Elisabeth (1970). *Encuestas en la sociedad de masas*. Madrid: Alianza.
- Schaeffer, Nora Cate et al. (2010). "Interviewers and Interviewing". In: Marsden, P. V. and Wright, J. D. (eds.). *Handbook of Survey Research* (segunda edición). Bingley (Reino Unido): Emerald.
- Schräpler, Jorg-Peter et al. (2010). "Changing from PAPI to CAPI: Introducing CAPI in a Longitudinal Study". *JOSS*, 26, 2: 239-269.

Stoop, Inkele *et al.* (2010). *Improving Survey Response: Lessons Learned from the European Social Survey*. New York: Wiley.

Tabachnick, Barbara G. and Fidell, Linda S. (1989). *Using Multivariate Statistics*. New York: Harper & Publishers.

Zigerell, L. J. (2011). "Midpoint Misperceptions on 1-to-10 Scales". *International Journal of Public Opinion Research*, 23: 205-213.

RECEPTION: February 07, 2014

REVIEW: July 11, 2014

APPROBANCE: April 24, 2015

Calidad de los datos de preguntas de batería en encuestas presenciales: una comparación de un estudio con cuestionario en papel y en formato electrónico

Quality of Data from Matrix Questions in Face-to-face surveys: A Compared Study Using Paper and Electronic Questionnaires

Vidal Díaz de Rada

Palabras clave

- Cuestionarios
- Efectos de respuesta
- Ordenadores
- Recogida de datos

Key words

- Questionnaires
- Response effects
- Computers
- Data Collection

Resumen

En este trabajo se considera la calidad de respuesta de ocho preguntas de batería (36 variables) incluidas en un cuestionario que es administrado en papel y con ordenador en dos muestras equivalentes. Se comparan tres criterios: 1) promedio de preguntas no respondidas en cada modalidad; 2) repetición –o no– de la misma respuesta en varios ítems de una pregunta (diferenciación); y, 3) elección de extremos y punto central (respuestas fáciles). Los cuestionarios en papel proporcionan un mayor número de no respuestas, falta de respuesta que no varía cuando se considera la amplitud de la escala: 4, 5, 10 y 11 categorías. El cuestionario con ordenador presenta una –ligera– mayor diferenciación y unas respuestas menos extremas.

Abstract

This study compares the quality of responses to eight matrix questions (36 variables) included in a survey administered on paper and by computer in two equivalent samples. Three criteria have considered: 1) the average of non-response questions for each mode; 2) the repetition (or not) of the same response in different question items (differentiation); and, 3) the choice of extreme and middle response alternatives on the scale (easy responses). The paper questionnaires provided a greater number of non-response questions, with no change in non-response rates when taking into account scale range: 4, 5, 10 and 11 categories. The electronic questionnaire showed slightly greater differentiation and fewer extreme responses.

Cómo citar

Díaz de Rada, Vidal (2015). «Calidad de los datos de preguntas de batería en encuestas presenciales: una comparación de un estudio con cuestionario en papel y en formato electrónico». *Revista Española de Investigaciones Sociológicas*, 152: 167-178. (<http://dx.doi.org/10.5477/cis/reis.152.167>)

La versión en inglés de este artículo puede consultarse en <http://reis.cis.es>

Vidal Díaz de Rada: Universidad Pública de Navarra | vidal@unavarra.es

INTRODUCCIÓN¹

La encuesta presencial ha dominado las modalidades/modos de recogida de información hasta finales del siglo XX, aunque nuevas innovaciones han provocado un descenso de su utilización en la actualidad. La utilización de ordenadores cada vez más pequeños y potentes ha posibilitado sustituir el papel (del cuestionario) por un ordenador, lo que implica notables mejoras en la información recogida al permitir una mayor concentración del encuestador en la situación de entrevista y una comprobación de la información cuando es recogida.

En este trabajo se compara la *calidad* de información en preguntas de batería en cuestionarios de papel y cuestionarios en ordenador en las investigaciones 2893 y 2894 del Centro de Investigaciones Sociológicas (en adelante CIS). La medición de la calidad se realizará considerando tres indicadores: 1) promedio de preguntas de batería no respondidas en los cuestionarios realizados con cada modalidad diferenciando entre «no sabe» y «no responde»; 2) capacidad de *diferenciación*, esto es, que los entrevistados no seleccionen sistemáticamente la misma categoría de respuesta; y 3) elección de «respuestas fáciles».

ENCUESTA PRESENCIAL ASISTIDA POR ORDENADOR

La encuesta presencial es el modo de recogida de datos mediante el cual un entrevistador administra un cuestionario estructurado (o parcialmente estructurado) a un entrevistado dentro de un limitado período de tiempo y en presencia (generalmente en el hogar) del en-

trevistado (De Leeuw, 1992). La aplicación de avances informáticos a la disciplina desarrolla la entrevista *computerizada* (Keller, 1995) que busca «integrar y automatizar al máximo el proceso de encuesta» (Martin y Manner, 1995). En el siglo XXI los avances del *hardware* se han concretado en el desarrollo de (mini) ordenadores con escaso peso y larga autonomía, que han permitido aplicar el software de la entrevista telefónica computerizada (*Computer Assisted Telephone Interviewing*, en adelante CATI) a la entrevista presencial; sustituyendo el tradicional cuestionario de «papel y lápiz» a un cuestionario «integrado» en un ordenador que permite comprobar y grabar la información a la vez que el cuestionario es respondido (Martin y Manners, 1995).

El CAPI mejora la recogida de datos reduciendo los errores de medición (Cea D'Ancona, 2012) al eliminar la «contaminación» de unas preguntas por otras e integrar el procesamiento de datos en la misma entrevista. Es decir, todas las características del sistema CATI —relacionadas con el proceso de la entrevista— son utilizadas por el equipo informático del CAPI (Couper, 1996). No obstante, la novedad de la entrevista personal asistida por ordenador es que se unen estas ventajas del CATI a la realización de una entrevista personal, eliminando con ello determinados elementos negativos de las encuestas telefónicas, como la imposibilidad de mostrar objetos al entrevistado (Gordoni *et al.*, 2012), gran número de entrevistas interrumpidas, falta de cobertura de determinados colectivos (Cea D'Ancona, 2012), la utilización del contestador automático para filtrar llamadas, etc.

DISEÑO DE LA INVESTIGACIÓN: ESTUDIO POSTELECTORAL DE ELECCIONES EN LA COMUNIDAD DE MADRID (ESTUDIOS CIS 2893 Y 2894)

Objetivos y técnicas utilizadas

Como ya ha sido señalado, el objetivo de este trabajo es comparar la calidad de res-

¹ Este texto es parte de una investigación financiada por el Ministerio de Economía y Competitividad, ref. CSO2012-34257. Agradezco las sustanciales mejoras aportadas por tres evaluadores anónimos y por el Consejo Editorial de la *REIS*. Gracias a Isabel do Campo, Araceli Mateos y Berta Álvarez-Miranda (CIS).

puesta de las preguntas de batería de un mismo cuestionario administrado en papel (en adelante PAPI) y mediante ordenador portátil (en adelante CAPI) a dos muestras similares que han sido seleccionadas de la misma forma y los trabajos de campo se han llevado a cabo simultáneamente por los mismos entrevistadores (CIS, 2012).

Las preguntas de batería pueden definirse como un conjunto de preguntas *encadenadas* que se complementan entre sí al estar referidas a una misma temática y utilizar la misma respuesta. Azofra (1999: 15) las define como preguntas que utilizan la misma formulación para obtener información sobre aspectos diferentes, y señala que dan fluidez al cuestionario. Son muy frecuentes en los cuestionarios por su capacidad para profundizar en las temáticas objeto de estudio, aunque su abuso puede cansar al entrevistado² y perjudicar la calidad de la información por el riesgo de elegir —de forma «automática» y sin pensar— la misma respuesta.

Para la medición de la calidad en la respuesta del cuestionario se seguirán las aportaciones de publicaciones relevantes³ que evalúan la respuesta con los siguientes indicadores:

- a) Número de *no respuestas* en el cuestionario (considerando «no sabe» y «no contesta»), y constatar si varía en función del instrumento empleado en la recogida (PAPEL/CAPI) o del nivel educativo, consecuencia del efecto *satisficing-complacencia*.

Es un tema de gran trascendencia en el ámbito de la investigación con encuestas no solo por la pérdida (o no recogida) de una información en ocasiones muy valiosa y fundamental (Tabachnick y Fidell, 1989), sino también por las implicaciones que ge-

nera en el análisis de datos (De Leeuw, 1992); especialmente cuando se trabaja con técnicas de análisis multivariantes. Muchas de estas —cada vez más utilizadas— plantean problemas de aplicación cuando el cuestionario cuenta con un gran número de preguntas sin respuesta. Así, por ejemplo, Jaccard y Wan (1996) consideran que una regresión múltiple con 15 términos independientes el tamaño muestral puede reducirse a la mitad si cada una de las variables empleadas tiene, al menos, un 5% de entrevistados que no han respondido.

- b) Repetición de la misma respuesta en varios ítems de una pregunta de batería.
- c) Elección de las respuestas más fáciles (categorías extremas y centrales) en preguntas de actitud con elevado número de categorías (Zigerell, 2011).

Estos indicadores, variables con una métrica cuantitativa, precisarían de una comparación de los valores medios entre cada muestra, por lo que será utilizado un test de hipótesis (o test de significación), concretamente las diferencias de medias para determinar si las desigualdades son significativas, tal y como se ha procedido en investigaciones similares (Buelens y van den Brakel, 2010; Millar y Dillman, 2011; Schräpler, Schupp y Wagner, 2010). En un segundo momento se utilizará la regresión para conocer conjuntamente la influencia de las variables sociodemográficas y la modalidad empleada en la administración del cuestionario. Es decir, una vez constatado que hay diferencias en, por ejemplo, el número de no respuestas, este será considerado como término dependiente en la regresión. Los términos independientes son la modalidad utilizada para la administración del cuestionario, las características sociodemográficas y socioeconómicas de los entrevistados. Esto es, sexo, edad y nivel de estudios terminados dentro de las primeras, y la relación con la actividad, situación profesional, ingresos y estatus socioeconómico dentro de las características socioeconómicas.

² «Hace el cuestionario interminable», señala M. J. Azofra (1999).

³ Entre otras, Heerwegh (2009), Schaeffer *et al.* (2010), Stoop *et al.* (2010).

Fuentes de datos

El universo es la población residente en la Comunidad de Madrid entre el 27 de mayo y el 1 de julio del año 2011 con derecho a voto en las elecciones autonómicas de la comunidad. Los 922 cuestionarios de papel y los 450 cuestionarios en formato electrónico se aplicaron siguiendo el procedimiento muestral habitual del CIS.

El cuestionario estaba formado por 73 preguntas que proporcionan 95 respuestas-variables. Además de las 22 preguntas que recogen información sociodemográfica, hay ocho preguntas de batería que proporcionan 36 respuestas, y que son las que constituyen el objetivo de la presente investigación.

RESULTADOS: CALIDAD EN LA RESPUESTA DE PREGUNTAS DE BATERÍA

Las preguntas de batería proporcionan una gran cantidad de información, pero, en numerosas ocasiones, generan sensaciones negativas en los entrevistados por la gran cantidad de información solicitada. Las ocho preguntas de batería del cuestionario producen 36 respuestas; más de un tercio del cuestionario (sin considerar la información sociodemográfica). Éstas presentan gran variación en los formatos de respuesta al contar con preguntas con cuatro, cinco, diez y once categorías de respuesta. La amplitud en la respuesta será utilizada como *guía* de la exposición, comenzando con las preguntas que tienen menos opciones y terminando con las que ofrecen un mayor número de categorías de respuesta.

Una vez constatadas que las muestras objeto de estudio son *equivalentes*⁴ se procede-

rá con el análisis de las tres preguntas de batería que presentan a los entrevistados menos categorías de respuesta; concretamente cuatro respuestas, dos relativas a actitudes (utilidad y nivel de acuerdo) y una a comportamientos. Comenzaremos por la última, que plantea a los entrevistados la *Frecuencia con la que ha hablado de las elecciones (con la familia o los amigos) durante la campaña electoral*, presentando cuatro opciones: *habitualmente, de vez en cuando, rara vez y nunca o casi nunca* (tabla 1). Únicamente 18 personas dejan sin responder algún ítem, con un promedio de ítems no respondidos de 0,018; valor promedio que desciende al 0,0044 en la encuesta CAPI, y se quintuplica en la PAPI. La ausencia de respuestas «no sé» se explica por la mayor respuesta que proporcionan las preguntas de comportamiento.

El segundo de los aspectos considerados, que tampoco presenta diferencias, es la *diferenciación*, indicador referido a la variabilidad en la respuesta en cada uno de los ítems que forman la pregunta. Para ello se ha calculado el índice de diferenciación de respuestas P_d (ρ) siguiendo los criterios de McCarty y Shrum (2000). Valores elevados del índice P_d indican que el entrevistado ha empleado muchas opciones de respuesta, escasa repetición de las respuestas ya proporcionadas, lo que indica una mayor diferenciación. La ausencia de un valor límite nos ha llevado a realizar una modificación de este siguiendo la transformación propuesta por Narayan Annand (2006), que hace más sencilla su interpretación al oscilar entre un valor mínimo de 0 y un máximo de 1. De esta consideración se desprende que el valor 0,341 de la segunda línea del cuadro 1 está mostrando un bajo nivel de diferenciación, que además no varía según la modalidad de recogida empleada.

⁴ Únicamente con muestras equivalentes podrá asegurarse que las diferencias obtenidas (entre ambas submuestras) estarán producidas por la modalidad utilizada en la recogida de información.

La equivalencia de las muestras se puede observar en http://www.cis.es/cis/opencm/ES/1_encuestas/estudios/resBusqueda.jsp?nEstudioMin=2893&nEstudioMax

=2894&mesIni=&anioIni=&mesFin=&anioFin=&muestralni=&muestraFin=&publicado=3&cualitativo=3&ptitulo=&contenido=&ambito=0&universo=0&sexo=0&edad=0&tnBuscarE=Buscar (acceso 9 de agosto de 2014).

TABLA 1. Calidad en la respuesta de preguntas de batería: número medio de no respuestas e índice de diferenciación

Cuatro categorías de respuesta				
Respuestas: <i>mucho, bastante, poco y nada</i> (pregunta 11).				
Respuestas: <i>habitualmente, de vez en cuando, rara vez, nunca y casi nunca</i> (pregunta 14).				
Respuestas: <i>muy de acuerdo, de acuerdo, en desacuerdo y muy en desacuerdo</i> (pregunta 28).				
	PAPI	CAPI	T de Student PAPI-CAPI	Promedio (sin considerar modo)
Preg. 14: Frecuencia con la que ha hablado de las elecciones durante la campaña electoral, 2 ítems.				
No respuesta				
No responde	0,025	0,004	2,83**	0,018
Diferenciación	0,341	0,324	0,61	0,336
Preg. 11: Utilidad de diversos aspectos de la campaña electoral, 4 ítems.				
No respuesta				
No responde	0,025	0,000	2,59*	0,017
No sabe	0,018	0,022	0,23	0,020
Diferenciación	0,309	0,307	0,16	0,309
Preg. 28: Acuerdo con distintas frases sobre la política, 4 ítems.				
No respuesta				
No responde	0,093	0,007	5,43**	0,065
No sabe	0,165	0,127	1,57	0,152
Diferenciación	0,603	0,610	0,44	0,606
Cinco categorías de respuesta				
Respuestas, en preguntas 5 y 15: <i>muy buena, buena, regular, mala y muy mala</i> .				
En pregunta 27: <i>muy cercano, cercano, ni cercano ni distante, distante y muy distante</i> .				
	PAPI	CAPI	T de Student PAPI-CAPI	Promedio (sin considerar modo)
Preg. 5: Calificación labor del gobierno Comunidad de Madrid, 8 ítems.				
No respuesta				
No responde	0,083	0,007	3,32**	0,058
No sabe	0,458	0,411	0,84	0,442
Diferenciación	0,335	0,348	1,49	0,651
Preg. 15: Valoración de la campaña electoral de partidos políticos en las elecciones autonómicas de la Comunidad de Madrid de 2011, 4 ítems (referido a los cuatro partidos).				
No respuesta				
No responde	0,083	0,002	4,54**	0,057
No sabe	0,946	0,950	0,04	0,947
Diferenciación	0,336	0,349	0,306	0,345
Preg. 27: Cercanía hacia partidos políticos de la Comunidad de Madrid, 4 ítems.				
No respuesta				
No responde	0,113	0,046	2,329**	0,091
No sabe	0,117	0,132	0,527	0,122
Diferenciación	0,575	0,625	2,517*	0,592
(*)	Diferencia significativa al 0,05.			
(**)	Diferencia significativa al 0,01.			

Fuente: Elaboración propia con estudios CIS 2893 y 2894. N = 1.377.

Las otras dos preguntas con cuatro categorías de respuesta hacen referencia a la influencia de la campaña electoral y el nivel de acuerdo con cuatro frases referidas a las preocupaciones de los políticos. En la prime-

ra no han respondido 18 personas (11 respuestas «no sabe» y 7 «no respuestas»), que supone un promedio de no respuesta parcial de 0,020 «no sabe» y 0,0167 «no respuestas». El número de no respuestas, de nuevo,

es superior en el PAPI, mientras que el CAPI presenta un ligero mayor número de «no sabe», diferencia que no llega a ser significativa. El valor de diferenciación vuelve a ser bajo, y apenas cambia según la modalidad de recogida. La tendencia es muy similar en la tercera pregunta, con la diferencia de que en este caso el valor de diferenciación aumenta notablemente, hasta alcanzar el valor 0,6, que no varía según la modalidad utilizada. El número de no respuestas es muy bajo, aunque siempre ligeramente inferior en CAPI, al tiempo que presenta un ligero mayor número de «no sabe».

Se empleará una regresión para detectar la influencia conjunta de los rasgos sociodemográficos, socioeconómicos y la modalidad de respuesta. El escaso número de no respuestas en las preguntas con cuatro categorías precisa dicotomizar la variable dependiente considerando si el entrevistado responde –o no– a las variables de esta pregunta. Esta situación recomienda el empleo de una regresión logística, considerando como variables independientes los rasgos sociodemográficos, socioeconómicos y la modalidad de respuesta y como dependiente el hecho de responder/no responder (diferenciando entre «no sabe» y «no responde»). Ninguna variable influye en el número de «no sabes» de las preguntas con cuatro categorías de respuesta. Respecto a la falta de respuesta (categoría «no respuesta») solo la modalidad de cuestionario influye en las preguntas 14 y 28, añadiéndose los ingresos de la unidad familiar en la pregunta 14. El análisis de los coeficientes desvela que los entrevistados mediante ordenador (y con altos ingresos en la pregunta 14) tienen un menor número de no respuestas, una mejora en la cumplimentación del cuestionario.

Las preguntas con cinco categorías de respuesta recogen las opiniones de los entrevistados sobre la labor realizada por el gobierno de la Comunidad de Madrid en educación, sanidad, vivienda, etc. (*muy buena, buena, regular, mala y muy mala*); una

valoración de la campaña electoral de los partidos políticos más importantes (*muy buena, buena, regular...*) y, por último, su sensación sobre la cercanía hacia esos partidos (*muy cercano, bastante, ni cercano ni distante, distante y muy distante*). Se procederá exactamente igual que en la pregunta anterior, aunque añadiendo un aspecto que permite la codificación impar. Concretamente la existencia de un punto intermedio en las respuestas (valoración «regular» —en preguntas 5 y 15— y «ni cercano ni distante» en la pregunta 27), una forma de proceder rechazada por algunos expertos sobre elaboración de cuestionarios porque consideran que recogen una gran parte de respuestas de personas que no desean pensar o, cuando tienen opinión, no desean decirla (Converse y Presser, 1986). La medida calculada es el número de elecciones del valor central respecto al total de valores elegidos, lo que indica el porcentaje de respuestas en la categoría central, un 30% en la pregunta 5 (tabla 2).

Una visión panorámica de la segunda parte de la tabla 1 desvela una tasa de «no responde» significativamente superior en la encuesta PAPI y un aumento promedio del indicador de diferenciación, aumento que es superior en la modalidad CAPI (aunque solo en la última pregunta hay diferencia significativa). Más sorprendente es el elevado número de entrevistados (aproximadamente un cuarto de la muestra) que no han proporcionado una respuesta definida (tabla 2), decantándose por la opción central (opción «regular», magnitud que apenas varía según modalidades. Es un hallazgo importante que, consideramos, debe tenerse en cuenta cuando se redactan cuestionarios, tratando siempre de evitar una categoría central «indefinida».

El análisis de regresión con las variables sociodemográficas, socioeconómicas y la modalidad de recogida (tabla 3) desvela que cuando se considera «no responde» como término dependiente la modalidad de administración influye en las tres preguntas. La

TABLA 2. *Porcentaje de entrevistados que eligen categorías «fáciles»: centrales y extremas*

Cinco categorías de respuesta				
Respuestas, en preguntas 5 y 15: <i>muy buena, buena, regular, mala y muy mala.</i>				
En pregunta 27: <i>muy cercano, cercano, ni cercano ni distante, distante y muy distante.</i>				
	PAPI	CAPI	T de Student PAPI-CAPI	Promedio (sin considerar modo)
Preg. 5: <i>Calificación labor del gobierno Comunidad de Madrid, 8 ítems.</i>				
Elección categoría central	29,7	30,6	0,75	30,0
Preg. 15: <i>Valoración de la campaña electoral de partidos políticos en las elecciones autonómicas de la Comunidad de Madrid de 2011, 4 ítems (referido a los cuatro partidos).</i>				
Elección categoría central	25,1	27,7	1,39	26,0
Preg. 27: <i>Cercanía hacia partidos políticos de la Comunidad de Madrid, 4 ítems.</i>				
Elección categoría central	25,7	23,7	1,130	25,0
Diez categorías de respuesta				
Respuestas: escala de 1 a 10 mostrada en una tarjeta: 1 izquierda y 10 derecha.				
	PAPI	CAPI	T de Student PAPI-CAPI	Promedio (sin considerar modo)
Preg. 35: <i>Escala de autoubicación ideológica y de partidos políticos de la Comunidad de Madrid, 4 ítems (referido a los cuatro partidos).</i>				
Elección extremos (1 y 10)	10,9	7,7	3,53**	9,9
Elección categorías centrales (5 y 6)	12,6	14,4	2,676**	13,2
Elección categorías fáciles	23,5	22,2	1,321	23,1
Once categorías de respuesta				
Respuestas: escala valoración de 0 a 10, 0 muy mala valoración y 10 valoración muy buena.				
	PAPI	CAPI	T de Student PAPI-CAPI	Promedio (sin considerar modo)
Preg. 26: <i>Conocimiento y escala de valoración de líderes políticos nacionales y autonómicos, 6 ítems (cuatro líderes autonómicos y dos nacionales).</i>				
Elección extremos (0 y 10)	25,8	24,9	0,533	25,5
Elección categorías centrales (5)	13,8	13,1	0,560	13,6
Elección categorías fáciles	39,6	38,1	0,880	39,1

(*) Diferencia significativa al 0,05.

(**) Diferencia significativa al 0,01.

Fuente: Elaboración propia con estudios CIS 2893 y 2894. N = 1.377.

magnitud de los coeficientes indica que es la modalidad empleada el término que presenta mayor influencia. El «no sabe», por su parte, varía según el sexo (los hombres expresan pocos «no sabe» en las preguntas 5 y 15), edad en pregunta 5 (a más edad peor respuesta), estudios terminados y situación profesional en preguntas 15 y 27 (estudios altos mejor respuesta, asalariados mejor respuesta), ingresos en pregunta 15, y esta-

tus socioeconómico en pregunta 27. Los altos ingresos y el estatus alto están asociados a una mejor respuesta.

En la tabla 4 se presentan las dos preguntas de batería con más categorías, una escala de ubicación ideológica (1-10) de los cuatro principales partidos políticos de la Comunidad de Madrid, y una *escala (0-10) de valoración de seis líderes políticos nacionales y autonómicos*; dos y cuatro respectivamente. La

TABLA 3. Influencia de variables sociodemográficas, socioeconómicas y modalidad en responder/no responder las preguntas 5, 15 y 24 (regresión logística)

	Preg 5: Calificación de la labor del gobierno Comunidad de Madrid		Preg. 15: valoración de la campaña electoral		Preg. 27: cercanía hacia partidos políticos de la Com.							
	No responde	No sabe	No responde	No sabe	No responde	No sabe						
	Coef.	E.T.	Coef.	E.T.	Coef.	E.T.						
Modalidad de recogida (Referencia CAPI)	-1,819*	0,74	-0,052	0,142	-2,458*	1,024	-0,116	0,126	-1,43**	0,537	-0,12	0,22
Variables sociodemográficas												
Sexo: varones												
(Referencia mujeres)	0,275	0,42	0,259	0,143	-0,199	0,448	0,393*	0,126	0,305	0,37	-0,016	0,232
Edad	-0,011	0,018	-0,03**	0,006	-0,033	0,019	-0,002	0,005	0,004	0,015	-0,006	0,009
Estudios	0,187	0,194	0,028	0,058	0,016	0,179	0,161*	0,052	0,189	0,162	0,289*	0,102
Variables socioeconómicas												
Relación con la actividad (Referencia ama de casa)												
Ocupados	-0,768	0,896	-0,042	0,256	0,302	0,785	0,07	0,233	0,041	0,685	0,669	0,369
Parados	-0,901	0,983	-0,229	0,293	-0,242	0,893	-0,036	0,265	-0,627	0,728	0,537	0,423
Jubilados	-0,443	0,892	0,295	0,263	0,435	0,739	0,164	0,249	0,282	0,762	0,566	0,365
Estudiantes	-0,887	1,471	-0,582	0,461	-0,393	1,416	0,362	0,409	0,494	1,297	-0,191	0,634
Situac. profesional	-0,437	0,756	0,189	0,191	-0,658	0,754	0,454*	0,173	-0,328	0,621	0,568*	0,269
Ingreso familiar	-0,044	0,091	0,053	0,032	-0,185	0,102	0,116**	0,029	-0,017	0,081	0,089	0,047
Estatus socioeconómico	-0,22	0,167	-0,012	0,058	0,032	0,178	-0,014	0,051	-0,032	0,147	-0,24*	0,089
Constante	6,974*	1,991	2,04**	0,574	8,978**	2,156	-1,087*	0,506	4,205*	1,576	1,142	0,863
-2 log de Verosimilitud	236,272		1378,54		100,894		1656,528		304,77		685,049	
Pseudo R ² (Nagelkerke)	0,079		0,066		0,086		0,071		0,059		0,123	
% casos correctos	98		75,8		98,9		63,5		97,3		91,4	

Notas:

* Significación al 0,05.

** Significación al 0,01.

Fuente: Elaboración propia con estudios CIS 2893 y 2894. N = 1.377.

TABLA 4. *Calidad en la respuesta de preguntas de batería de valoración: número medio de preguntas y diferenciación*

Diez categorías de respuesta				
Respuestas: escala de 1 a 10 mostrada en una tarjeta: 1 izquierda y 10 derecha.				
	PAPI	CAPI	T de Student PAPI-CAPI	Promedio (sin considerar modo)
Preg. 35: Escala de ubicación ideológica de los partidos políticos de la Comunidad de Madrid, 4 ítems (referido a los cuatro partidos).				
No respuesta				
No responde	0,259	0,061	5,32**	0,052
No sabe	0,638	0,637	0,00	0,638
Diferenciación	0,912	0,890	1,70*	0,905
Once categorías de respuesta				
Respuestas: escala valoración de 0 a 10, 0 muy mala valoración y 10 valoración muy buena.				
	PAPI	CAPI	T de Student PAPI-CAPI	Promedio (sin considerar modo)
Preg. 26: Conocimiento y escala de valoración de líderes políticos nacionales y autonómicos, 6 ítems (cuatro líderes autonómicos –Esperanza Aguirre, Tomás Gómez, Gregorio Gordo y Luis de Velasco***– y dos nacionales: Mariano Rajoy y José Luis Rodríguez Zapatero).				
No respuesta				
No responde	0,070	0,015	3,13**	0,052
No sabe	0,428	0,393	0,72	0,417
Diferenciación (***)	0,701	0,743	2,92**	0,719
* Diferencia significativa al 0,05.				
** Diferencia significativa al 0,01.				
*** Considera únicamente dos líderes autonómicos (Esperanza Aguirre y Tomás Gómez) y los dos nacionales (Mariano Rajoy y José Luis Rodríguez Zapatero), por el gran número de entrevistados que no responden la valoración de Gregorio Gordo y Luis de Velasco por no conocerlos (un 44% y un 58% respectivamente).				

Fuente: Elaboración propia con estudios CIS 2893 y 2894. N = 1.377.

interpretación de estas preguntas de valoración con numerosas categorías muestra que la encuesta CAPI logra menor número de no respuestas y una mayor variabilidad en las respuestas en la pregunta con once categorías de respuesta. El análisis de regresión⁵ muestra una relación similar a la detectada en las baterías con cinco categorías de respuesta. La no respuesta presenta una gran relación con la modalidad de administración del cues-

tionario y —con un menor coeficiente— con los ingresos familiares y el estatus socioeconómico en la pregunta 35, no existiendo ninguna relación significativa en la pregunta 26. En la pregunta 35 los que responden con CAPI y los estratos más elevados presentan menos no respuestas. Cuando se coloca «no sabe» como término dependiente, el sexo, ingresos familiares y nivel de estudios terminados presentan la mayor relación en la pregunta 35, mientras que la pregunta 26 solo muestra relación con la edad. Los hombres, altos ingresos y altos estudios proporcionan menos «no sabe» en la pregunta 35, y los de edad avanzada presentan mayores tasas de desconocimiento en la pregunta 26.

⁵ No se muestra por motivos de espacio y para hacer más ligera la lectura. La similitud con los coeficientes mostrados en la tabla 3 también ha contribuido en la decisión de no mostrar los coeficientes.

CONCLUSIONES

Solo conociendo los límites y potencialidades de la encuesta será posible realizar una correcta interpretación de los resultados proporcionados por esta herramienta; «entendiendo el método desde su fuente de error» (Newman, 1970). En este trabajo se han comparado las respuestas obtenidas por un cuestionario que es administrado en papel y con ordenador en dos muestras equivalentes, considerando tres indicadores de calidad: promedio de preguntas de batería no respondidas en los cuestionarios realizados con cada modalidad diferenciando entre «no sabe» y «no responde», variabilidad de las respuestas de cada entrevistado en las escalas, y elección de categorías centrales o extremas; esto es, respuestas «fáciles».

Los resultados en la falta de respuesta en preguntas de batería apenas varían en función de la amplitud de la escala. Se concreta en un mayor número de elecciones «no sabe» y diferencia significativa entre PAPI y CAPI en la manifestación de no responder; presentando siempre PAPI una mayor falta de respuesta. El indicador de diferenciación mide la variabilidad en la respuesta en cada uno de los ítems que forman la pregunta, y presenta valores bajos, que indican que el entrevistado ha empleado pocas opciones de respuesta (gran repetición de las respuestas anteriores) y solo en tres comparaciones —de las 8 realizadas— varía según la modalidad de respuesta, presentando el CAPI una mayor diferenciación (en 2 de 3).

Las escalas con cinco categorías de respuesta cuentan con una categoría central «indiferente» o «ni uno ni otro», que recoge entre un 25-30% de las respuestas. Esta posición central queda más «diluida» en las preguntas con 11 categorías, ya que el valor central únicamente recibe un 13,6% de las respuestas; valor que apenas cambia en función de la modalidad. Estas respuestas, que algunos consideran como un «no

sabe» encubierto (entre otros, Lynn y Purdon, 1994), suponen un notable aumento del número de respuestas «no sabe». Para las variables de mayor amplitud se han elaborado dos indicadores específicos que recogen las *puntuaciones extremas* (porcentaje de elecciones de la primera y última categoría) y las *respuestas fáciles* (elección de extremos y de categoría central); indicadores que no proporcionan resultados consistentes al presentar altos valores del CAPI en una pregunta y ausencia de diferencias en la otra.

Terminar señalando que la mayor parte de la no respuesta en las preguntas de opiniones y actitudes se ha producido por «no sabe», y son respuestas que no varían en función de la modalidad utilizada, sino que están relacionadas con estudios terminados, situación profesional, ingresos familiares y sexo; tal y como se mostró en la tabla 3. El número de respuestas «no sabe» desciende en los entrevistados con estudios más elevados, altos ingresos y varones. Es posible que la propia temática de la encuesta, actitudes y opiniones políticas, sea responsable de esta situación; aunque quizás el clima de rechazo social hacia la política pudiera arrojar luz sobre esta realidad.

No disponer de indicadores como la duración de las entrevistas, o el número de incidencias por entrevista conseguida, es una de las grandes limitaciones del trabajo, pero el archivo de datos no proporciona tales informaciones. Solicitada esta información al personal del CIS, destacaron que la poca capacidad de la batería de los equipos informáticos generó que —en numerosas ocasiones— los entrevistadores se vieron obligados a reiniciarlos en mitad de la entrevista, alargado con ello la duración de la entrevista. El hecho de que uno de los dos estudios (nº 2894) sea un estudio piloto sobre la modalidad de administración lleva a recomendar cautela a la hora de generalizar los resultados obtenidos en este trabajo.

BIBLIOGRAFÍA

- Azofra, M. J. (1999). *Cuestionarios*. Madrid: CIS.
- Buelens, Bart y van den Brakel, Jan (2010). «On the Necessity to Include Personal Interviewing in Mixed-mode Survey». *Survey Practice*, October (en línea). <http://surveypractice.org/index.php/SurveyPractice/article/view/145>, acceso 16 de junio de 2015.
- Cea D'Ancona, M^a Ángeles (2012). *Fundamentos y aplicaciones en metodología cuantitativa*. Madrid: Síntesis.
- Centro de Investigaciones Sociológicas (2012). *Informe breve estudio piloto implantación CAPI. Postelectoral elecciones autonómicas y municipales 2011. Comunidad de Madrid* (en línea). http://www.cis.es/cis/opencms/ES/12_NotasInvestigacion/Investigaciones/2012/NotaInvestigacion0007.html (en línea), acceso 16 de junio de 2015.
- Converse, Jean M. y Presser, Stanley (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Newbury Park, CA: Sage, Quantitative Research Methods Series, 63.
- Couper, Mick P. (1996). «Changes in Interviewing Setting under CAPI». *Journal of Official Statistics*, 12(3): 301-316.
- De Leeuw, Edith D. (1992). *Data Quality in Mail, Telephone and Face to Face Surveys*. Amsterdam: TT-Publikaties.
- Gordoni, Galit M. et al. (2012). «Measurement Invariance across Face-to-Face and Telephone Modes: The Case of Minority-Status Collectivistic-oriented Groups». *International Journal of Public Opinion Research*, 24: 185-207.
- Heerwegh, Dirk (2009). «Mode Differences between Face-to-face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects». *International Journal of Public Opinion Research*, 21: 111-120.
- Jaccard, James J. y Wan, Choi K. (1996). *LISREL Approaches to Interaction Effects in Multiple Regression*. Thousand Oaks: Sage.
- Keller, Wouter J. (1995). «Changes in Statistical Technology». *Journal of Official Statistics*, 1(1): 115-127.
- Lynn, Peter y Purdon, Susan (1994). «Time Series and Lap-tops: The Change to Computer Assisted Interviewing». En: Jowell, R. et al. *The British Social Attitudes*. Aldershot: Dartmouth.
- Martin, Jean; O'Muircheartaigh, Colm y Curtice, John (1993). «The Use of CAPI for Attitude Surveys: An Experimental Comparison with Traditional Methods». *Journal of Official Statistics*, 9: 641-661.
- Martin, Jean y Manners, Tony (1995). «Computer Assisted Personal Interviewing in Survey Research». En: Lee, R. M. (ed.). *Information Technology for the Social Sciences*. London: UCL Press.
- McCarty, John A. y Shrum, Larry J. (2000). «The Measurement of Personal Values in Survey Research: A Test of Alternative Rating Procedures». *Public Opinion Quarterly*, 64: 271-298.
- Millar, Morgan M. y Dillman, Don A. (2011). *Do Mail and Web Produce Different Answers? Mode Differences in Question Responses and Item Non-Response Rates*. Paper presented at the 66th annual conference of the AAPOR.
- Narayan Anand, Sowmya et al. (2006). *Non-differentiation in Ratings: Tests of Satisficing Theory Predictions*. Ohio: Department of Psychology, Ohio State University.
- Noelle, Elisabeth (1970). *Encuestas en la sociedad de masas*. Madrid: Alianza.
- Schaeffer, Nora Cate et al. (2010). «Interviewers and Interviewing». En: Marsden, P. V. y Wright, J. D. (eds.). *Handbook of Survey Research* (segunda edición). Bingley (Reino Unido): Emerald.
- Schräpler, Jorg-Peter et al. (2010). «Changing from PAPI to CAPI: Introducing CAPI in a Longitudinal Study». *JOSS*, 26, 2: 239-269.
- Stoop, Inkele et al. (2010). *Improving Survey Response: Lessons Learned from the European Social Survey*. New York: Wiley.
- Tabachnick, Barbara G. y Fidell, Linda S. (1989). *Using Multivariate Statistics*. New York: Harper & Publishers.
- Zigerell, L. J. (2011). «Midpoint Misperceptions on 1-to-10 Scales». *International Journal of Public Opinion Research*, 23: 205-213.

RECEPCIÓN: 07/02/2014

REVISIÓN: 11/07/2014

APROBACIÓN: 24/04/2015